The American Journal of Human Genetics, Volume *108*

Supplemental Data

De novo TRIM8 variants impair its protein localization to nuclear bodies and cause developmental delay, epilepsy, and focal segmental glomerulosclerosis

Patricia L. Weng, Amar J. Majmundar, Kamal Khan, Tze Y. Lim, Shirlee Shril, Gina Jin, John Musgrove, Minxian Wang, Dina F. Ahram, Vimla S. Aggarwal, Louise E. Bier, Erin L. Heinzen, Ana C. Onuchic-Whitford, Nina Mann, Florian Buerger, Ronen Schneider, Konstantin Deutsch, Thomas M. Kitzler, Verena Klämbt, Amy Kolb, Youying Mao, Christelle Moufawad El Achkar, Adele Mitrotti, Jeremiah Martino, Bodo B. Beck, Janine Altmüller, Marcus R. Benz, Shoji Yano, Mohamad A. Mikati, Talha Gunduz, Heidi Cope, Vandana Shashi, Undiagnosed Diseases Network, Howard Trachtman, Monica Bodria, Gianluca Caridi, Isabella Pisani, Enrico Fiaccadori, Asmaa S. AbuMaziad, Julian A. Martinez-Agosto, Ora Yadin, Jonathan Zuckerman, Arang Kim, UCLA Clinical Genomics Center, Ulrike John-Kroegel, Amanda V. Tyndall, Jillian S. Parboosingh, A. Micheil Innes, Agnieszka Bierzynska, Ania B. Koziell, Mordi Muorah, Moin A. Saleem, Julia Hoefele, Korbinian M. Riedhammer, Ali G. Gharavi, Vaidehi Jobanputra, Emma Pierce-Hoffman, Eleanor G. Seaby, Anne O'Donnell-Luria, Heidi L. Rehm, Shrikant Mane, Vivette D. D'Agati, Martin R. Pollak, Gian Marco Ghiggeri, Richard P. Lifton, David B. Goldstein, Erica E. Davis, Friedhelm Hildebrandt, and Simone Sanna-Cherchi

Table S1. Clinical features of 12 families with dominant truncating *TRIM8* variants

Family	Renal Disease	Neurologic Disease	Other Organ Involvement
S1666	Onset: 2.2 years. Presentation: Nephrotic Syndrome. Chemistries at onset: Serum Albumin < 15 g/dL; Crea <100 umol/L. Urinalysis: UPC >2.5 g/g Crea. Renal Biopsy: FSGS. Glomeruli showed global sclerosis (majority), segmental sclerosis with hyalinization and capsular reaction (>4), or mesangial matrix expansion. Extensive tubular atrophy with moderate interstitial infiltration by chronic inflammatory cells. Interstitial fibrosis with dilated tubules, some with protein droplets. Thick walled vessels. Therapy: Resistant to corticosteroids. CKD/ESRD: ESRD onset at 3 years. Tx: Age 7.6 years. NRP at 10 days post-Tx. Plasma exchange for 3 months. Tx rejection at >8 years but not FSGS recurrence by biopsy.	Seizures: Onset 4.5 years. Generalized tonic-clonic seizures. DD: Onset 2.2 years. Cognitive and motor developmental delays. Brain Imaging: MRI showed slightly reduced cerebral and cerebellar volume. No evidence of a space occupying lesion. EEG: Benign focal epilepsy left central temporal region and occasionally on right.	Cardiac: Echocardiogram normal. Endocrine: Male pattern hirsutism. Ovarian US suggestive of polycystic ovarian syndrome at age 16. Other organ involvement: Mildly elevated bladder pressures during pre-transplant evaluation, leading to ureterostomy and subsequent bladder augmentation and Mitrofanoff.
F827	Onset: 4.5 years. Presentation: Nephrotic Syndrome. Chemistries at onset: Albumin 1.3 g/dL, Crea 0.78 mg/dL (eGFR 77 mL/min/1.73m^2), complement C3 and C4 normal. Urinalysis: UPC 24 g/g Crea, microscopic glomerular hematuria. Renal US: Bilaterally enlarged echogenic kidneys (97th percentile) with no corticomedullary differentiation. Renal Biopsy: Focally global and segmental glomerulosclerosis Tubular atrophy and interstitial fibrosis (50% of the cortical tubulointerstitium). Therapy: Resistant to corticosteroids. CKD/ESRD: ESRD at 4.8 years. Transplant: Age 5 years with no rejection episodes through age 24 years.	Seizures: Onset 4.5 years. Generalized seizures. DD: Onset 1 year. Motor delay first. Hypotonia at age 3 years. Progressive worsening intellectual disability and speech disorder. Brain Imaging: MRI at 4.5 years showed cortical frontotemporal atrophy. EEG: Central left accentuated seizure pattern	Facial Dysmorphisms: Notable for broad forehead, plagiocephaly, high arched palate, low set ears. Ophthalmologic: Hyperopia, astigmatism. ENT: Sensorineural hearing loss, age 21 years. Cardiac: Echocardiogram normal. Gastrointestinal: Pilonidal abscess, age 21 years. Post-Tx diarrhea. Dermatologic: Seborrheic eczema, age 16 years. MSK: Carpo-metacarpal growth dissociation, scoliosis, hallux valgus. GU: Inguinal testicular hernia.

A4582	**Onset: Birth** **Presentation: Nephrotic Syndrome.** **Chemistries at onset: Increased Crea.** **Urinalysis: Nephrotic Range Proteinuria.** **Renal US: Small, atrophic hyperechogenic kidneys with no cysts or obstruction at 2.75 years.** **Renal Biopsy: DMS at 1.5 years.** **Therapy: Resistant to corticosteroids. Other therapies such as MMF, tacrolimus, CsA were not tried as rapid progression to ESRD.** **CKD/ESRD: ESRD age 1.1 years.** **Tx: Cadaveric transplant at 4.9 years. No rejection to age 14 years (when passed away).**	**Seizures: Onset age 2 years that were recurrent but did respond to levetiracetam (Keppra).** **DD: Onset <1 year as unable to sit, stand, or walk. Developed psychomotor retardation, cerebral palsy, and hypotonia.** **Brain Imaging: MRI showed Cerebral atrophy with small mid-brain structures and normal cerebellum at age 2.7 years.**	**Cardiac: Atrial septal defect status post spontaneous closure.** **Pulmonary: Recurrent pneumonia from age 13 years.** **Gastrointestinal: Gastrostomy tube for feeding and medications since age 5 years.** **MSK: Scoliosis, osteopenia, and pathologic femur fracture secondary to immobilization.** **Endocrine: Precocious puberty.** **Oncology: Osteoblastic osteosarcoma of right scapula, for which passed away at 14 years of age.**
FSGSGE126	**Onset: 13.7 years** **Presentation: Nephrotic Syndrome.** **Chemistries at onset: normal Crea.** **Urinalysis: Nephrotic Range Proteinuria.** **Renal US at presentation: normal.** **Renal Biopsy: FSGS at 16.7 years.** **Therapy: Resistant to corticosteroids and CNI.** **CKD/ESRD: ESRD age 19.7 years.** **Tx: Renal transplant at 20 years, no recurrence to date. AMR Parvovirus B19 interstitial nephritis at age 22 years. Currently alive.**	**Seizures: Onset age 1.5 years resistant to levetiracetam (Keppra).** **DD: Onset 1.5 years, developed spastic tetraparesis with tonic clonic seizures, aggressive behavior.** **Brain Imaging: MRI showed cerebral atrophy with posterior fossa dilation at age 15.**	**Gastrointestinal: feeding difficulties**
UC-023-1	**Onset: 4.9 years** **Presentation: Nephrotic syndrome** **Chemistries at onset: Albumin 1.5 g/dL, Crea 0.70 mg/dL (eGFR 86 mL/min/1.73m^2), complement C3 and C4 normal.** **Urinalysis: UPC 16 g/g Cr, microscopic glomerular hematuria.** **- Renal biopsy findings: FSGS. extensive, but subtotal foot process effacement by electron microscopy. 20% glomeruli completely sclerosed, non-sclerotic glomeruli showed mild mesangial expansion mild patchy tubulointerstitial scarring** **- Initial therapy: resistant to corticosteroids. Treated with ACE inhibitor (enalapril). Other therapies not tried due to rapid progression to ESRD.** **- CKD/ESRD: ESRD at age 5.3 years** **- Renal transplantation: 6.3 years**	**Seizure: onset age 2.8 years. Focal tonic with secondary generalization. Seizures resistant to lamotrigine and topiramate but responsive to levetiracetam.** **Neurologic features** **DD: Onset 5 months rolled over at 5 months, sitting at 14 months, crawling at 18 months, pulling to stand at 2 years and >2 years for walking. Currently non-verbal and has limited understanding of speech but is able to follow certain simple commands. Hypotonia and psychomotor retardation.** **Brain Imaging: MRI at 2.2 years with mild cerebral atrophy** **EEG: mild, diffuse slowing of the background. Abundant, frequently near**	**Dysmorphisms: broad nasal bridge, bilateral low set ears, widespread nipples** **MSK: bilateral 5th toe clinodactyly, mild clubbing bilaterally, bilateral hand tremors** **Dermatology: mild hirsutism to upper lip** **Gastrointestinal: Gastrostomy tube since 6.8 years for hydration and medications**

		continuous independent and synchronous left centrotemporal to right centrotemporal spike-wave discharges, occasionally with rolandic morphology	
HN-F65	**Onset: 7.9 years.** **Presentation: Hypertension, Proteinuria.** **Chemistries at onset: Crea 0.27 mg/dl (normal), BUN 24 mg/dl (elevated), Elevated triglycerides and cholesterol levels, Albumin 2.8 g/dL.** **Urinalysis: UPC 10 g/g crea, Microscopic glomerular hematuria.** **Renal Biopsy: FSGS (4/36 glomeruli) and slight global glomerulosclerosis (10/36 glomeruli). 15-20% tubular atrophy/interstitial fibrosis.** **Therapy: Resistant to corticosteroids. Treated with ACE inhibitor (ramipril) since diagnosis. Slight improvement of proteinuria with CsA.** **CKD/ESRD: Continuous decrease of eGFR, last at CKD stage 3 as 9.8 years.** **Transplant: No**	**Seizures: Onset at age 2.5 years. Primary, atypical absences and atonic seizures that were treated through age 6 years with anti-epileptic medications. Off therapy, developed recurrence of seizures at age 7 years and medications were resumed.** **DD: Onset at 2.5 years. Expressive speech delay, autism-spectrum disorder, aggressive behavior, hyperactivity, and encephalopathy.** **HC: Secondary microcephaly (50 cm at 7.9 years).** **Brain Imaging: MRI at 2.5 years was normal.**	**Gastrointestinal: Esophagitis and Duodenitis at age 8 years.**
B1117	**Onset: 2.5 years.** **Presentation: Nephrotic Syndrome.** **Renal Biopsy: FSGS.** **CKD/ESRD: ESRD at >5 years.**	**Seizures: Onset at 2.5 years.** **DD: Developmental Delay before age 1.**	
B3883	**Onset: 6 years.** **Presentation: Asymptomatic nephrotic range proteinuria.** **Labs: Albumin 2.8 g/dL, Crea 0.45 mg/dL,** **Urinalysis: UPC 6 g protein/g Crea.** **Renal Biopsy: FSGS (7/42 glomeruli) and global glomerulosclerosis (20/42 glomeruli). 20% tubular atrophy/interstitial fibrosis. EM shows podocyte foot process effacement and microvillous degeneration of the posterior apical surfaces. In some cases, prominent vacuolization of cytoplasm. Thin basement membranes focally.** **Therapy: ACE inhibitor initiated at age 7 years.** **CKD/ESRD: CKD/ESRD: Crea 0.42 mg/dL at age 7, consistent with CKD stage 2.** **Transplant: No.**	**Seizures: Onset at 1.5 years with focal pattern. Responsive to valproic acid and cannabidiol.** **DD: Onset 0.5 years. Global developmental delay. Encephalopathy. Hypotonia.** **Brain Imaging: MRI at age 6 showed supratentorial, cerebellar and brainstem volume loss.**	**MSK: Contractures.**
FG-FA **(affected monozygotic twins)**	**Onset: 6 years (both twins)** **Presentation: nephrotic range proteinuria in both.** **Renal biopsy: one twin only, at age 6, FSGS** **ESRD: Renal transplantation: living-related (one from mom one from dad) at 14 years of age (both twins); no rejection,**	**Seizures: no overt epileptic disease or seizures in both twins; never on anti-seizure medications.** **DD: Onset at 1.5 years in both twins.**	**NA**

