PNAS RESEARCH ARTICLE | GENETICS

# Genetic risk factors for Mesoamerican nephropathy

David J. Friedman[a,b,c,1], Dominick A. Leone[d], Juan José Amador[d,e], Joseph Kupferman[a,b], Lauren J. Francey[a], Damaris Lopez-Pilarte[e], Jorge Lau[f], Iris Delgado[d], W. Katherine Yih[b,g,h], Alejandro Salinas[f,2], Minxian Wang[c], Giulio Genovese[c], Shrijal Shah[a], Jessica Kelly[a], Calum F. Tattersfield[a], Nathan H. Raines[a,b], Magaly Amador[e], Leny Dias[a], Achilleas Pitsillides[i], Oriana Ramirez-Rubio[d,j], Alda G. Amador[e], Marissa Cortopassi[k], Katie M. Applebaum[l], Seth L. Alper[a,b,c], Alex S. Banks[b,k], Michael D. McClean[m], Jessica H. Leibler[m], Madeleine K. Scammell[m], Josée Dupuis[i,n], and Daniel R. Brooks[d,1]

Affiliations are included on p. 8.

Edited by Peter Agre, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD; received March 8, 2024; accepted September 14, 2024

**Mesoamerican nephropathy (MeN) is a progressive kidney disease found on the Pacific coast of Central America primarily in young male agricultural workers without typical kidney disease risk factors. While it is generally accepted that environmental exposures contribute to MeN, we hypothesized that there was also an important genetic component. We performed a genome-wide association study comparing individuals with MeN versus individuals with normal kidney function. We found that Native American ancestry was strongly associated with increased risk of MeN. We also identified candidate variants in the *OPCML* gene, which encodes a protein that binds opioid receptors, that were associated with ~sixfold reduced odds of MeN (allele frequency 0.029 in controls and 0.005 in cases, OR = 0.16; $P = 4 \times 10^{-8}$). Sugarcane workers with the protective *OPCML* variants had markedly increased urine osmolality suggesting greater ability to defend against hypovolemia. Experiments with *Opcml* knock-out mice revealed roles for *OPCML* in fluid balance and temperature regulation consistent with our findings in humans. Our data suggest that heritable differential sensitivity to heat stress and dehydration contributes to high rates of kidney disease in Central America.**

kidney disease | genetics | Mesoamerican | *OPCML*

## Significance

There is an extreme clustering of kidney disease in otherwise healthy young manual laborers in Central America known as Mesoamerican nephropathy (MeN). Environmental factors are the main suspect, particularly exposure to heavy labor in high heat. We found that this disease also has a strong genetic component. Native American ancestry is a strong risk factor for MeN. Genetic variants in a gene called *OPCML* may help protect against MeN, potentially by preventing dehydration and increasing tolerance to heat. Our data suggest that genetic factors are important for tolerance of heat stress and that populations have differential susceptibility to kidney injury based on these genetic factors.

Rates of chronic kidney disease (CKD) in Nicaragua and El Salvador are among the highest in the world, driven by exceptionally high rates of disease along the Pacific coast (1–3). CKD prevalence is highest in Northwestern Nicaragua and the Bajo Lempa region of El Salvador, with other hotspots of disease reported in Costa Rica, Guatemala, Honduras, Mexico, and Panama (1). Kidney disease in these high-risk regions disproportionately affects young (age < 45) male agricultural workers and other manual laborers, with rates at least 25 times higher than in age-matched US populations (3–5). This disease clustering in Pacific-coast agricultural regions, termed Mesoamerican nephropathy (MeN), is not associated with typical CKD risk factors such as diabetes or hypertension. Kidney biopsies show tubulointerstitial injury and proteinuria is minimal at disease onset (6, 7). The strong male preponderance of MeN and the high rates in some industries such as sugarcane farming, mining, and brickmaking have suggested occupational environmental risk factors. Most research has explored the role of exposures such as heat stress, agrichemicals, infectious agents, and heavy metals (8–11). However, definitive identification of the environmental contributors to MeN has remained elusive. Yet, all high-risk regions and occupations share the common requirement of strenuous activity in extreme heat.

