Award Number: W81XWH-16-1-0451

TITLE: Multispecies, Integrative GWAS for Focal Segmental Glomerulosclerosis

PRINCIPAL INVESTIGATOR: **Simone Sanna-Cherchi**

CONTRACTING ORGANIZATION: The Trustees of Columbia University in the City of New York
New York, NY

REPORT DATE: December 2020

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
Decenber 2020	Final	15Aug2016-14Aug2020

4. TITLE AND SUBTITLE

Multispecies, Integrative GWAS for Focal Segmental Glomerulosclerosis

5a. CONTRACT NUMBER
W81XWH-16-1-0451

5b. GRANT NUMBER
W81XWH-16-1-0451,PR151419P1

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Simone Sanna-Cherchi, Ali G. Gharavi

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Trustees of Columbia University in the City of New York
630 W. 168th St, FL4
New York, NY
10032-3725

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research
And Material Command
Fort Detrick, Maryland
21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT

14. ABSTRACT
Focal Segmental Glomerulosclerosis (FSGS) is one of the most common causes of nephrotic syndrome and several studies reported an increase of FSGS diagnosis with up to 18.7% and 47% of cases in adults and children receiving a kidney biopsy, respectively. From a clinical perspective, idiopathic FSGS is characterized by high morbidity, poor response to medical therapy, and high rate of progression to end-stage renal disease requiring dialysis or transplantation. We approached this phenotype in a comprehensive fashion using both human genetics and mouse studies. In a human GWAS of FSGS, we detected genome-wide significant signals in the in the HLA and APOL1 loci, as well as in multiple novel loci after analysis of subgroups of FSGS stratified by age of onset and response to immunosuppressive therapy. In a mouse model of collapsing FSGS, we detected a signal on Chr13. Fine mapping studies and bioinformatics analyses suggested Ssbp2 as the lead candidate. Analysis of Ssbp2 null mice demonstrated that they develop spontaneous FSGS with aging. These data identify multiple risk loci and candidate genes for FSGS in human and mice and implicate new pathways in disease pathogenesis.

15. SUBJECT TERMS
None listed.

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	10	**19b. TELEPHONE NUMBER** *(include area code)*

TABLE OF CONTENTS **Page**

1. INTRODUCTION:

Focal Segmental Glomerulosclerosis (FSGS) is one of the most common causes of nephrotic syndrome and several studies reported an increase of FSGS diagnosis with up to 18.7% and 47% of cases in adults and children receiving a kidney biopsy, respectively[1] . From a clinical perspective, idiopathic FSGS is characterized by high morbidity, poor response to medical therapy, and high rate of progression to end-stage renal disease requiring dialysis or transplantation. We approached this phenotype in a comprehensive fashion using both human genetics and mouse studies. In the original grant, we pursued a multispecies GWAS with the following aims:

Specific aim 1: A Genome-wide association study (GWAS) for common single nucleotide polymorphisms and rare copy number variations in 7,559 FSGS and over 50,000 controls

Specific aim 2: A GWAS for FSGS in a mouse model of collapsing FSGS

Specific aim 3. Cross annotation between human and mouse GWAS and identification of downstream dysregulated pathways and networks

2. KEYWORDS: Provide a brief list of keywords (limit to 20 words).

APOL1 apolipoprotein 1
CKD chronic kidney disease
DO Diversity Outcross strains
ESRD End stage renal disease
FSGS focal segmental glomerulosclerosis
GWAS Genome wide association study
HIVAN HIV- associated nephropathy
HLA human leukocyte antigen
NS nephrotic syndrome
QTL quantitative trait loci
SNP single nucleotide polymorphism
SNV single nucleotide variant
Ssbp2 single stranded DNA binding protein 2
TgFVB Transgenic mouse with the proviral HIV genome with deletion of the gag and pol sequences in the FVB background
WES Whole exome sequencing

3. ACCOMPLISHMENTS: The PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction.

What were the major goals of the project?

Major task 1. A Genome-wide association study for common single nucleotide polymorphisms and rare copy number variations in 7,559 FSGS and over 50,000 controls.