	no recurrence, still functioning after 11 years.	**Mild global developmental delay, hypotonia, and Tourette's syndrome like features (arms shaking, verbal and physical tics in both twins). Preserved verbal IQ, difficulties with fine movements (as handwriting). Good long term memory, poor short term memory.**	
RAP027	**Onset: 3.5 years** **Type of presentation Nephrotic Syndrome** **- Renal biopsy: 5 years.** **Findings: Severe glomerulosclerosis pattern consistent with diffuse mesangial sclerosis. Medullary cystic disease.** **- ESRD: 5 years** **- Renal transplantation: No.**	**Seizures: Atypical Febrile Seizures 3.5 years. No treatment** **DD: Mild developmental delay. Delayed motor milestones. Speech delay. Shy. Delayed motor skills** **Brain imaging: Age 5. Mild diffuse global cerebral and cerebellar atrophy**	**Other findings:** **Endocrine: Hypothyroidism (elevated TSH, 18 months). Benign breast buds during infancy.** **Dermatologic: Infected verruca plantaris**
DUKEPIMIK01	**Onset: 11 years in setting of rhinovirus infection with AKI and hypernatremia** **Presentation: Nephrotic syndrome** **- Renal biopsy: at 12 years. Findings: FSGS.** **Rapid progression to ESRD, initially on peritoneal dialysis.** **- Renal transplantation: 12.5 years**	**Seizure: onset age 2 years. Complex partial seizures and electrographic status epilepticus in sleep (ESES). Seizures resistant to lamotrigine, pregabaline, carbamazepine, clonazepam and folinic acid therapy. He had some relief with topiramate, levetiracetam and steroid therapy.** **Neurologic features** **DD: Onset 6 months, never developed speech, able to walk but not run** **Brain Imaging: MRI with nonspecific T2 signal abnormalities. EEG: Electrical Status Epilepticus In Sleep**	**Gastrointestinal: feeding difficulties, GERD, gastrostomy tube dependent.**

UDN171252	**Onset: 3 years** **Presentation: Nephrotic Syndrome.** **Chemistries at onset: Increased Crea.** **Urinalysis: Nephrotic Range Proteinuria.** **Renal US: crossed fused ectopia and vesicoureteral reflux (VUR).** **Renal Biopsy: FSGS at 3 years.** **Progression to ESRD at age 5 years.** **Tx: renal transplant at 6.5 years.**	**Seizures: Onset age 2.5 years, multidrug resistant** **DD: Onset 3-6 months. Aggressive and self-injurious behavior.** **Brain Imaging: MRI showed mesial temporal sclerosis**	**Facial Dysmorphisms: micrognatia.** **Ophthalmologic: astigmatism, amblyopia and pseudostrabismus** **Respiratory: reactive airway disease, obstructive sleep apnea, (underwent tonsillectomy and adenoidectomy)** **Gastrointestinal: GERD, esophagitis, dysphagia, and diarrhea.** **Endocrine: hypothyroidism at age 1 year.** **Dermatologic: jaundice at birth requiring phototherapy.**

Abbreviations: ACE, angiotensin-converting enzyme; BUN, blood urea nitrogen; CKD, chronic kidney disease; Crea, creatinine; CsA, cyclosporine A; DMS, diffuse mesangial sclerosis; DD, developmental delays; EEG, electroencephalography; eGFR, estimated GFR; EM, electron microscopy; ESRD, end-stage renal disease; FSGS, focal segmental glomerulosclerosis; GU, genitourinary; HC, head circumference; MMF, mycophenolate mofetil; MRI, magnetic resonance imaging; MSK, musculoskeletal; NRP, nephrotic-range proteinuria; UPC, urine protein-to-creatinine ratio; US, ultrasound; Tx, transplant; yrs, years. [&]FG-FA corresponds to a pair of monozygotic twins (genetically confirmed) who share the same *TRIM8* variant and a nearly identical phenotype.

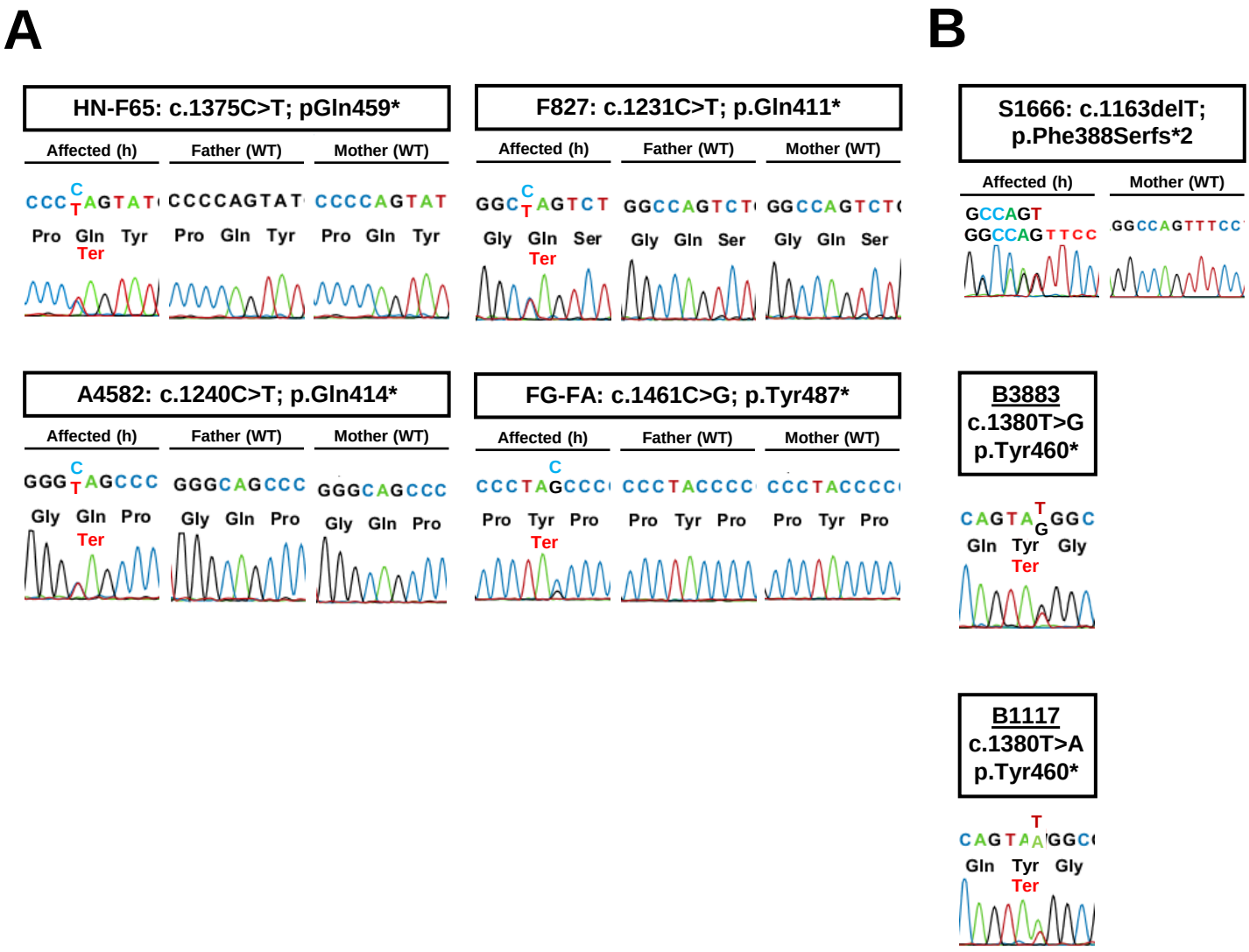
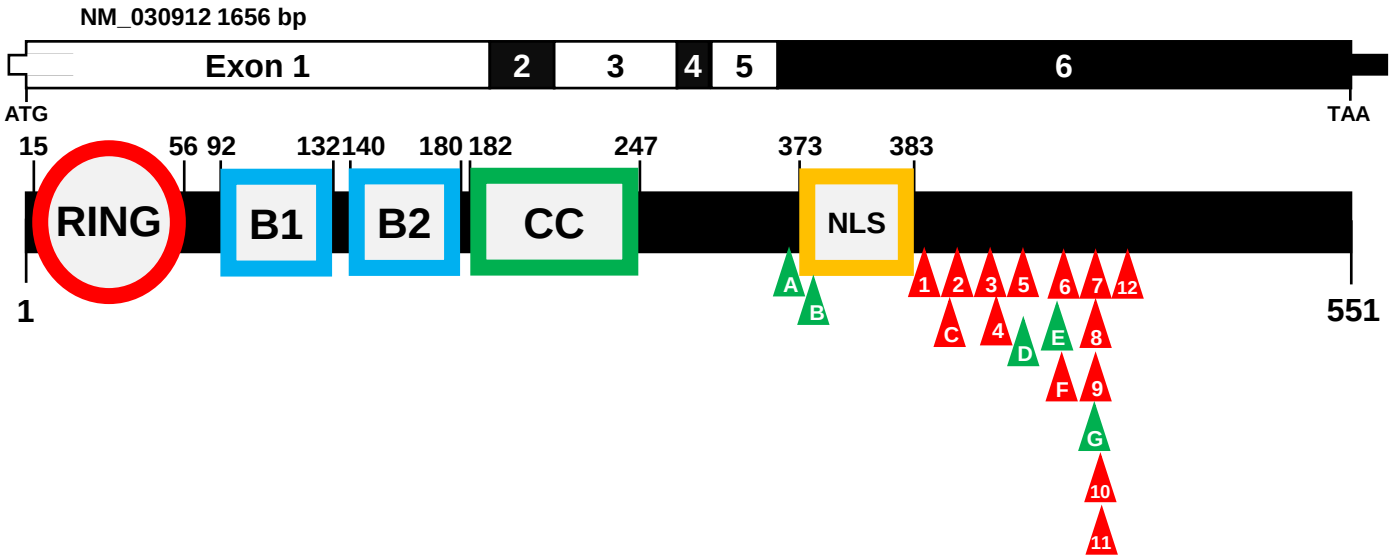


Figure S1. Sanger sequencing of *TRIM8* mutations in families with nephrotic syndrome and neurologic disease.

(A) Sanger sequencing chromatograms of *TRIM8* is shown from families with complete subject and parental DNA. All cases demonstrated heterozygous truncating *TRIM8* mutations in the affected subjects and wildtype sequence in the parents, consistent with *de novo* inheritance.

(B) Sanger sequencing chromatograms of *TRIM8* is shown from families where subject DNA was available but not both parents. All cases demonstrated heterozygous truncating *TRIM8* mutations in the affected subjects and wildtype sequence in the parental DNA, when it was available.

Abbreviations: h, heterozygous; WT, wildtype.



Current study

(1) S1666: c.1163del; p.Phe388Serfs*2

(2) RAP027: c.1201_1202delGGInsTA; p.Gly401*

(3) F827: c.1231C>T; p.Gln411*

(4) A4582: c.1240C>T p.Gln414*

(5) DUKEPIMIK01: c.1267C>T; p.Gln423*

(6) FSGSGE126: c.1333C>T; p.Gln445*

(7) HN-F65: c.1375C>T; p.Gln459*

(8) UC-023-1: c.1375C>T; p.Gln459*

(9) UDN171252: c.1375C>T; p.Gln459*

(10) B3883: c.1380T>G; p.Tyr460*

(11) B1117[#]:c.1380T>A; p.Tyr460* (Warren et al, 2020)

(12) FGFAA: c.1461C>G; p.Tyr487*

Previous reports

(A) Sakai et al., 2016: c.1099insG; p.Cys367Trpfs*43

(B) Assoum et al., 2018: c.1117delG; p.Ala374Argfs*16

(C) McClatchey et al., 2020: c.1198_1220del; p.Tyr400Argfs*2

(D) Assoum et al., 2018: c.1267C>T; p.Gln423*

(E) Assoum et al., 2018: c.1331C>A; p.Ser444*

(F) Assoum et al., 2018: c.1338T>A; p.Tyr446*

(G) Assoum et al., 2018: c.1375C>T; p.Gln459*

Figure S2. Dominant truncating mutations in *TRIM8* in 19 families with neurologic and renal disease.