Several observations suggest a role for genetic susceptibility in MeN. First, none of the environmental risk factors proposed to date including high heat exposure are unique to this geographical location (12, 13). Second, some laborers develop MeN after only brief periods of work in these high-risk occupations while others living in the same communities and working at the same jobs for several decades show no signs of disease. Published reports show that family history is a risk factor for MeN (14, 15). The most parsimonious explanation is that both genetic susceptibility and environmental exposures are important and that MeN disease prevalence is highest where both factors share the greatest geographic overlap. To test for the possibility of a genetic component of MeN, we performed a genome-wide association study (GWAS) in a region of Northwestern Nicaragua with extremely high MeN prevalence. We then explored the potential function of candidate GWAS hits in human biosamples and a mouse model.

## Results

Participants were recruited in Northwestern Nicaragua between 2014 and 2020. We enrolled 1) active sugarcane workers with no apparent kidney disease and former sugarcane workers with CKD who had worked at the same two sugarcane companies, 2) individuals

Author contributions: D.J.F., J. Kupferman, J.D., and D.R.B. designed research; D.J.F., D.A.L., J.J.A., J. Kupferman, L.J.F., D.L.-P., J.L., I.D., W.K.Y., A.S., M.W., G.G., S.S., J. Kelly, C.F.T., N.H.R., M.A., L.D., A.P., O.R.-R., A.A., M.C., K.M.A., S.L.A., A.S.B., M.D.M., M.K.S., J.D., and D.R.B. performed research; D.A.L. contributed new reagents/analytic tools; D.J.F., D.A.L., J.J.A., J. Kupferman, L.J.F., I.D., W.K.Y., M.W., G.G., J. Kelly, C.F.T., A.P., M.C., S.L.A., A.S.B., M.D.M., J.H.L., M.K.S., J.D., and D.R.B. analyzed data; and D.J.F., D.A.L., J.D., and D.R.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.



[1]To whom correspondence may be addressed. Email: dfriedma@bidmc.harvard.edu or danbrook@bu.edu.

[2]Deceased January 11, 2021.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2404848121/-/DCSupplemental.

Published November 25, 2024.

from a cohort study of brickmakers, 3) residents of a mining community, and 4) individuals from a longitudinal study in an occupationally diverse community. MeN was defined as elevation in serum creatinine in the absence of another cause of kidney disease (such as polycystic kidney disease, glomerulonephritis, or urinary tract abnormalities) and absence of risk factors including self-reported diabetes, random glucose >200 mg/dL, history of hypertension preceding kidney disease, HIV, or hepatitis B/C. Male cases had creatinine ≥1.6 mg/dL and controls had creatinine ≤1.2 mg/dL, while creatinine for female cases was ≥1.3 mg/dL and ≤1.0 mg/dL for controls. Additional details about the study population are included in the *Methods* section.

**Discovery Cohort.** For genome-wide association testing, our study design has aspects of an "extreme phenotype study," with controls who appear highly resistant to renal injury and cases who appear highly susceptible. For cases, we prioritized candidates with younger age and fewer years of work in high-risk jobs, thereby selecting for those most susceptible to renal injury. Conversely, for controls we prioritized individuals with normal kidney function despite older age and a longer work history, selecting for individuals who had not developed kidney disease despite prolonged occupational exposure. Demographic and laboratory data on the 429 cases and 385 controls are shown in *SI Appendix*, Table S1.

In our discovery cohort of 429 cases and 385 controls, we genotyped ~1.5 million SNPs on Illumina MEGA (Multi-Ethnic Global Array) genotyping arrays. We imputed additional SNPs using the Trans-Omics for Precision Medicine (TOPMED) imputation server (16). To evaluate population structure (PS), we tested the top 10 principal components (PC) for association with MeN case status. PC1 was strongly associated with case status ($P = 3.1 \times 10^{-7}$), so we adjusted for PC1 as a fixed effect in our statistical models. We incorporated an individual-level random effect, with a correlation structure specified by pairwise genetically determined kinship coefficients, to account for relatedness. We utilized a logistic mixed model (LMM) including PC to test for genetic associations between SNP dosages and odds of MeN. No variants reached genome-wide statistical significance ($P < 5 \times 10^{-8}$) after correction for ancestry and accounting for relatedness. Top loci (best $P = 1.8 \times 10^{-6}$) adjusted for PC1 are shown in *SI Appendix*, Table S2A.