- Genotyping completed for 75% of cases and 100% of controls.
- GWAS analysis completed and manuscript in preparation.
- CNV analysis completed and manuscript in preparation

Major task 2. Generate 1000 F1 hybrids between TgFVB and >30 inbred mice strains and conduct a GWAS of clinical and molecular phenotypes to identify genes and loci predisposing to FSGS in mice.

- F1 hybrid GWAS completed at 90%

- GWAS analysis completed and manuscript submitted for publication
-

Major task 3. Cross annotation between human and mouse GWAS and identification of downstream dysregulated pathways and networks.

- Cross annotation of loci completed, partially included in the mouse GWAS manuscript already submitted, partially included in the human GWAS paper, currently in preparation

What was accomplished under these goals?

During this award we made significant progress towards the completion of the proposed aims. Here we describe the progress most directly related to the aims of the original proposal.

1. Collection of the largest NS cohort worldwide with comprehensive genetic data: We have collected a cohort of more than 12,000 independent patients with FSGS (Table 1), and over 6,000 family members. DNA is available for all cases. Of these, 7,236 have already been subjected to Illumina SNP genotyping (MEGA arrays), and 7,149 to whole-exome sequencing (WES) or whole-genome sequencing (WGS). Hence minimal funding will be requested to the DoD for generating remaining genomic data genomic **data.**

Cohort	Pediatric FSGS (N)	Adult FSGS (N)	Total (N)	SNP Array	WES/WGS
Columbia Cohort: Sanna-Cherchi	1,326	1,314	**2,640**	2,527	2,527
NEPTUNE Cohort: Sampson	180	220	**400**	400	400
GICKD Cohort: Sampson	230	29	**259**	259	0
Harvard Cohort: Hildebrandt	3,600	400	**4,000**	1,300	2,000
Harvard Cohort: Pollak	200	1,800	**2,000**	650	622
Duke Cohort: Gbadegesin	1,100	450	**1,550**	900	400
Total	**6,305**	**3,701**	**10,849**	**5,636**	**5,949**
Additional Cohort: CureGN – Kretzler/Kiryluk	400	800	**1,200**	1,200	1,200
Total	**7,036**	**4,863**	**12,049**	**7,236**	**7,149**

Table 1. Cohorts for genetic studies on FSGS. The PI accrued a cohort of 12,049 FSGS patients including 1,200 cases from the CureGN cohort. Approximately 60% of patients have been already subjected to genome-wide Illumina SNP genotyping and WES/WGS.

Importantly, 1,200 patients with nephrotic syndrome from the NIH-funded CureGN Study are available to us for replication and joint analyses. The CureGN cohort has been subjected to Illumina SNP genotyping, whole genome sequencing (WGS), and blood RNA-Seq, and primary data is available to. We have access to glomerular and tubulointerstitial RNA-Seq datasets on ~350 patients with FSGS. We also have access to over 100,000 controls with WES/WGS data through the Institute for Genomic Medicine at CUMC.

2. Multiethnic GWAS identifies known and novel associations for adult and pediatric FSGS (*in preparation*): We conducted a GWAS comparing 3,434 FSGS cases and 26,020 controls. The FSGS cohort was composed of four major subpopulations: Europeans (N=2,280), Africans (N=426), Admixed Americans (N=528), and Asians (N=200). Therefore, we performed both ancestry specific and trans-ethnic GWAS. The raw genotyped data was first imputed using TOPMed reference panel. Thereafter, genotype-phenotype associations were performed using additive logistic mixed model R package SAIGE. Trans-ethnic association was performed using fixed-effect model.

In the combined meta-analysis (**Figure 1**; 3,434 cases), we discovered significant associations for *APOL1* (OR= 3.49, P= 9.75x10^{-43}), that was driven by African and Admixed individuals, the *HLA-DQB1* locus (OR=2, P= 6.35x10^{-16}), one novel locus on chromosome 15 (*ADAMTSL3*; OR=1.53, P=2.73x10^{-8}), and one suggestive new association on chromosome 2 (*AC068276.9*; OR=1.21, P=4.01x10^{-7}). After removal of cases harboring pathogenic mutations causing Mendelian forms of FSGS (N=268, resulting in 3,166 unsolved cases), we observed stronger association for the newly identified