Coding exon (upper bar) and protein domain (lower bar) structures of *TRIM8* are shown with numbered arrowheads indicating position of mutations identified in 12 subjects from the current study and 7 reported patients. All subjects had developmental delay and seizures. Those with isolated neurologic disease are colored in **green**, while those with nephrosis are colored in **red**. *TRIM8* disease mutations are all truncating alleles, which exclusively reside in the last exon.

B1, B-box domain 1; B2, B-box domain 2; CC, coiled-coil domain; RING, ring finger domain. [#]This patient was independently recruited and published, while the current manuscript was in preparation.

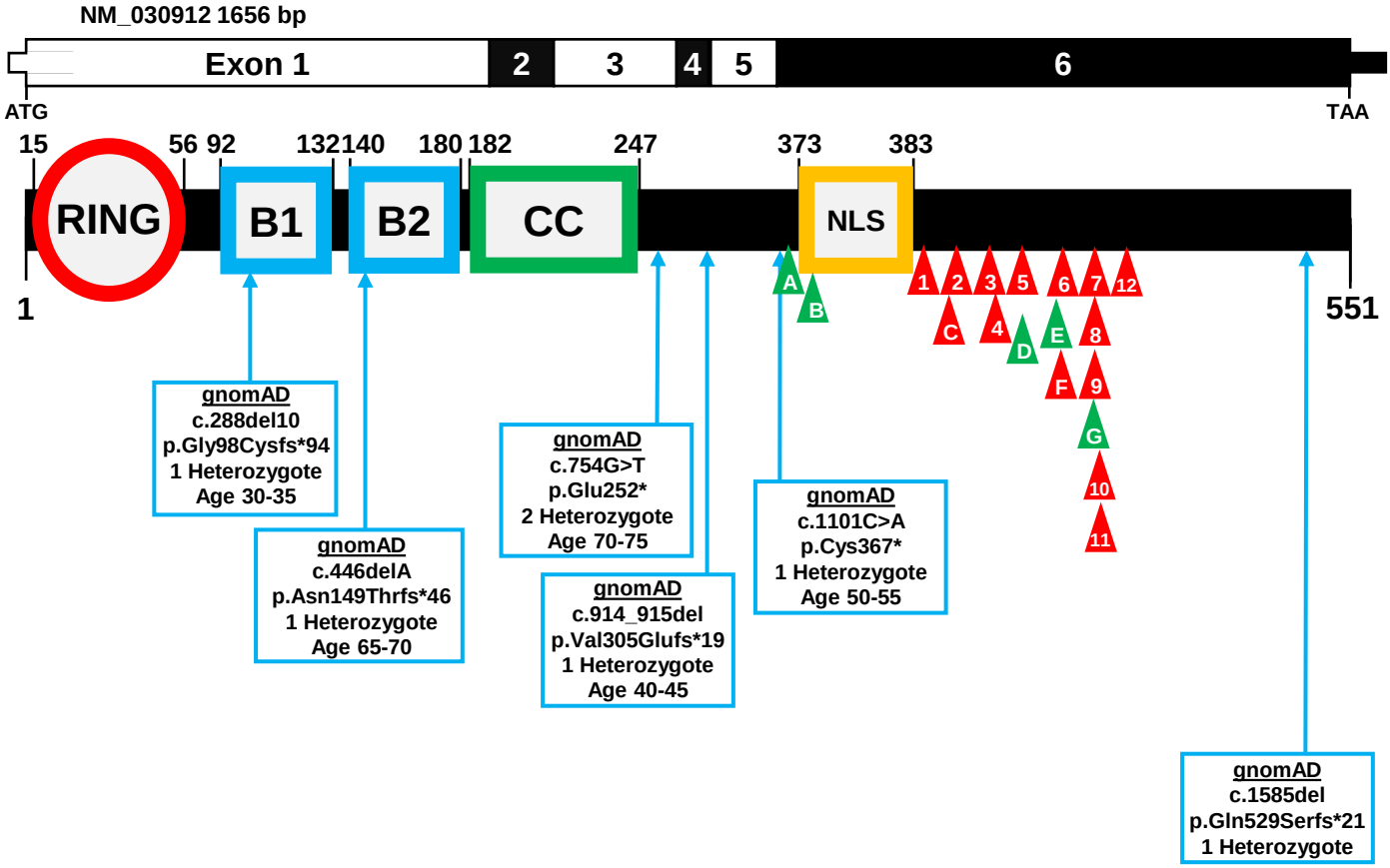


Figure S3. Dominant truncating mutations in *TRIM8* are spatially distinct from gnomAD subjects. Coding exon (upper bar) and protein domain (lower bar) structures of *TRIM8* are shown with numbered arrowheads indicating position of mutations identified in 12 subjects from the current study and 7 reported patients. All subjects had developmental delay and seizures. Those with isolated neurologic disease are colored in **green**, while those with neurologic disease and nephrosis are colored in **red.** *TRIM8* disease mutations are all truncating alleles, which exclusively reside in the last exon. (See **Figure S2** for precise disease mutation positions.) In addition, 7 gnomAD subjects with heterozygous truncating *TRIM8 alleles* (thin **blue** border) are shown, in which mutations arise earlier (98-367) or, in one subject, later (amino acid 529) than disease mutations. The gnomAD subjects are adults ages 30-75 years who do not have pediatric disease, suggesting that earlier or later TRIM8 truncating events are less likely to cause neurologic or renal disease. B1, B-box domain 1; B2, B-box domain 2; CC, coiled-coil domain; RING, ring finger domain.

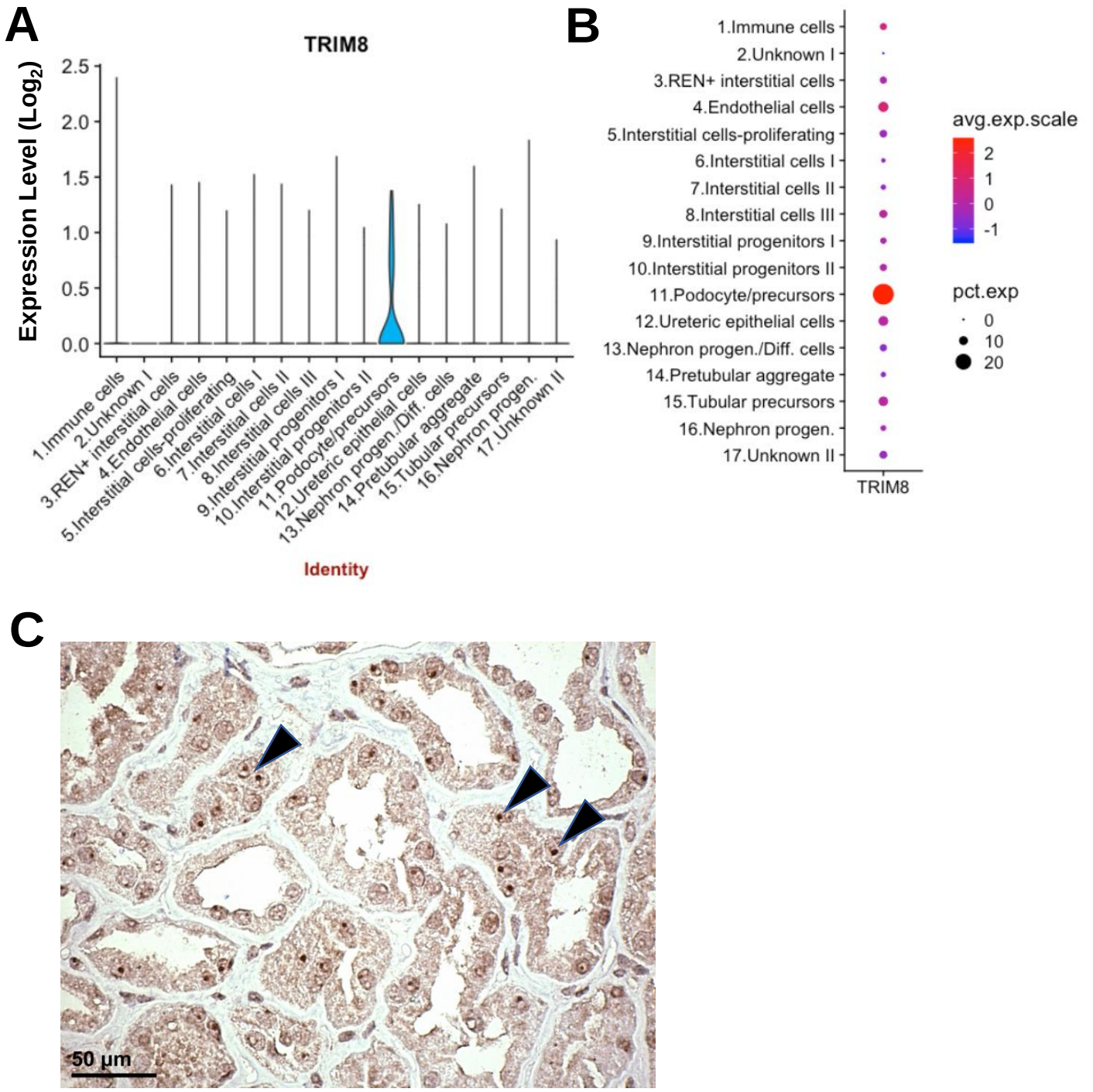


Figure S4. *TRIM8* mRNA expression in podocytes of human fetal kidney and protein localization in human kidney tissue by immunohistochemistry.

(A) Single cell mRNA sequencing data from 16 week human fetal kidney (Lindstrom et al. *JASN* 2018) was queried (http://humphreyslab.com/SingleCell) and revealed predominant expression of *TRIM8* mRNA in the podocyte/precursor cluster. This is shown by the violin plot reflecting the distribution of log-base 2 expression levels within each cluster.

(B) The expression of *TRIM8* mRNA is visualized within each fetal kidney cell cluster based on (i) the percent of cells within the cluster expressing *TRIM8* and (ii) the average gene expression (scaled to the other clusters).