PC1, which was strongly associated with case status, separated Native American from both European and African ancestry (ancestry information in *SI Appendix*, Fig. S1). We did observe a strong relationship between genome-wide ancestry proportions and MeN risk. Each additional 1% Native American ancestry increased the OR for case status by 1.046 (CI 1.028 to 1.064; Fig. 1*A*), which equates to a >twofold increase in risk comparing individuals 1 SD above to 1 SD below the mean of Native American ancestry, and a >fivefold risk comparing individuals 2 SD above to 2 SD below the mean.

**Replication Cohort and Meta-Analysis.** To further pursue genetic variants associated with MeN, we recruited another cohort of participants from the same general population (including sugarcane workers, brickmakers, and individuals from mining and diverse occupational communities) and using the same selection criteria (*SI Appendix*, Table S1). The top 1,000 variants from the discovery GWAS were tested in the replication cohort of 414 cases and 230 controls. Three SNPs within an intron of the *OPCML* gene showed association in the replication set with best $P = 9 \times 10^{-5}$ after adjustment for ancestry and relatedness, surpassing the LD-adjusted significance threshold of $1.3 \times 10^{-4}$

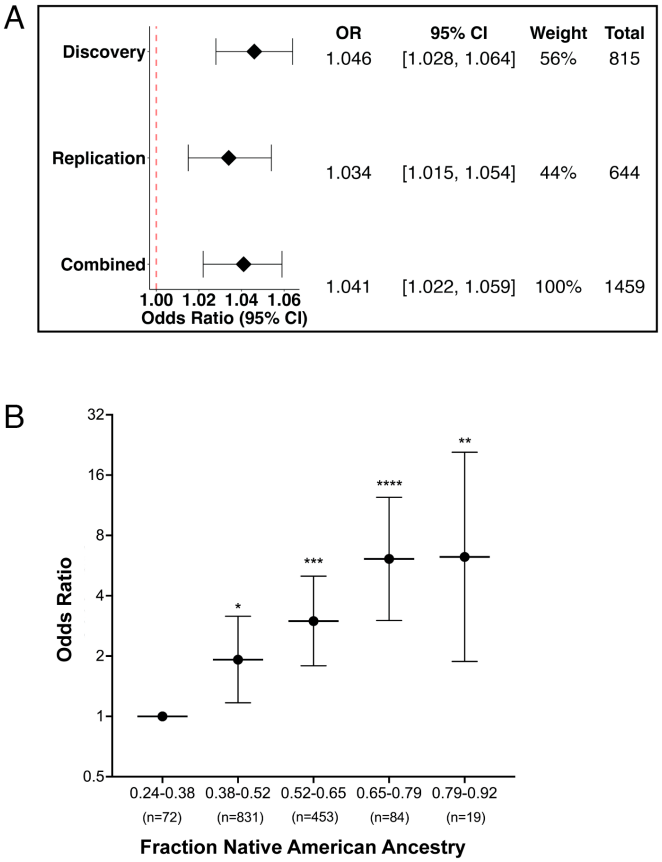


**Fig. 1.** Estimating the effect of Native American ancestry on MeN risk. (*A*) Increase in odds ratio (OR) for each additional 1% Native American ancestry for Discovery, Replication, and Combined datasets. In the combined dataset, each 10% increase (e.g., from 30 to 40%) in percent Native American ancestry translates into a ~1.5 fold increase in the OR of MeN. There was no heterogeneity between discovery and replication cohorts ($P = 0.38$). CI = confidence interval. (*B*) To assess for nonlinear effects of fraction Native American ancestry, the study population was divided into five evenly sized intervals by fraction Native American ancestry. The group with the lowest (0.24 to 0.38) Native American ancestry was the referent. With increasing Native American ancestry, group 2 OR was 1.92 (1.17 to 3.16; $P = 0.01$); group 3 OR, 2.99 (1.79 to 5.01; $P = 0.00002$); group 4 OR, 6.11 (3.01 to 12.39; $P = 2.6 \times 10^{-7}$); and group 5 OR, 6.25 (1.88 to 20.81; $P = 0.0017$).

(Table 1). The OR of the lead SNPs at this locus was 0.17 in the discovery set and 0.14 in the replication set. Joint analysis of the discovery and replication sets using METAL software (17) revealed the lead *OPCML* SNP was associated with MeN at $P = 4 \times 10^{-8}$ (2.9% in cases and 0.005% in controls: Table 2 and *SI Appendix*, Fig. S3 and Table S2B). Repeating the analyses with adjustment for PCs 1-10 instead of just PC1 slightly strengthened the meta-analysis association at the lead *OPCML* SNPs to $P = 2 \times 10^{-8}$. No associations at genome-wide significance were observed for variants on X, Y, and mitochondrial chromosomes in the joint analysis (*SI Appendix*, Tables S3–S6).