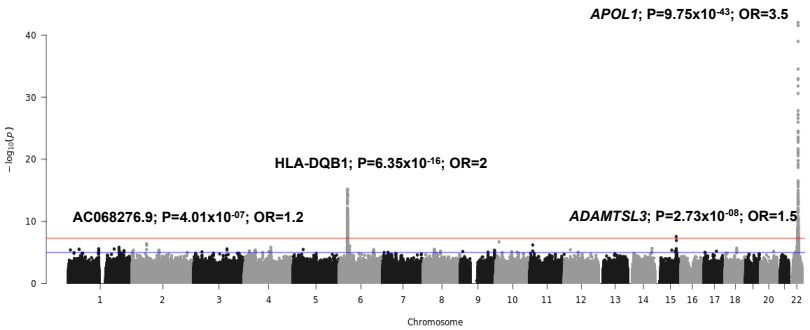

Figure 1. Manhattan plot of GWAS in the entire cohort (3,434 cases vs 26,020 controls) showing highly significant association at the *HLA* and *APOL1* loci, and two novel associations on chromosomes 2 and 15.

signals (*ADAMTSL3*; OR=1.58, P=9.66x10^{-9}; *AC068276.9*; OR=1.22, *P*=1.76x10^{-7}), suggesting true positive associations. We also conducted meta-analysis between the two-super population with significant African ancestry (Africans and Admixed Americans, totaling 846 cases & 20,645 controls). We identified a novel significant association, beyond *APOL1*, on chromosome 9 (*AL353742.1*; OR=1.41, P=2.93x10^{-8}). Furthermore, ancestry-specific and sub-phenotype analyses according to age of onset and response to corticosteroid treatment identified multiple significant and suggestive signals. For example, we identified a novel locus on chromosome 1 (*AL161637.1*; OR=1.27, *P*=4.38x10^{-8}) for European ancestry (2,280 cases and 4,398 controls). Interestingly, we also observed significant association on chromosome 6 (*HLA-DQA1*; OR= 1.28, *P*=6.43x10^{-9}) for steroid resistant pediatric patients (445 cases & 4,398 controls) suggesting immune-related mechanisms converging to FSGS. In Admixed Americans, we identified a suggestive signal on chromosome 5 (*AC093298.2*; OR=2.67, *P*=5.15x10^{-8}) for steroid resistant patients (310 cases & 13,300 controls) and on chromosome 16 (*DNAH3*; OR= 2.39, *P*=9.5x10^{-8}) for steroid sensitive patients (155 cases & 13,300 controls). In conclusion, our results demonstrate multiple novel FSGS loci, pleiotropic risk alleles that predispose to NS across different sub phenotypes, and signals specific to race, age of onset and response to therapy. Importantly, we can now implicate variation in the *HLA* locus in children who have steroid resistant FSGS but do not carry highly penetrant mutations in Mendelian NS genes, thus potentially identifying a population that might benefit from aggressive immunosuppression, and pre- and post- transplant plasma exchange to prevent recurrence. Fine-mapping the HLA and non-HLA risk loci and integrating these GWAS alleles with other FSGS-associated genetic factors (Mendelian alleles, CNVs) holds promise in further elucidating the genetic architecture of FSGS. These results are included in a manuscript that will be submitted in the next 3 months and are the backbone of a DoD expansion award that is currently under review for funding.

3. Mapping genetic susceptibility to collapsing FSGS in mouse model (*manuscript submitted for publication*). Although mice do not possess an *APOL1* ortholog, HIV-1 transgenic mice on the FVBN/J background (TgFVB) display virtually all the clinical and molecular features of collapsing variant of FSGS, indicating that alternative genetic lesions, in the absence of *APOL1*, can predispose to this disease (**Figure 2**). To map genes for glomerulosclerosis, we generated F1 hybrids between TgFVB and 20 inbred strains (over 750 F1 hybrids generated). F1 hybrids between 129S1/SvImJ, Balb/CJ, C57BL/6J, C57BL/6NJ, C57BL/10J, C57BL/J, C58/J and CAST/EiJ are completely protected from disease, while the other F1 hybrids have evidence of disease (**Figure 2**). To map loci predisposing to collapsing FSGS, we performed a GWAS using a mixed linear model method with 160,000 SNPs, searching for haplotype distribution patterns that matched the high/low strain susceptibility pattern. This search identified a major locus on Chr 13 which is genome-wide significant (p=1x10-6, rs48893293, **Figure 3**). Remarkably, the top signal falls within a QTL, *HIVAN2*, which we had previously mapped using segregating crosses. Notably, the Chr 13 interval contains the *Ssbp2* gene, encoding a DNA binding protein that stabilizes transcriptional complexes by prevent proteosomal degradation. *Ssbp2* is highly expressed in podocytes based on recent single cell transcriptomic data published by the Susztak and McMahon labs (**Figure 4**). In addition, it has been shown to interact *LMX1B* a known FSGS genes, in yeast two hybrid assay. Finally, *Ssbp2* null mice have been previously described and have been shown