(C) Immunohistochemistry of adult human kidney tissue demonstrate TRIM8 shows positivity most predominantly within nuclei but also cytoplasm of tubular epithelial cells. The nuclear staining is localized to discrete foci within the nuclei of these tubular epithelial cells. (Scale Bar: 50 µm)

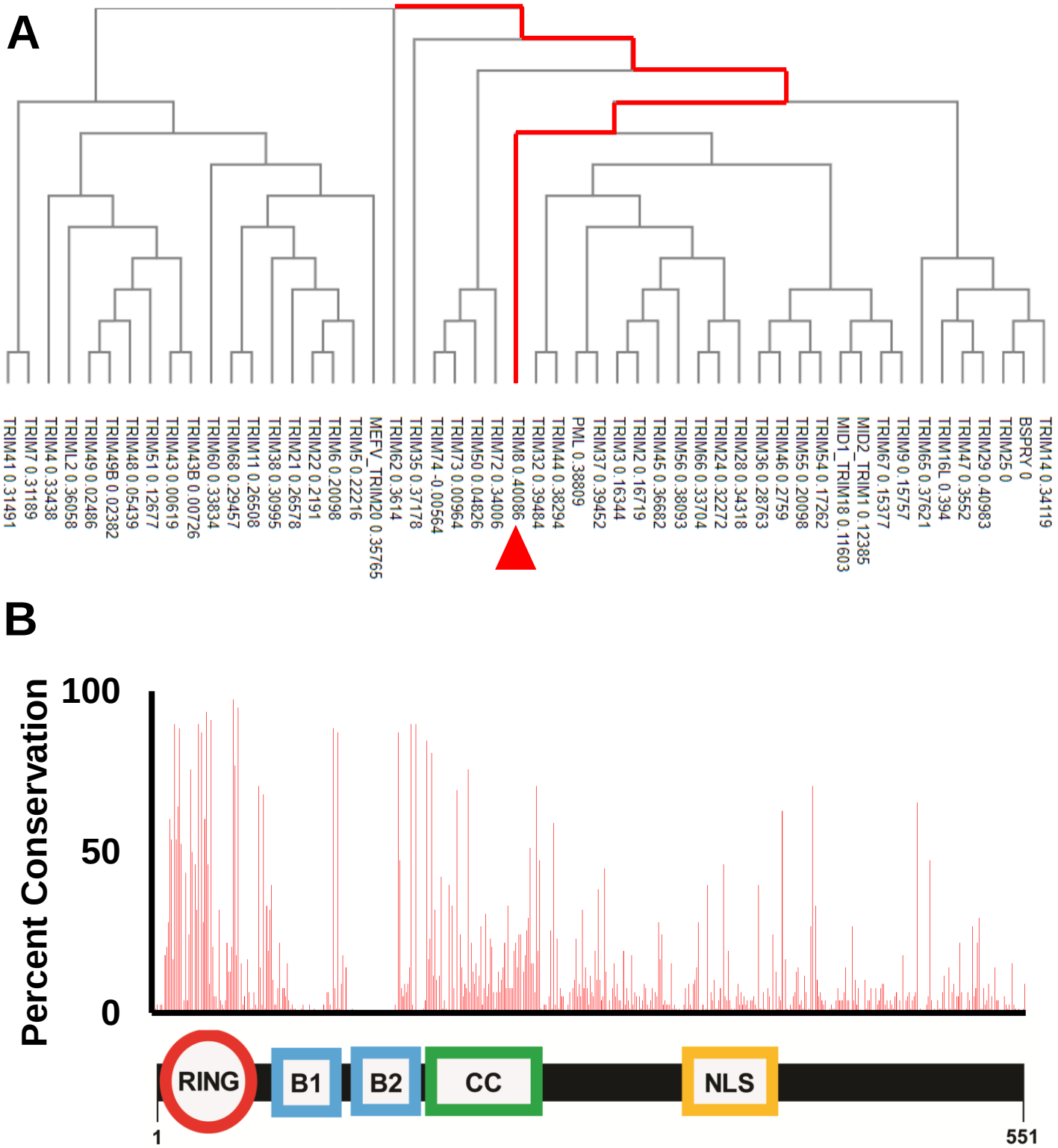


Figure S5. TRIM8 C-terminal region is not well conserved across 51 human TRIM8 paralogues.

(A) 51 human TRIM8 paralogues were assembled from Ensembl (ensemble.org) and previous literature (Reymond et al., *EMBOJ* 2001). The sequences were aligned using Clustal Omega, and the resulting dendogram is shown. The path of TRIM8 is highlighted in **red** (red arrow).

(B) The conservation of TRIM8 amino acid residues across 51 human TRIM paralogues is shown graphically as a percentage of amino acid identity across the 51 paralogues at each residue from N-terminus to C-terminus. The 551 amino acid protein domain structure of TRIM8 is displayed below as reference. The N-terminus, which contains the four conserved tri-partite domains, is more well conserved across paralogues than the C-terminus of TRIM8. B1, B-box domain 1; B2, B-box domain 2; CC, coiled-coil domain; NLS, nuclear localization signal; RING, ring finger domain.

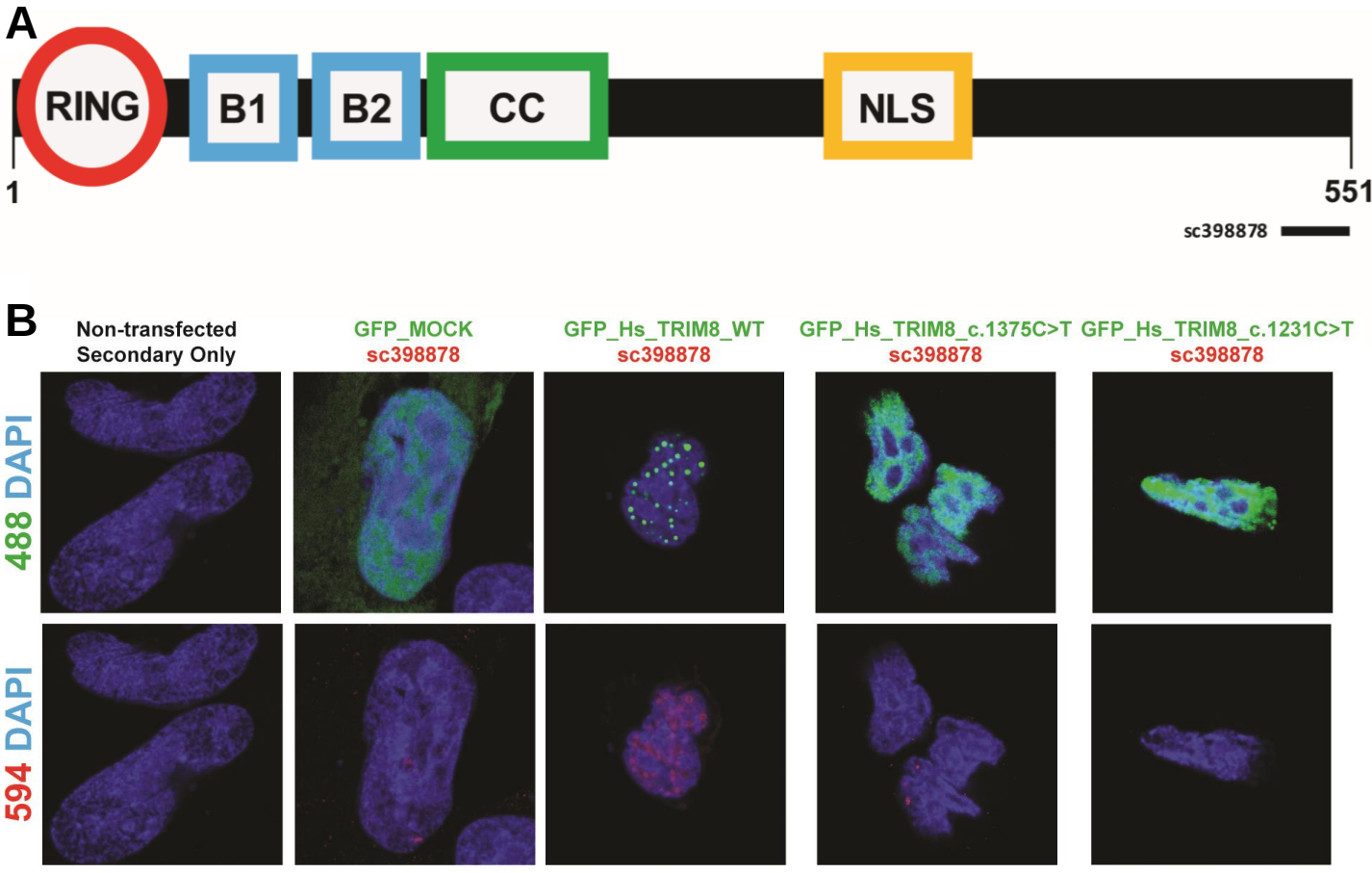


Figure S6. TRIM8 antibody detects GFP-tagged full-length TRIM8 over-expressed protein but not protein products of constructs reflecting patient mutations.

(A) TRIM8 protein domain structure is shown in relation to the immunogen against which the monoclonal mouse TRIM8 antibody sc398878 was generated. B1, B-box domain 1; B2, B-box domain 2; CC, coiled-coil domain; NLS, nuclear localization signal; RING, ring finger domain.

(B) Co-immunofluorescence with sc398878 reveals that this antibody identifies GFP-tagged TRIM8 upon overexpression in a human podocyte cell line (column 3) with overlapping staining of GFP-TRIM8 nuclear bodies. Secondary only and GFP-MOCK transfected cell controls show no background signal from the antibody (columns 1 and 2). The antibody does not recognize GFP-tagged TRIM8 constructs reflecting patient mutations (columns 4 and 5), which demonstrated pan-nuclear localization.