The replication cohort also demonstrated a strong correlation of both PC1 and Native American ancestry with case status ($P = 3 \times 10^{-4}$). In the full combined cohort, each additional 1% Native American ancestry increased the OR of MeN by 1.041 (Fig. 1*A* and *SI Appendix*, Fig. S2), translating to a ~1.5 fold increase in the OR for each 10% increase in percent Native American ancestry. To allow for nonlinear effects of the percent Native American ancestry on the risk of MeN, we binned all study participants into five evenly spaced intervals based on percent Native American ancestry from lowest to highest in our population. Compared to the group with

**Table 1. Replication of top SNPs from discovery set**

| Rank | Lead SNP | Chr. | position | Ref | Alt | Case AF | Control AF | OR | *P*-value | Nearest gene |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | rs59337101 | 11 | 132583931 | T | A | 0.006 | 0.033 | 0.14 | 9.5E-05 | *OPCML* |
| 2 | rs73588969 | 11 | 132591037 | C | T | 0.006 | 0.033 | 0.14 | 9.5E-05 | *OPCML* |
| 3 | rs73588951 | 11 | 132568249 | T | C | 0.006 | 0.033 | 0.14 | 9.5E-05 | *OPCML* |
| 4 | rs10820187 | 9 | 102460452 | T | C | 0.048 | 0.091 | 0.45 | 1.6E-03 | *LINC00587* |
| 5 | rs12380682 | 9 | 102456329 | G | A | 0.048 | 0.091 | 0.45 | 1.6E-03 | *LINC00587* |
| 6 | rs12550977 | 9 | 102439649 | C | A | 0.048 | 0.091 | 0.45 | 1.6E-03 | *LINC00587* |
| 7 | rs74633492 | 9 | 102455243 | A | G | 0.048 | 0.091 | 0.45 | 1.7E-03 | *LINC00587* |
| 8 | rs10990101 | 9 | 102450544 | A | C | 0.048 | 0.091 | 0.45 | 1.7E-03 | *LINC00587* |
| 9 | rs74496863 | 9 | 102455067 | G | T | 0.048 | 0.091 | 0.45 | 1.7E-03 | *LINC00587* |
| 10 | rs10990104 | 9 | 102452649 | A | C | 0.048 | 0.091 | 0.45 | 1.7E-03 | *LINC00587* |

The top 1,000 SNPs from the discovery GWAS were carried forward to a second cohort of 414 cases and 230 controls. Because of linkage disequilibrium, these SNPs represented ~384 independent tests (*Methods*). Thus, *P*-value threshold was set to 0.05/384 = 0.00013. Three intronic SNPs in the *OPCML* gene in near-perfect linkage disequilibrium met this statistical threshold. Genome coordinates are GRCh38/hg38. Chromosome (Chr.), Reference Allele (Ref), Alternate Allele (Alt), Odds Ratio (OR).

the lowest fraction of Native American ancestry (referent), OR for disease in the subsequent intervals 2 to 5 were 1.92 (1.17 to 3.16; $P = 0.01$), 2.99 (1.79 to 5.01; $P = 2 \times 10^{-5}$), 6.11 (3.01 to 12.39; $P = 2.59 \times 10^{-7}$), and 6.25 (1.88-20.81; $P = 0.0017$) (Fig. 1*B*).

**OPCML Variants and Urinary Concentrating Ability.** We next considered the possible functional role of OPCML in the MeN disease process. OPCML is an opioid receptor binding and cell adhesion protein (18, 19). *OPCML* expression in the kidney, at least under basal conditions, appears low in gene expression databases (*SI Appendix*, Fig. S4). OPCML expression is highest in the brain (*SI Appendix*, Fig. S4) with function best characterized in the hypothalamic cells that release arginine vasopressin (AVP), also known as antidiuretic hormone (20). Because both natural and synthetic opioids can influence salt and water balance in animal models and humans, especially under stress conditions, we hypothesize that *OPCML* SNPs might affect fluid balance (21).