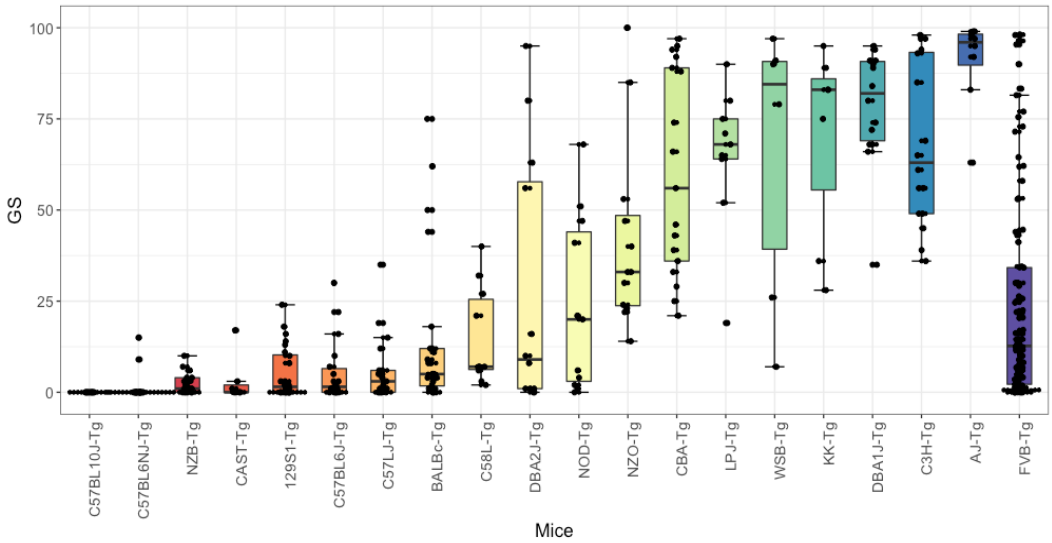

Figure 2: Significant interstrain variation in glomerulosclerosis in 20 inbred strains (N=10-20 per strain). There were no differences by sex.

to develop cancer at multiple sites, but also develop chronic glomerulopathy. Finally, there is a cis-eQTL for *Sspb2* in this interval in our F2 cross. There are two other genes in the vicinity, *Agt10* and *Acot12*, but they are not highly expressed in podocytes and have no known involvement in kidney disease. Taken together, these data strongly suggest that *Ssbp2* is the culprit gene producing susceptibility to HIVAN in the mouse.