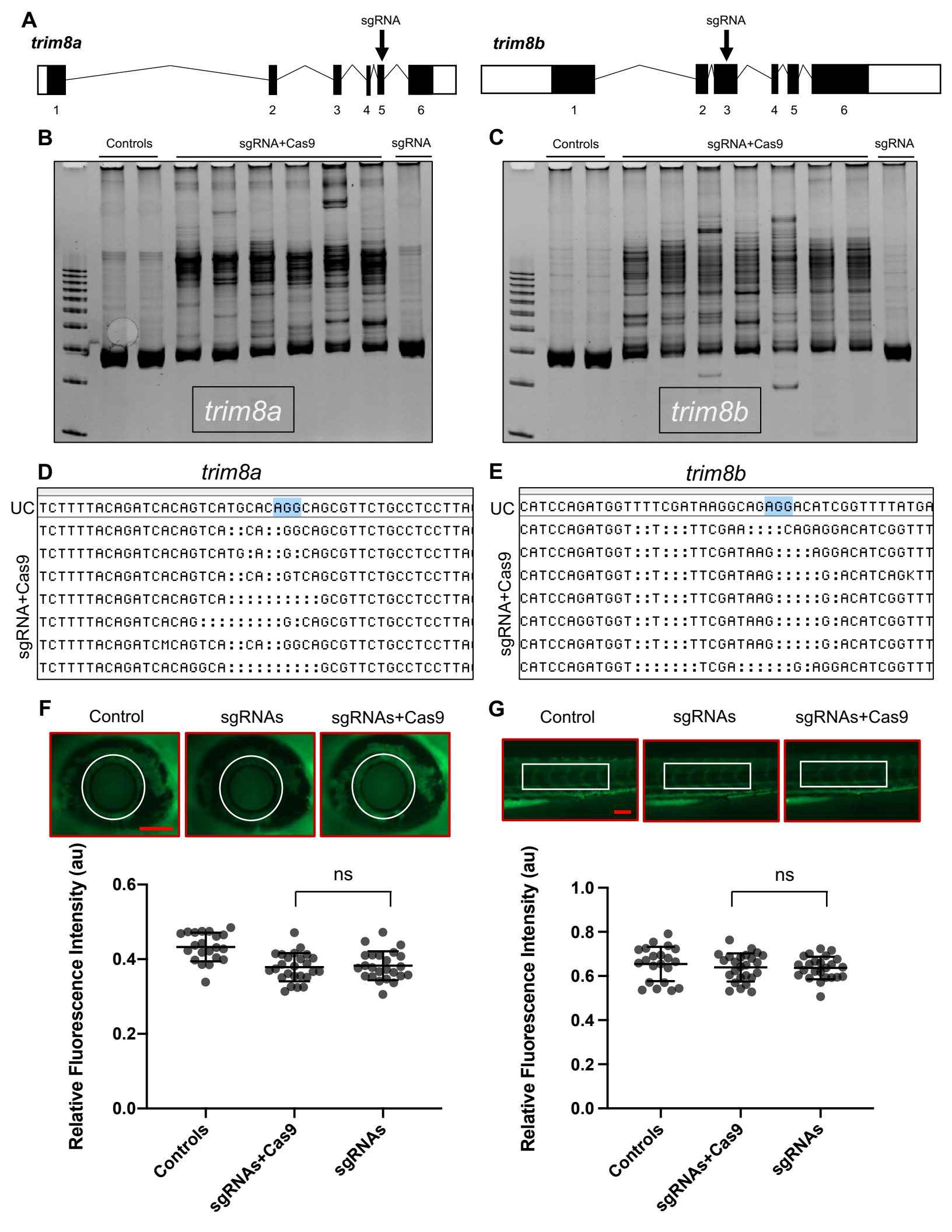

A
trim8a
sgRNA
1
2
3
4
5
6
trim8b
sgRNA
1
2
3
4
5
6
B
Controls
sgRNA+Cas9
sgRNA
trim8a
C
Controls
sgRNA+Cas9
sgRNA
trim8b
D
trim8a
UC
sgRNA+Cas9
TCTTTTACAGATCACAGTCATGCACAGGCAGCGTTCTGCCTCCTTA
TCTTTTACAGATCACAGTCA::CA::GGCAGCGTTCTGCCTCCTTA
TCTTTTACAGATCACAGTCATG:A::G:CAGCGTTCTGCCTCCTTA
TCTTTTACAGATCACAGTCA::CA::GTCAGCGTTCTGCCTCCTTA
TCTTTTACAGATCACAGTCA:::::::::GCGTTCTGCCTCCTTA
TCTTTTACAGATCACAG:::::::::G:CAGCGTTCTGCCTCCTTA
TCTTTTACAGATCMCAGTCA::CA::GGCAGCGTTCTGCCTCCTTA
TCTTTTACAGATCACAGGCA:::::::::GCGTTCTGCCTCCTTA
E
trim8b
UC
sgRNA+Cas9
CATCCAGATGGTTTTCGATAAGGCAGAGGACATCGGTTTTATGA
CATCCAGATGGT::T:::TTCGAA::::CAGAGGACATCGGTTT
CATCCAGATGGT::T:::TTCGATAAG::::AGGACATCGGTTT
CATCCAGATGGT::T:::TTCGATAAG:::::G:ACATCAGKTT
CATCCAGATGGT::T:::TTCGATAAG:::::G:ACATCGGTTT
CATCCAGGTGGT::T:::TTCGATAAG:::::G:ACATCGGTTT
CATCCAGATGGT::T:::TTCGATAAG:::::G:ACATCGGTTT
CATCCAGATGGT:::::::TCGA::::::G:AGGACATCGGTTT
F
Control
sgRNAs
sgRNAs+Cas9
Relative Fluorescence Intensity (au)
0.6
0.4
0.2
0.0
ns
Controls
sgRNAs+Cas9
sgRNAs
G
Control
sgRNAs
sgRNAs+Cas9
Relative Fluorescence Intensity (au)
1.0
0.8
0.6
0.4
0.2
0.0
ns
Controls
sgRNAs+Cas9
sgRNAs

Figure S7. Zebrafish *trim8a* and *trim8b* double F0 mutants did not display gross morphological or glomerular filtration defects.

(A) Schematic of zebrafish *trim8a* (Ensembl ID: ENSDART00000128249.4; GRCz11) and *trim8b* (Ensembl ID: ENSDART00000085888.6; GRCz11) loci. Exons, black rectangles; introns, black lines; UTRs, white rectangles; Black arrows indicate position of target sites for CRISPR single guide RNAs (sgRNA).

(B, C) Heteroduplex analysis of PCR products amplified from DNA extracted from individual uninjected controls or embryos injected with sgRNA+Cas9 and sgRNA alone on 15% polyacrylamide gels. Embryos were harvested at 2 dpf for genomic DNA extraction and target sites were PCR amplified using locus specific primers.

(D, E) Representative sequence alignments of CRISPR target sites for *trim8a*/*trim8b* F0 mutants and uninjected controls (UC) show that 100% of amplification products assessed were targeted. PCR products from individual embryos were TOPO-TA cloned and sequence confirmed (n=3 embryos per condition; 24 clones per embryo). Protospacer adjacent motif (PAM) sequences for both guides are marked with a blue rectangle.

(F) Top: representative live lateral fluorescent images of larval eyes at 6 dpf (4 days post-injection with 70 kDa Dextran FITC conjugates). Bottom: Embryos were injected with 70 kDa Dextran FITC conjugates in the cardiac venous sinus at 2 dpf and live fluorescent images of the eye were acquired at 3 dpf and 6 dpf. FITC signal was quantified using ImageJ; 6 dpf values were normalized to baseline (3 dpf) and statistical differences were calculated between controls and F0 mosaic mutants with a Student's t-test. The region-of-interest (ROI) is outlined with a white circle. Error bars show standard deviation of the mean; scale bar, 100 μm; ns, not significant.

(G) Top: representative live lateral fluorescent images of the larval trunk at 6 dpf (4 days post-injection with 70 kDa Dextran FITC conjugates). Bottom: Embryos were injected with 70 kDa Dextran FITC conjugates in the cardiac venous sinus at 2 dpf and live fluorescent images of the trunk were acquired at 3 dpf and 6 dpf. FITC signal was quantified using ImageJ; 6 dpf values were normalized to baseline (3 dpf) and statistical differences were calculated between controls and F0 mosaic mutants with a Student's t-test. The ROI is indicated with a white rectangle. Error bars show standard deviation of the mean; scale bar, 100 μm; ns, not significant.

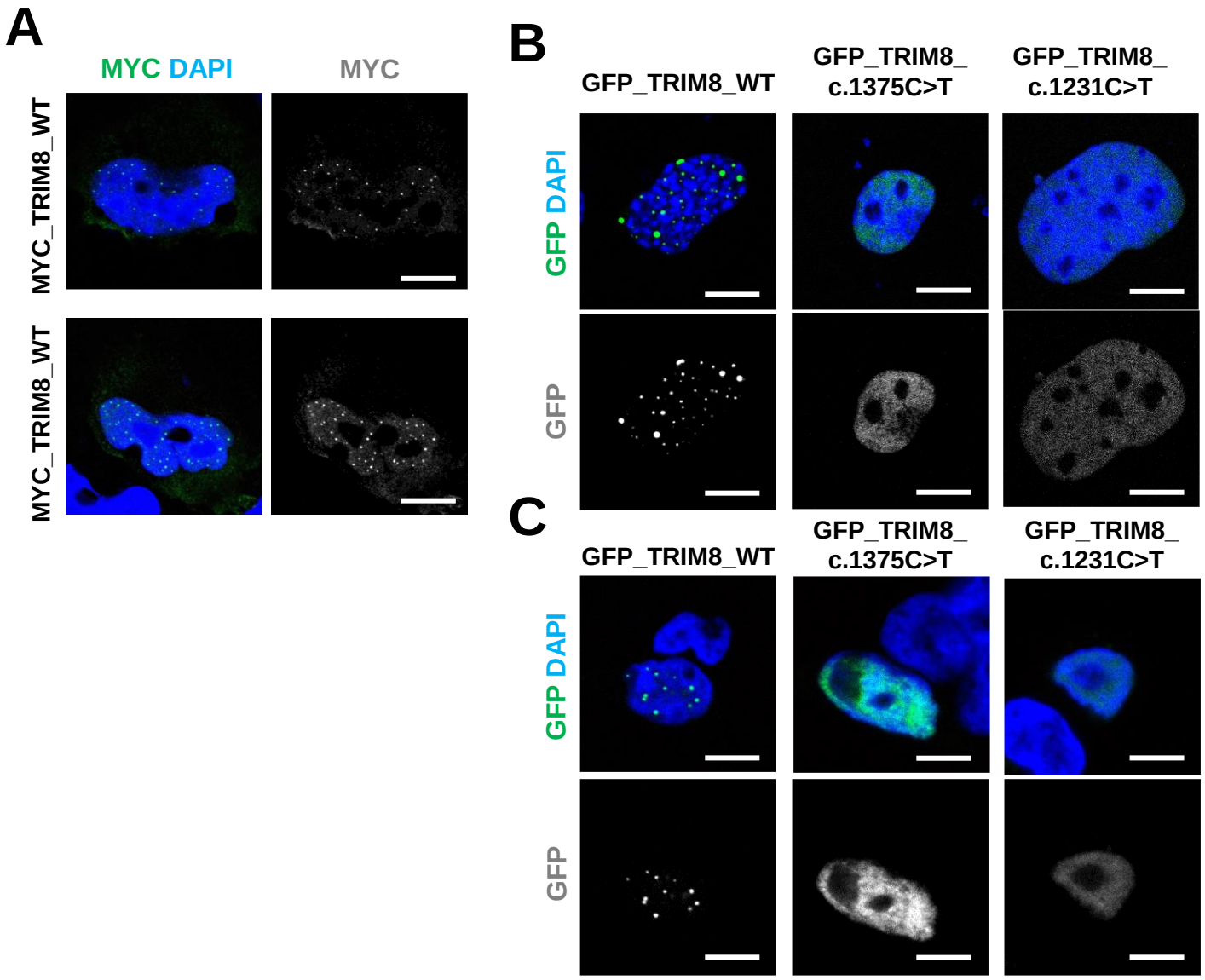


Figure S8. TRIM8 localizes to nuclear bodies with distinct N-terminal tag and in different cell lines.

(A) An immortalized human podocyte cell line was transfected with N-terminal MYC tagged wildtype *TRIM8* construct. Anti-MYC immunofluorescence was performed. Cells were imaged by confocal microscopy. Representative images of MYC-tagged protein and DAPI localization are shown, demonstrating that wildtype TRIM8 localizes to nuclear bodies. (Scale Bars: 7.5 µm).

(B) An immortalized mouse podocyte cell line was transfected with N-terminal GFP tagged wildtype *TRIM8* or *TRIM8* mutant constructs based on NS patient variants c.1375C>T and c.1231C>T. Cells were imaged by confocal microscopy. Representative images of GFP-tagged protein and DAPI localization are shown, revealing that wildtype TRIM8 localizes to nuclear bodies, while patient mutants exhibit pan-nuclear staining overlapping with DAPI signal. (Scale Bars: 10 µm).

(C) The neuroblastoma BE(2)-M17 cell line was transfected, processed and imaged as in (B). Representative images of GFP-tagged protein and DAPI localization are shown, revealing that wildtype TRIM8 localizes to nuclear bodies, while patient mutants exhibit pan-nuclear staining overlapping with DAPI signal. (Scale Bars: 5 µm).

SUPPLEMENTARY APPENDIX

Research subjects

The Sanna-Cherchi laboratory Pediatric SNRS cohort comprised 369 patients of less than 21 years of age recruited from the CUIMC Nephrology Division and 25 collaborating medical centers in 7 nations. The clinical diagnosis of Steroid Resistant Nephrotic Syndrome and FSGS was based on published clinical criteria and renal biopsy. Blood and clinical data were obtained following written informed consent from patients or their legal guardians. Epilepsy data and CUIMC controls were collected from previously conducted exome or genome sequencing, which data were hosted in an internal database at the Institute for Genomic Medicine (IGM). These data were previously consented to be available for control use. Controls with the broad clinical categories of brain malformation, congenital disorder, fetal ultrasound anomalies, epilepsy and other neurologic features were excluded from the controls dataset. The study was approved by the Columbia University Institutional Review Board and local ethics committee.

The Hildebrandt laboratory (BCH Cohort) obtained blood samples and pedigrees following informed consent from individuals with NS or their legal guardians. The diagnose of NS was based on published clinical criteria and renal biopsies criteria evaluated by renal pathologists[1]. Patients recruited for other renal disease entities including renal stone disease were used as controls, as a nephrologist excluded NS in these patients through their evaluation. Clinical data were obtained using a standardized questionnaire (http://www.renalgenes.org).

The Pollak laboratory (BIDMC) obtained blood or saliva samples for DNA extraction as well as clinical and family data following informed consent from individuals with FSGS and/or NS or their legal guardians.

Patients were recruited to the National Study of Nephrotic Syndrome (NephroS) via the United Kingdom Registry for Rare Kidney Diseases (RaDaR). Detailed phenotypic information was entered online (https://nww.radar.nhs.uk) and laboratory data were laboratory data were automatically populated via links to the UK Renal Registry (www.renalreg.org). Appropriate informed consent from parents and/or carers was collected, and assent for collection of data and genetic analysis obtained. The study was approved by the South West research ethics committee and the institutional review board at each recruiting center. Inclusion criteria for patients to enter RaDaR included: Children and

adults (no age restrictions), Idiopathic Nephrotic Syndrome (nephrotic range proteinuria and/or hypoalbuminaemia). This includes Congenital NS (presumed Steroid resistance), Childhood or adult onset with primary Steroid Resistance, Childhood or adult onset with late onset Steroid Resistance, Steroid Sensitive Nephrotic Syndrome but early in the disease course (i.e. after one episode of Nephrotic Syndrome), or syndromes (e.g. Nail Patella Syndrome and Denys-Drash Syndrome). Those with a biopsy diagnosis of FSGS or minimal change disease could be included if they fall in the above categories, but biopsy is not a prerequisite for inclusion. Patients with Secondary causes of Nephrotic Syndrome e.g. primary diagnosis of Glomerulonephritis (IgA Nephropathy, Membranoproliferative Glomerulonephritis, Membranous Nephropathy), Vasculitis, Systemic Lupus Erythematosus, Diabetes, Obesity, Hypertension were excluded.