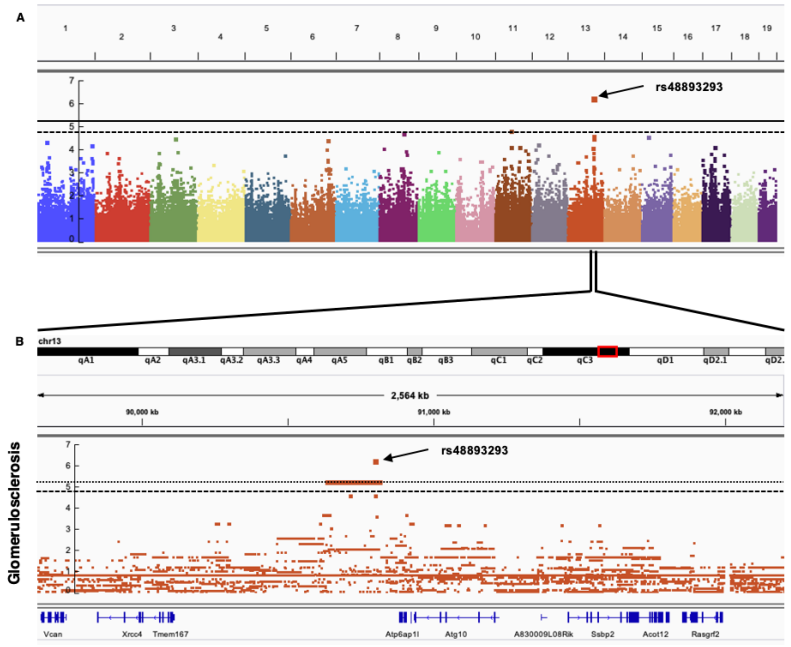

Figure 3. A GWAS of F1 hybrids identifies a signal on Chr13 with coincides with the HIVAN2 locus. **B.** The region contains 3 candidate genes, *Agt10*, *Ssbp2* and *Acot12*

How do studies of animal models viral collapsing glomerulopathy inform about FSGS. FSGS can arise due to developmental, environmental, as well as genetic defects. The HIVAN mouse model of FSGS is informative not only about HIV induced kidney injury but also about APOL1-assocaited glomerulopathies which contribute to a substantial proportion of kidney failure among African Americans. In addition, we have recently reported multiple cases of FSGS-collapsing glomerulopathy in the setting of SARS-CoV-2 in patients with high risk *APOL1* genotypes[58-61]. Therefore, these studies can also illuminate the pathogenesis kidney disease due to SARS-CoV-2 infection, which occurs in 30% of ICU cases. Elucidating novel candidate genes and dissecting mechanisms of glomerular damage will provide critical information for the design of novel therapy, and patient treatment and management.

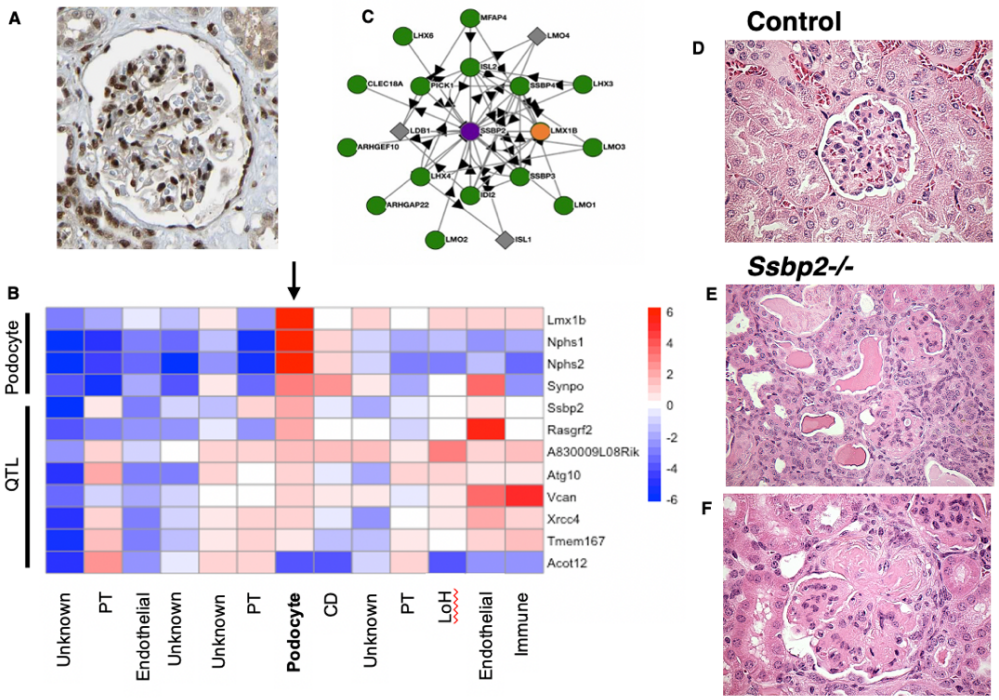

Figure 6. A SSBP2 is expressed in human podocytes (Protein Atlas). B. Single cell sequencing of kidney revealed 13 populations, and *Sspb2* is expressed in the population containing podocyte genes *Nphs1*, *Nphs2* and *Lmx1b,* whose orthologs are known genes for FSGS**. C**. Interrogation of Bioplex data identifies an interaction between SSBP2 and *LMX1B.* D. Histology slide of WT control mice. E. and F. histology slides of *Sspb2* null mice showing FSGS lesions with tubular casts, focal tubular atrophy and interstitial fibrosis. PT – proximal tubule, LoH – Loup of Henle, CD – collecting duct

Rare variant gene burden analysis identifies a candidate modifier signal for *APOL1* nephropathy

Through this award, we studied a large number of patients with FSGS and contributed samples to a collaborative study aiming to discover rare genetic variants modifying the development to CKD[15]. In this study, we identified 137 CKD cases with 2 *APOL1* risk variants. Comparing the 137 CKD cases with the *APOL1* risk genotypes to 2,209 ancestry-matched controls, we identified a suggestive signal in the *AT-Hook DNA Binding Motif Containing 1* gene (*AHDC1*; MIM:615790) in the rare non-benign model (p = 5.9x10-7; OR 14.0 [95% CI 5.0 – 37.9]). This

association was driven by nonsynonymous and indel variants. *AHDC1* encodes a nuclear protein with a predicted DNA binding domain that is ubiquitously expressed. *De novo* mutations cause Xia-Gibbs syndrome, an intellectual disability syndrome that also features sleep apnea and various dysmorphologies (OMIM 615829). The function of *AHDC1* is not completely understood, but this suggestive signal encourages further investigation to confirm this association, and, if replicated, to examine the potential links to *APOL1* biology. We are currently studying larger cohorts both for rare and common variants via exome-wide collapsing analysis and GWAS, respectively, as part of a new expansion award proposal to the DoD.

What opportunities for training and professional development has the project provided?

Dr. Dina Ahram. Dr. Ahram, PhD, joined my lab at the end of 2016 and she is currently a postdoc scientist in my lab. Dr. Ahram has been working on several projects aimed at gene identification for FSGS using genome-wide association analysis (GWAS) for common variants, as well as modern collapsing rare variant association studies suing exome sequencing data from large populations. She rapidly developed an interest for clinical genetics and she successfully applied to a fellowship in molecular genetic pathology and clinical molecular genetics (Director Vimla S. Aggarwal) in the Laboratory of Personalized Genomic Medicine at CUMC. She is dedicating 50% effort to research and 50% to her clinical genetic training, which she will complete in the summer of 2021. She coauthored six publications (related or unrelated to this grant): Am J Hum Genet (2021, in press), Am J Hum Genet (2020), Am J Kidney Dis (2020), Clin J Am Soc Nephrol (2020), Hum Mutat (2018), Nat Genet (2019).

Tze Y. Lim. Ms. Lim, MS, joined my lab in 2017 as lead bioinformatician. Ms. Lim has been working and helping postdoc and scientists in the lab with the computational biology component of projects related to this grant: analysis of exome sequencing, DNA microarrays for GWAS, RNA sequencing etc. Ms. Lim has also been heavily involved in training with the ultimate goal of proceeding to graduate school in order to obtain a PhD degree in computational biology/genetics applied to urinary tract malformations. During the tenure of this award she attended several courses, including: The Advanced Sequencing Technology & Applications, Cold Spring Harbor Laboratory (2019); The Medical Genetics for Internal Medicine "How to Integrate Genetics into Internal Medicine Practice" by The Center for Precision Medicine and Genomics course (October 2020); and the Getting started with statistical software, Columbia University Department of Biostatistic (November 2020). She coauthored three publications: Am J Hum Genet (2021, in press), Am J Hum Genet (2020), Nat Genet (2019).

All trainees participated and presented original work at national and international meetings, including the American Society of Nephrology Kidney Week (2015, 2016, 2017, and 2018); the American Society of Human Genetics (2017, 2018); the CAIRIBU meeting (2019, 2020); and others.

How were the results disseminated to communities of interest?

Nothing to Report.

4. IMPACT:

At the time of writing, we are 3 years into the funding cycle. To date, we have presented multiple abstracts at the American Society of Nephrology and the American Society of Human Genetics; we have one manuscript published, one in press, two submitted, and two in advanced phase of preparation that are deriving from this award. The work we have undertaken is comprehensive in scope and we have a preference to publish complete stories rather than incremental gains. Therefore, we anticipate publications reporting the bulk of the work resulting from this award in the coming year.