For the Technical University Munich Cohort, subject TUMG_1 was recruited as part of an exome sequencing (ES) study on hereditary kidney disease. All recruited index cases had either a) the clinical tentative diagnosis of a hereditary kidney disease or b) renal involvement without overlap with a specific hereditary kidney disease but met at least one the following inclusion criteria: (1) first manifestation before the age of 18 years; (2) Syndromic disease, i.e. involvement of an additional organ system apart the kidney. If there was only one extrarenal manifestation which had a prevalence of more than 1% in the general population this case was not classified as a syndromic case; (3) familial occurrence (>1 equally affected individual in a family); (4) reported consanguinity. Patients were recruited, applying the criteria mentioned above, either directly at the Institute of Human Genetics of the Klinikum rechts der Isar, Technical University of Munich, which is a tertiary care center. Or they were recruited by external human geneticists and (pediatric) nephrologists and referred to the Institute. Phenotypes were ascertained by reviewing medical reports and filling out a standardized questionnaire. Written informed consent for sequencing and publication of results was obtained from both parents. DNA was extracted from peripheral blood using the Gentra Puregene Blood Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

The Alberta Children's Hospital cohort was compiled through two exome sequencing research projects: (i) Enhanced Care for Rare Genetic Diseases in Canada (Care4Rare) research project; and (ii) Rapid Access to Pediatric Diagnoses (RAPiD) kidomics. Participants enrolled in the Care4Rare clinical diagnostic arm consisted of pediatric or adult index cases who had an undiagnosed suspected rare (typically multisystem) genetic condition. The RAPiD participants included children admitted to an acute care facility suspected to have an underlying but undiagnosed genetic disorder. Included in the

Alberta Children's Hospital cohort are all probands enrolled in Care4Rare and RAPiD who presented with suspected seizures or epilepsy amongst their diagnostic features at the time of enrollment (N=20). Peripheral whole blood samples were collected from the participants and genomic DNA was extracted using Gentra Puregene blood kit (QIAGEN).

ES and mutation calling

For the Sanna-Cherchi laboratory (CUIMC Cohort), genomic DNA was isolated from whole blood according to standard protocols and was subjected to exome sequencing with the Agilent, Roche or Integrated DNA Technologies (IDT) capture kits. Sequencing data were processed and aligned to Genome Reference Consortium Human Genome Build 37 (GRCh37 hg19) using the Dynamic Read Analysis for Genomics (DRAGEN) platform. Variants were called in accordance to the best practices outlined in Genome Analysis Tool Kit (GATK v3.6) and annotated with Ensembl-GRCh37.73 annotations and external data from the gnomad 2.2.1 reference population databases. The in-house Analysis Tools for Annotated Variants (ATAV, https://github.com/igm-team/atav) developed and maintained by the IGM was used for the variant prioritization analysis. Briefly, we prioritized variants that were either absent or ultra-rare, < 0.1% in the gnomAD v2.1.1 databases, with minimum sequencing depth of 10 reads, minimum genotype quality of 50, and were predicted to be cause loss-of-function; These include non-synonymous, frameshift, splice-site donor, splice-site acceptor and stop gained variants.

For the Hildebrandt laboratory (BCH cohort), exome sequencing (ES) was performed through (i) the Yale Genomics Center using Agilent SureSelect™ human exome capture arrays (Thermo Fisher Scientific) with next generation sequencing (NGS) on an Illumina™ platform or (ii) through the Broad Institute Center for Mendelian Genomics. For WES data from Yale, sequence reads were mapped against the human reference genome (NCBI build 37/hg19) using CLC Genomics Workbench (version 6.5.1) (CLC bio). Genetic location information is according to the February 2009 Human Genome Browser data, hg19 assembly (http://www.genome.ucsc.edu). Downstream processing of aligned BAM files were done using Picard and samtools[2], and SNV calling was done using GATK5. For WES through the Broad Institute, data processing was performed by the Genomics Platform at the Broad Institute of Harvard and MIT (Broad Institute, Cambridge, MA). Exome sequencing (>250 ng of DNA, at >2 ng/μl) was performed using Illumina exome capture (38 Mb target). Single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) were jointly called across all samples using the Genome Analysis Toolkit (GATK) HaplotypeCaller. Default filters were applied to

SNP and indel calls using the GATK Variant Quality Score Recalibration approach. Lastly, variants were annotated using the Variant Effect Predictor. For additional information, please refer to the Supporting Information Section S1 in the exome aggregation consortium (ExAC) study[3]. The variant call set was uploaded on to Seqr (https://seqr.broadinstitute.org) and analysis of the entire WES output was performed. From both platforms, mutation calling was performed in line with proposed guidelines[4], and the following criteria were employed as previously described[5,6]. The variants included were rare in the population with mean allele frequency <0.1% and with 0 homozygotes in the adult reference genome databases ExAC and gnomAD. Additionally, variants were non-synonymous and/or located within splice-sites. Based on an autosomal homozygous recessive hypothesis, homozygous variants were evaluated. Subsequently, variant severity was classified based on prediction of protein impact (truncating frameshift or nonsense mutations, essential or extended splice-site mutations, and missense mutations). Splice-site mutations were assessed by *in silico* tools MaxEnt, NNSPLICE, HSF, and CADD splice-site mutation prediction scores[7–10]. Missense mutations were assessed based on SIFT, MutationTaster and PolyPhen 2.0 conservation prediction scores[11–13] and evolutionary conservation based on manually derived multiple sequence alignments.

For the Pollak laboratory (BIDMC cohort), exome sequencing was performed as previously[14].

For the NephroS cohort, exome sequencing was performed in the Genomics Core Facility of the Biomedical Research Centre at Guy's and St Thomas' Hospitals and King's College (London) and the whole genome sequenced at the NIHR BioResource (Cambridge). Sequencing was performed on Illumina platforms with mean coverage of 120X for the exome and 35X for the genome. Variant calling and annotation was performed using the King's College London BRC Genomics and NIHR Cambridge BioResource pipelines.

For the Technical University Munich Cohort, Trio ES was performed using a Sure Select Human All Exon 60 Mb V6 Kit (Agilent, USA) and a HiSeq4000 (Illumina, USA) as previously described[**15**]. Mitochondrial DNA was derived from off-target exome reads as previously described[**16**]. Reads were aligned to the human reference genome (UCSC Genome Browser build hg19) using Burrows-Wheeler Aligner (v.0.7.5a). Detection of single-nucleotide variants and small insertions and deletions (indels) was performed with SAMtools (version 0.1.19). ExomeDepth was used for the detection of copy number variations (CNVs). A noise threshold of 2.5 was accepted for diagnostic analysis[**17**]. Called CNVs were visualized by the Integrative Genomics Viewer (IGV,

https://software.broadinstitute.org/software/igv/) to check for sufficient coverage of the inspected region and plausibility of the CNV. CNVs were compared with publicly available control databases like the Genome Aggregation Database (gnomAD, https://gnomad.broadinstitute.org/about), the Database of Genomic Variants (DGV, http://dgv.tcag.ca/dgv/app/home) and databases for pathogenic CNVs like DECIPHER (https://decipher.sanger.ac.uk/) and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/). For the analysis of de novo, autosomal dominant and mitochondrial variants, only variants with a minor allele frequency (MAF) of less than 0.1% in the in-house database of the Helmholtz Center Munich containing over 19,000 exomes were considered. For the analysis of autosomal recessive and X-linked variants (homozygous, hemizygous or compound heterozygous) variants with a MAF of less than 1.0% were considered. Identified variants were compared with publicly available databases for pathogenic variants like ClinVar, the Human Gene Mutation Database (HGMD®, http://www.hgmd.cf.ac.uk) and the Leiden Open Variation Database (LOVD, https://www.lovd.nl).

For the Undiagnosed Disease Network (UDN) individuals, exome sequencing was conducted as previously described[18].

For the Care4Rare enrolled participants of the Alberta Children's Hospital cohort, sequencing was conducted using the Illumina TruSight One gene panel, Agilent SureSelect Clinical Research Exome v2 and sequenced on the Illumina MiSeq or the NextSeq 550. Bioinformatics including read alignment, variant calling and annotation were completed as described previously[19]. For the RAPiD participants, exome library capture was performed using the xGen Exome Research Panel v.1.0 (Integrated DNA Technologies (IDT)) and sequenced on the Illumina NextSeq 550. Data were uploaded to Illumina BaseSpace and the DRAGEN Enrichment app (v.3.5.7) was used to align data (to hg19 reference genome). Vcf files produced from BaseSpace were uploaded to Fabric Genomics for variant interpretation and report generation. Variants were excluded if they met at least one of the following criteria: 1) non-coding (excluding the first 6 nucleotides and last 15 nucleotides of introns); 2) have a minor allele frequency greater than 0.5% from any general population database (e.g. gnomAD); 3) fewer than 4 variant reads in the proband. The remaining variants were assessed for pathogenicity primarily using the Fabric Genomics variant annotation platform and the VAAST Variant Prioritizer to assign prioritization score to each resulting variant[20].

Accession numbers

Human TRIM8 full-length protein (GenBank accession NP_112174) encoded by GenBank accession NM_030912.

Molecular cloning and site directed mutagenesis

We obtained a full length wild type *TRIM8* human open reading frame clone (Genebank: NM_030912.2) in a Gateway entry vector backbone (GeneArt, ThermoFisher; pENTR221; clone ID: IOH13368). We performed PCR based site directed mutagenesis (2x Phusion Master Mix; ThermoFisher) using an in-house developed protocol[21], and synthesized TRIM8 clones carrying two randomly selected variants from our case cohort (c.1231C>T; p.Gln411* and c.1375C>T; p.Gln459*). We confirmed all constructs using Sanger sequencing (BigDye3.1 chemistry on an ABI 3730xl, Applied Biosystems) and cloned it into the pCDNA6.2 N-DEST-Em-GPF vector backbone using LRII clonase-mediated recombination (ThermoFisher). *TRIM8* cDNA was cloned into pRK5-N-MYC tag construct using a LRII clonase-mediated recombination as above. Prior to *in vitro* cell studies, all constructs were confirmed by bidirectional Sanger sequencing and restriction enzyme digestion.

Cell lines

Human immortalized podocytes were a gift of Moin Saleem (University of Bristol, Bristol, United Kingdom)[22] and were cultured as previously described[23–25]. Immortalized wildtype mouse podocytes were a gift from Dr. Minoru Takemoto (International University of Health and Welfare, Chiba, Japan) and were cultured as previously described[26]. The neuroblastoma cell line BE(2)-M17 cells were a kind gift from the laboratory of Timothy Yu (Boston Children's Hospital, Boston, Massachusetts, USA) and were cultured as described by ATCC (CRL-2267).

Antibodies and immunostaining reagents

The following primary antibodies were used: rabbit anti-SMN1 (Novus, NBP1-03326), rabbit anti-p80 Coilin (Sigma, PLA0290), mouse anti-TRIM8 (Santa Cruz, sc-398878), mouse anti-myc (abcam, 9E10). Donkey anti-mouse and anti-rabbit Alexa 594–conjugated secondary antibodies and DAPI staining reagents were obtained from Invitrogen (Thermo Fisher Scientific).

Immunohistochemistry

Immunostaining was conducted on human kidney biopsies using 3,3'-diaminobenzidine (DAB). TRIM8 antibody "TRIM8 (B-3): sc-398878" was purchased from Santa Cruz Biotechnology, Inc.