Our work has direct impact on understanding the etiology and genetic architecture of FSGS and, more broadly, chronic kidney disease.
Moreover, in our discovery of TRIM8 mutation in pediatric patients with FSGS, developmental delay, and epilepsy, we uncovered an unexpected role of nuclear bodies and subnuclear liquid phase biology both in the pathogenesis of FSGS and developmental brain disease.

What was the impact on technology transfer?

Nothing to Report.

What was the impact on society beyond science and technology?

Nothing to Report.

5. CHANGES/PROBLEMS:

The major challenges were essentially all logistical. For the human genetic studies, samples are now all in hand, but due to the budgeting structure, we had to perform the genotyping in batches across the 3 years, and are still waiting for the final batch to complete the human GWAS. Nonetheless the interim analysis is very promising as it reveals multiple novel associations for human FSGS, with HLA locus and other immune as well as structural/developmental loci for the first time, indicating a strong immune mediated component to disease. A second challenge arose for the procurement of mice. Our original plans for the mouse GWAS involved analysis of diversity outcross mice but because these mice were still under development at the initiation of the award, this created major delays. We finally had to switch plans and perform analysis of classic inbred strains (after consultation and approval from DOD program officials). This turned to our advantage as the mouse GWAS was highly informative. We now can turn our attention to validation of GWAS signals by generating and analyzing Sspb2 null mice.

6. PRODUCTS:

Publications and Meeting Presentations

1. Cameron-Christie S, Wolock C, Groopman E, Petrovski S, Kamalakaran S, Povysil G, Vitsios D, Zhang M, Fleckner J, March RE, Gelfman S, Marasa M, Li Y, **Sanna-Cherchi S**, Kiryluk K, Allen AS, Fellström B, Haefliger C, Platt A, Goldstein DB, **Gharavi AG**. Exome-based rare-variant analyses in chronic kidney disease. *J Am Soc Nephrol* 2019 June; 30(6): 1109-1122

2. Steers NJ, Na YJ, DeMaria ND, Lam WY, D'Agati VD, **Gharavi AG**. "Interstrain variation in severity of nephropathy and immunoglobulin levels in HIV-1 transgenic mice". American Society of Nephrology Kidney Week 2018

3. HLA Alleles Confer Risk to Primary Idiopathic Nephrotic Syndrome in Individuals of Caucasian Ancestry. Ahram D, Gilles C, Mitrotti A, …. **Gharavi AG**, Hildebrandt F, Sampson MG, **Sanna-Cherchi S**. ASN SA-PO593. American Society of Nephrology Kidney Week, 2017

4. Steers N, Gupta Y, DeMaria ND, Lim TY, Lam WY, Mo A, Liang J. Stevens K., D'Agati VD, **Sanna-Cherchi S**, **Gharavi AG**. GWAS in the mouse identifies loci for collapsing glomerulopathy (submitted for publication).

5. Weng PL, Majmundar AJ, Khan K, … , **Gharavi AG**, Jobanputra V, Pierce-Hoffman E, O'Donnell-Luria A, Rehm HL, Mane S, D'Agati VD, Pollak MR, Ghiggeri GM, Lifton RP, Goldstein DB, Davis E, Hildebrandt F, **Sanna-Cherchi S**. *De novo* truncating *TRIM8* variants impair its protein localization to nuclear bodies and cause a pediatric syndrome of developmental delay, epilepsy and focal segmental glomerulosclerosis. *Am J Hum Genet 2021* (*in press*).

6. Marasa M, Ahram DF, Rehman AU, Mitrotti A, Abhyankar A, Jain NG, Weng PL, Piva SE, Fernandez HE, Uy NS, Chatterjee D, Kil BH, Nestor JG, Felice V, Robinson D, Whyte D, **Gharavi AG**, Appel GB, Radhakrishnan J, Santoriello D, Bomback A, Lin F, D'Agati VD, Jobanputra V, and **Sanna-Cherchi S**. Rapid genome sequencing to guide clinical decision making in patients with proteinuric kidney disease. American Society of Nephrology Kidney Week 2019 (submitted for publication).

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to Report.

What other organizations were involved as partners?
Nothing to Report.

8. SPECIAL REPORTING REQUIREMENTS

Nothing to Report.

9. APPENDICES:

Nothing to Report.