Immunofluorescence in immortalized human podocytes

Human immortalized podocytes were seeded on coverslips and grown at a permissive temperature. For overexpression studies, podocytes were transiently transfected using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions with 250 ng of GFP-tagged *TRIM8* construct plasmids and 100 ng of MYC-tagged *TRIM8* construct plasmid. Immortalized mouse podocytes and BE(2)-M17 cells were similarly transfected with 500 ng and 300 ng, respectively, of GFP-tagged *TRIM8* construct plasmids. Experiments were performed 24 hours after transfection. Cells were fixed for 15 minutes using 4% paraformaldehyde and permeabilized using 0.1% Triton-X 100. After blocking, sections were incubated overnight at 4°C with primary antibody (if used). If primary antibody used, the cells were incubated in secondary antibodies for 90 minutes at room temperature followed by mounting in hardening medium with DAPI. Confocal imaging was performed using the Leica SP5X system with an upright DM6000 microscope, and images were processed with the Leica AF software suite.

Assessment of adjacent nuclear bodies

For visualization of nuclear bodies, three independent experiments performed, and data pooled. Z-stack image sets of individual nuclei at 100X magnification were obtained which spanned the entire thickness of the nuclei. All Gemin or Cajal bodies within the nuclei were scored for contact a TRIM8 NB. Contact between nuclear bodies was determined through (i) three-dimensional reconstruction and rotation of the rendering to ensure contact between adjacent nuclear bodies and (ii) confirmation by identifying specific image(s) in Z-stack where contact observed.

CRISPR-Cas9 genome editing in zebrafish

Using reciprocal protein BLAST with human protein, we identified two TRIM8 orthologues in zebrafish: *trim8a* (Ensembl: ENSDART00000128249.4; GRCz11; 67% identity and 79% similarity) and *trim8b* (Ensembl: ENSDART00000085888.6; GRCz11; 67% identity and 78% similarity). We identified CRISPR single guide RNA (sgRNA) target sites for each independent ortholog using ChopChop2[27]: *trim8a:* 5'-TACAGATCACAGTCATGCACAGG-3' and *trim8b:* 5'-GATGGTTTTCGATAAGGCAGAGG-3'. We synthesized sgRNAs for both *trim8a* and *trim8b* using the GeneArt Precision gRNA synthesis kit according to manufactures' instructions (ThermoFisher). To test the efficiency of sgRNAs, we independently injected one-cell stage zebrafish embryos with cocktails containing 100 pg of sgRNA and 200 pg of Cas9 protein (PNA Bio, CP01). Embryos were

harvested at 2 dpf and genomic DNA was extracted with proteinase K digestion (Life Technologies, AM2548). The target sites were amplified for both *trim8a* and *trim8b* using flanking primers, PCR product was denatured, slowly reannealed, and heteroduplexes were detected using electrophoresis on a 15% polyacrylamide gel as described[28]. To estimate mosaicism, PCR products were cloned into the TOPO-TA cloning vector (ThermoFisher Scientific) and sequenced confirmed using bidirectional Sanger sequencing (n=3 embryos per condition, 24 clones per embryo; 100% mosaicism identified in both guides). The efficiency of sgRNAs (*trim8a*/*trim8b*) in double F0 zebrafish mutants was re-confirmed by heteroduplex analysis.

Zebrafish embryo injections and phenotyping

All zebrafish experiments were approved by the Institutional Animal Care and Use Committees (IACUC) at Duke University and Northwestern University. To simultaneously target both *trim8a* and *trim8b* loci, we injected zebrafish embryos at the one-cell stage with CRISPR-Cas9 cocktail containing 100 pg of each *trim8a* and *trim8b* sgRNAs and 200 pg of Cas9 protein. Double F0 mosaic mutants were assessed for survival and gross morphological phenotypes in comparison with embryos injected with sgRNAs alone daily until 6 dpf (n=30-45 larvae/condition, repeated twice). To investigate glomerular filtration defects, we injected 70 kDa Dextran-FITC conjugate (Sigma) into the cardiac venous sinus of double F0 mosaic mutants at 2 dpf. We acquired live lateral fluorescence images of the eye and trunk at 3 dpf (baseline) and 6 dpf using a ZEISS Axio V16 microscope and Axiocam 503 monochromatic camera facilitated by Zen Pro software. The intensity of FITC signal was quantified in eye and trunk regions of interest using Image J software and statistical differences (Student's t-test) were calculated by comparing relative FITC signal intensity in F0 mutants and controls at 6 dpf (normalized to baseline at 3 dpf; n=21-25, repeated). Experiments were performed with investigator blinded to injection cocktail.

Single-cell mRNA sequencing data analysis

Heatmap results depicting differential mRNA expression levels (from z-scores) was based on single-cell transcriptomics data from week 12-19 human fetal kidneys, E14.5 mouse fetal kidneys, or 8-week-old wild-type CD1 male mice[29–31]. Processed data from each set was queried for percent expression in pre-defined cell clusters. Queried data was normalized using z-score calculation as previously described[32]. Data was also viewed from week 16 human fetal kidneys[33] using the Humphreys Laboratory website (URL below).

Statistics

Burden of *TRIM8* truncating mutations was conducted using Fischer's Exact Test (R version 4.0.1). For the *de novo* analysis, the expected probability ''mu'' of *de novo* truncating mutations in *TRIM8* was estimated based on Samocha et al[34]. We then assumed as if we had trio data available for all 2,501 index case subjects. Therefore, the probability of observing at least six *de novo* truncating mutations in 2,501 independent trios was P(X>=6) where X~Poisson(2*2,501*mu) = 2.21×10^{-15}. A reference set of 12,840 LD-pruned informative SNPs (MAF>0.05) was used to infer relatedness in cases and controls. Briefly, variants were extracted from the VCF files of each sample, merged and converted into the PLINK binary format with PLINK v1.90b3.38 (www.cog-genomics.org/plink/1.9/)[35]. King 1.4 (http://people.virginia.edu/~wc9c/KING/) was used to estimate pairwise kinship coefficients in the cohort using the –kinship option, one of each pair of samples with an estimated second-degree or greater relationship (> 0.0884) was removed to retain unrelated cases and controls in the cohort[36].

Genetics Study approval

For Sanna-Cherchi laboratory (CUMC Cohort), human subject's research performed in this study was in accordance with the ethical standards of and approved by the Institutional Review Board of Columbia University and collaborating institutions.

For the Hildebrandt laboratory (BCH Cohort), human subject's research performed in this study was in accordance with the ethical standards of and approved by the Institutional Review Boards of the University of Michigan, Boston Children's Hospital, and local IRB equivalents.

The NephroS study was in accordance with the ethical standards of and approved by the South West research ethics committee and the institutional review board at each recruiting center.

For the Technical University Munich Cohort, this study was approved by the local Ethics Committee of the Technical University of Munich and performed according the standard of the Helsinki Declaration of 2013.

For the Pollak laboratory (BIDMC), human subject research was in accordance with the ethical standards of and approved by the Institutional Review Boards at Beth Israel Deaconess Medical Center.

All participating individuals through the University of Calgary provided informed written consent for a study, which was in accordance with the ethical standards of and approved by the University of Calgary Conjoint Health Research Ethics Board.

Supplementary Acknowledgement for ES and GS Data Sharing

Additional funding Support

N.M. is supported by funding from the National Institutes of Health T32-DK007726 grant. F.B. was supported by a fellowship grant (404527522) from the German Research Foundation (DFG). A.C.O.-W. is supported by the NIH F32 Ruth L. Kirschstein Postdoctoral Individual National Research Service Award (DK122766). V.K. is supported by the Deutsche Forschungsgemeinschaft (VK-403877094). T.M.K. was supported by a Post-Doctoral Fellowship award from the KRESCENT Program, a national kidney research training partnership of the Kidney Foundation of Canada, the Canadian Society of Nephrology, and the Canadian Institutes of Health Research. J. A. and B.B.B. are supported by the Deutsche Forschungsgemeinschaft Clinical research unit (KFO) 329 (grants AL901/2-1and AL901/3-1 to J.A. and grants BE6072/2-1 and BE6072/3-1 to B.B). We thank Xenia Latypova for technical assistance with zebrafish studies. Studies conducted by A.V.T., J.S.P., and A.M.I. was funded by the Alberta Children's Hospital Foundation Research Excellence Award and the Alberta Children's Hospital Research Institute. A.V.T. as supported through Alberta Children's Hospital Research Institute Clinical Research Fellowship. A.B. is funded by Kidney Research UK (Post-doctoral Fellowship).

Non-author contributions

The full list of contributors of sequencing data for cases and controls at the Institute for Genomic Medicine (IGM) and Undiagnosed Disease Network (UDN) can be found in the Supplementary Appendix. We want to thank Rajasree Menon and Jason Spence, University of Michigan, Ann Arbor, MI, USA for providing us with the data of their single-cell RNA sequencing study in human fetal kidneys.

List of the Institute for Genomic Medicine (IGM) collaborators for genomic data sharing.

K. Welsh-Bomer; C. Hulette; J. Burke; D. Valle; J. Hoover-Fong; N. Sobriera; A. Poduri; M. Connors; S. Palmer; R. Buckley; D. Murdock; the Murdock study community registry and biorepository Pro00011196; Kristen Newby; National Institute of Allergy and Infectious Diseases Center for HIV/AIDS Vaccine Immunology (CHAVI); R. Ottman; V. Shashi; Epilepsy Genetics Initiative, A Signature Program of CURE; E. Holtzman; S. Berkovic; I. Scheffer; B. Grinton; C. Depondt; S.

Sisodiya; G. Cavaller; N. Delanty; the Epi4K Consortium and Epilepsy Phenome/Genome Project; H. Hinose; the ALS Sequencing Consortium; Washington University Neuromuscular Genetics Project; Genomic Translation for ALS Care (GTAC) Group; P. Lugar; G. Nestadt; J. Samuels; Y. Wang; S. Schuman; E. Nading; C. Garcia; W.Lowe; D. Levy; E. Pras; D. Lancet; T. Young; K. Whisenhunt; C. Chen; R. Wapner; C. Moylan; A. Mae Diehl; M. Abdelmalek; DUHS (Duke University Health System); Nonalcoholic Fatty Liver Disease Research Database and Specimen Repository; D. Daskalakis; M. Winn; R. Gbadegesin; M. Hauser; S. Delaney; A. Need; J. McEvoy; Eleanor and Lou Gehrig ALS Center at Columbia University; the Columbia University Precision Medicine Initiative; A. Holden; E. Behr; M. Walker; M. Sum; Undiagnosed Diseases Network; R. Mayeux; National Institute on Aging; WHICAP project; Utah Foundation for Biomedical Research; S. Kerns, H. Oster. We also thank the investigators, staff, and participants in these studies.

The collection of control samples and data was funded in part by: Bryan ADRC NIA P30 AG028377; NIH RO1 HD048805; NIAID/NIH Grant# UO1AIO67854; Gilead; National Institute of Allergy and Infectious Diseases Center for HIV/AIDS Vaccine Immunology (CHAVI) (U19-AI067854); NINDS Award# RC2NS070344; NIH U54 NS078059; NIH P01 HD080642; The J. Willard and Alice S. Marriott Foundation; The Muscular Dystrophy Association; The Nicholas Nunno Foundation; The JDM Fund for Mitochondrial Research; The Arturo Estopinan TK2 Research Fund ; NIMH Grant RC2MH089915; UCB; Epi4K SBB and Admin Core grant; Epi4K Gene Discovery in Epilepsy study (NINDS U01-NS077303) and the Epilepsy Genome/Phenome Project (EPGP - NINDS U01-NS053998); ALS Association (National and the Greater New York Chapter); Biogen; Ellison Medical Foundation New Scholar award AG-NS-0441-08; National Institute Of Mental Health of the National Institutes of Health (K01MH098126); B57 SAIC-Fredrick Inc. M11-074; OCD Rare 1R01MH097971-01A1; Duke Chancellor's Discovery Program Research Fund 2014; NIAID/NIH Grant# 1R56AI098588-01A1; American Academy of Child and Adolescent Psychiatry (AACAP) Pilot Research Award; NIMH Grant RC2MH089915; Endocrine Fellows Foundation Grant; the NIH Clinical and Translational Science Award Program (UL1TR000040); NIH U01HG007672; the Washington Heights Inwood Columbia Aging Project (R01AG037212, P01AG007232, PO1AG07232, R01AG037212, RF1AG054023); the Stanley Institute for Cognitive Genomics at Cold Spring Harbor Laboratory; New York-Presbyterian Hospital; the Columbia University College of Physicians and Surgeons; the Columbia University Medical Center.

Members of the Undiagnosed Diseases Network.

Maria T. Acosta, Margaret Adam, David R. Adams, Pankaj B. Agrawal, Mercedes E. Alejandro, Justin Alvey, Laura Amendola, Ashley Andrews, Euan A. Ashley, Mahshid S. Azamian, Carlos A. Bacino, Guney Bademci, Eva Baker, Ashok Balasubramanyam, Dustin Baldridge, Jim Bale, Michael Bamshad, Deborah Barbouth, Pinar Bayrak-Toydemir, Anita Beck, Alan H. Beggs, Edward Behrens, Gill Bejerano, Jimmy Bennet, Beverly Berg-Rood, Jonathan A. Bernstein, Gerard T. Berry, Anna Bican, Stephanie Bivona, Elizabeth Blue, John Bohnsack, Carsten Bonnenmann, Devon Bonner, Lorenzo Botto, Brenna Boyd, Lauren C. Briere, Elly Brokamp, Gabrielle Brown, Elizabeth A. Burke, Lindsay C. Burrage, Manish J. Butte, Peter Byers, William E. Byrd, John Carey, Olveen Carrasquillo, Ta Chen Peter Chang, Sirisak Chanprasert, Hsiao-Tuan Chao, Gary D. Clark, Terra R. Coakley, Laurel A. Cobban, Joy D. Cogan, Matthew Coggins, F. Sessions Cole, Heather A. Colley, Cynthia M. Cooper, Heidi Cope, William J. Craigen, Andrew B. Crouse, Michael Cunningham, Precilla D'Souza, Hongzheng Dai, Surendra Dasari, Joie Davis, Jyoti G. Dayal, Matthew Deardorff, Esteban C. Dell'Angelica, Shweta U. Dhar, Katrina Dipple, Daniel Doherty, Naghmeh Dorrani, Argenia L. Doss, Emilie D. Douine, David D. Draper, Laura Duncan, Dawn Earl, David J. Eckstein, Lisa T. Emrick, Christine M. Eng, Cecilia Esteves, Marni Falk, Liliana Fernandez, Carlos Ferreira, Elizabeth L. Fieg, Laurie C. Findley, Paul G. Fisher, Brent L. Fogel, Irman Forghani, Laure Fresard, William A. Gahl, Ian Glass, Bernadette Gochuico, Rena A. Godfrey, Katie Golden-Grant, Alica M. Goldman, Madison P. Goldrich, David B. Goldstein, Alana Grajewski, Catherine A. Groden, Irma Gutierrez, Sihoun Hahn, Rizwan Hamid, Neil A. Hanchard, Kelly Hassey, Nichole Hayes, Frances High, Anne Hing, Fuki M. Hisama, Ingrid A. Holm, Jason Hom, Martha Horike-Pyne, Alden Huang, Yong Huang, Laryssa Huryn, Rosario Isasi, Fariha Jamal, Gail P. Jarvik, Jeffrey Jarvik, Suman Jayadev, Lefkothea Karaviti, Jennifer Kennedy, Dana Kiley, Isaac S. Kohane, Jennefer N. Kohler, Deborah Krakow, Donna M. Krasnewich, Elijah Kravets, Susan Korrick, Mary Koziura, Joel B. Krier, Seema R. Lalani, Byron Lam, Christina Lam, Grace L. LaMoure, Brendan C. Lanpher, Ian R. Lanza, Lea Latham, Kimberly LeBlanc, Brendan H. Lee, Hane Lee, Roy Levitt, Richard A. Lewis, Sharyn A. Lincoln, Pengfei Liu, Xue Zhong Liu, Nicola Longo, Sandra K. Loo, Joseph Loscalzo, Richard L. Maas, John MacDowall, Ellen F. Macnamara, Calum A. MacRae, Valerie V. Maduro, Marta M. Majcherska, Bryan C. Mak, May Christine V. Malicdan, Laura A. Mamounas, Teri A. Manolio, Rong Mao, Kenneth Maravilla, Thomas C. Markello, Ronit Marom, Gabor Marth, Beth A. Martin, Martin G. Martin, Julian A. Martínez-Agosto, Shruti Marwaha, Jacob McCauley, Allyn McConkie-Rosell, Colleen E. McCormack, Alexa T. McCray, Elisabeth McGee, Heather Mefford, J. Lawrence Merritt, Matthew Might, Ghayda Mirzaa, Eva Morava, Paolo M. Moretti, Deborah Mosbrook-Davis, John J. Mulvihill, David R. Murdock, Anna Nagy, Mariko Nakano-Okuno, Avi Nath, Stan F. Nelson, John H. Newman, Sarah K. Nicholas,

Deborah Nickerson, Shirley Nieves-Rodriguez, Donna Novacic, Devin Oglesbee, James P. Orengo, Laura Pace, Stephen Pak, J. Carl Pallais, Christina GS. Palmer, Jeanette C. Papp, Neil H. Parker, John A. Phillips III, Jennifer E. Posey, Lorraine Potocki, Bradley Power, Barbara N. Pusey, Aaron Quinlan, Wendy Raskind, Archana N. Raja, Deepak A. Rao, Genecee Renteria, Chloe M. Reuter, Lynette Rives, Amy K. Robertson, Lance H. Rodan, Jill A. Rosenfeld, Natalie Rosenwasser, Francis Rossignol, Maura Ruzhnikov, Ralph Sacco, Jacinda B. Sampson, Susan L. Samson, Mario Saporta, C. Ron Scott, Judy Schaechter, Timothy Schedl, Kelly Schoch, Daryl A. Scott, Vandana Shashi, Jimann Shin, Rebecca Signer, Edwin K. Silverman, Janet S. Sinsheimer, Kathy Sisco, Edward C. Smith, Kevin S. Smith, Emily Solem, Lilianna Solnica-Krezel, Ben Solomon, Rebecca C. Spillmann, Joan M. Stoler, Jennifer A. Sullivan, Kathleen Sullivan, Angela Sun, Shirley Sutton, David A. Sweetser, Virginia Sybert, Holly K. Tabor, Amelia L. M. Tan, Queenie K.-G. Tan, Mustafa Tekin, Fred Telischi, Willa Thorson, Audrey Thurm, Cynthia J. Tifft, Camilo Toro, Alyssa A. Tran, Brianna M. Tucker, Tiina K. Urv, Adeline Vanderver, Matt Velinder, Dave Viskochil, Tiphanie P. Vogel, Colleen E. Wahl, Stephanie Wallace, Nicole M. Walley, Chris A. Walsh, Melissa Walker, Jennifer Wambach, Jijun Wan, Lee-kai Wang, Michael F. Wangler, Patricia A. Ward, Daniel Wegner, Mark Wener, Tara Wenger, Katherine Wesseling Perry, Monte Westerfield, Matthew T. Wheeler, Jordan Whitlock, Lynne A. Wolfe, Jeremy D. Woods, Shinya Yamamoto, John Yang, Muhammad Yousef, Diane B. Zastrow, Wadih Zein, Chunli Zhao, Stephan Zuchner.

METHODS CITATIONS

1. Primary nephrotic syndrome in children: Clinical significance of histopathologic variants of minimal change and of diffuse mesangial hypercellularity. *Kidney Int.* **20**, 765–771 (1981).
2. 1000 Genome Project Data Processing Subgroup *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
3. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
4. MacArthur, D. G. *et al.* Guidelines for investigating causality of sequence variants in human disease. *Nature* **508**, 469 (2014).
5. Vivante, A. & Hildebrandt, F. Exploring the genetic basis of early-onset chronic kidney disease. *Nat. Rev. Nephrol.* **12**, 133 (2016).
6. van der Ven, A. T. *et al.* Whole-Exome Sequencing Identifies Causative Mutations in Families with Congenital Anomalies of the Kidney and Urinary Tract. *J. Am. Soc. Nephrol.* **29**, 2348 (2018).
7. REESE, M. G., EECKMAN, F. H., KULP, D. & HAUSSLER, D. Improved Splice Site Detection in Genie. *J. Comput. Biol.* **4**, 311–323 (1997).
8. Yeo, G. & Burge, C. B. Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. *J. Comput. Biol.* **11**, 377–394 (2004).
9. Desmet, F.-O. *et al.* Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37**, e67–e67 (2009).
10. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886–D894 (2018).
11. Schneider, G. *et al.* SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* **40**, W452–W457 (2012).
12. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* **11**, 361 (2014).

13. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248 (2010).
14. Wang, M. *et al.* Contributions of Rare Gene Variants to Familial and Sporadic FSGS. *J. Am. Soc. Nephrol.* **30**, 1625 (2019).
15. Kremer, L. S. *et al.* Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat. Commun.* **8**, 15824 (2017).
16. Griffin, H. R. *et al.* Accurate mitochondrial DNA sequencing using off-target reads provides a single test to identify pathogenic point mutations. *Genet. Med.* **16**, 962–971 (2014).
17. Plagnol, V. *et al.* A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* **28**, 2747–2754 (2012).
18. Yang, Y. *et al.* Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders. *N. Engl. J. Med.* **369**, 1502–1511 (2013).
19. Beaulieu, C. L. *et al.* FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. *Am. J. Hum. Genet.* **94**, 809–817 (2014).
20. Flygare, S. *et al.* The VAAST Variant Prioritizer (VVP): ultrafast, easy to use whole genome variant prioritization tool. *BMC Bioinformatics* **19**, 57 (2018).
21. Niederriter, A. R. *et al.* In vivo modeling of the morbid human genome using Danio rerio. *J. Vis. Exp. JoVE* e50338–e50338 (2013) doi:10.3791/50338.
22. Saleem, M. A. *et al.* A Conditionally Immortalized Human Podocyte Cell Line Demonstrating Nephrin and Podocin Expression. *J. Am. Soc. Nephrol.* **13**, 630 (2002).
23. Gee, H. Y. *et al.* ARHGDIA mutations cause nephrotic syndrome via defective RHO GTPase signaling. *J. Clin. Invest.* **123**, 3243–3253 (2013).
24. Rao, J. *et al.* Advillin acts upstream of phospholipase C ϵ1 in steroid-resistant nephrotic syndrome. *J. Clin. Invest.* **127**, 4257–4269 (2017).

25. Gee, H. Y. *et al.* FAT1 mutations cause a glomerulotubular nephropathy. *Nat. Commun.* **7**, 10822 (2016).
26. Ishibashi, R. *et al.* A novel podocyte gene, semaphorin 3G, protects glomerular podocyte from lipopolysaccharide-induced inflammation. *Sci. Rep.* **6**, 25955 (2016).
27. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272–W276 (2016).
28. Zhu, X. *et al.* An Efficient Genotyping Method for Genome-modified Animals and Human Cells Generated with CRISPR/Cas9 System. *Sci. Rep.* **4**, 6420 (2014).
29. Karaiskos, N. *et al.* A Single-Cell Transcriptome Atlas of the Mouse Glomerulus. *J. Am. Soc. Nephrol.* **29**, 2060 (2018).
30. Magella, B. *et al.* Cross-platform single cell analysis of kidney development shows stromal cells express Gdnf. *Dev. Biol.* **434**, 36–47 (2018).
31. Menon, R. *et al.* Single-cell analysis of progenitor cell dynamics and lineage specification in the human fetal kidney. *Development* **145**, dev164038 (2018).
32. Park, J. *et al.* Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758 (2018).
33. Lindström, N. O. *et al.* Progressive Recruitment of Mesenchymal Progenitors Reveals a Time-Dependent Process of Cell Fate Acquisition in Mouse and Human Nephrogenesis. *Dev. Cell* **45**, 651-660.e4 (2018).
34. Samocha, K. E. *et al.* A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.* **46**, 944–950 (2014).
35. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, (2015).

36. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).