To test the potential effect of *OPCML* SNPs under stress conditions, we measured urinary concentrating ability as a readout of AVP activity in a subset of sugarcane workers during work shifts in high heat. Individuals carrying protective variants had significantly higher urine osmolality (Uosm) (793 ± 62 mOsm) than participants lacking this variant (573 ± 253 mOsm) ($P = 0.003$, Mann–Whitney test) (Fig. 2*A*). We did not detect a statistically significant difference in kidney function between apparently healthy workers with and without the *OPCML* protective allele, although serum creatinine (Scr) trended lower in those with the protective allele ($P = 0.08$, Mann–Whitney test; Fig. 2*B*). There was no apparent relationship between Uosm and Scr in participants either with or without the protective variant (Fig. 2 *C* and *D*), suggesting a direct effect of OPCML on urinary concentrating ability. Since the *OPCML* protective variants are found almost exclusively in individuals of African ancestry, we tested whether there was a difference in fraction African ancestry between participants with and without the variants, but found no difference (0.159 ± 0.047 vs. 0.146 ± 0.038, respectively; $P = 0.6$, Mann–Whitney test).

**Physiological Role of OPCML in Mice.** In order to specifically test the potentially MeN-relevant physiological effects of OPCML, we studied wild-type (*Opcml*$^{+/+}$), heterozygous (*Opcml*$^{+/-}$), and *Opcml* null (*Opcml*$^{-/-}$) mice (*SI Appendix*, Fig. S5). These three groups were phenotypically indistinguishable in total body weight, fat mass, lean mass, or water content (*SI Appendix*, Fig. S6*A*). Furthermore, we observed no differences in food or water intake or physical activity (*SI Appendix*, Fig. S6*B*). Metabolic parameters were the same across all genotypes (*SI Appendix*, Fig. S6*C*)

We challenged these mice with several stressors. Using weight loss as a proxy for volume depletion, *Opcml*$^{-/-}$ mice exposed to 12-h food and water restriction lost significantly less weight than *Opcml*$^{+/+}$ mice, with the weight change in *Opcml*$^{+/-}$ mice midway in between (Fig. 3*A*). When adding furosemide treatment to food and water restriction to accentuate the volume depletion, the same pattern emerged, with weight loss being greatest in *Opcml*$^{+/+}$ mice and least in *Opcml*$^{-/-}$ mice (Fig. 3*B*). We also performed heat stress experiments over a 3 d period where we observed the same pattern of resistance to weight loss by *Opcml* genotype (Fig. 3*C*). As OPCML is also highly expressed in the neural centers that control metabolism and core body temperature, we assessed thermoregulation under ambient and heat stress conditions. Mean core body temperature (CBT) over the 24 h cycle was not different between

**Table 2. Meta-analysis of discovery and replication cohorts**

| | | | | | | | | Discovery | | | | Replication | | | | Combined | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| locus | Overall Rank | Lead SNP | Nearest gene | chr | position | Ref | Alt | Case AF | Control AF | OR | *P*-value | Case AF | Control AF | OR | *P*-value | Case AF | Control AF | OR | *P*-value |
| 1 | 1 | rs73588969 | *OPCML* | 11 | 132591037 | C | T | 0.003 | 0.026 | 0.17 | 0.00011 | 0.006 | 0.033 | 0.14 | 9.48E-05 | 0.005 | 0.029 | 0.16 | 3.97E-08 |
| 2 | 4 | rs62544395 | *CAAP1* | 9 | 26829324 | A | G | 0.015 | 0.048 | 0.33 | 0.00022 | 0.017 | 0.056 | 0.29 | 0.00045 | 0.016 | 0.051 | 0.31 | 3.58E-07 |
| 3 | 5 | rs10990094 | *LINC00587* | 9 | 102441455 | C | G | 0.041 | 0.082 | 0.42 | 8.07E-05 | 0.047 | 0.089 | 0.44 | 0.0017 | 0.044 | 0.085 | 0.43 | 4.80E-07 |
| 4 | 16 | rs75374284 | *NLGN1* | 3 | 174056345 | A | T | 0.058 | 0.019 | 3.26 | 1.44E-05 | 0.042 | 0.017 | 2.41 | 0.011 | 0.050 | 0.019 | 2.90 | 6.19E-07 |
| 5 | 51 | rs7613524 | *FHIT* | 3 | 60932619 | A | G | 0.958 | 0.992 | 0.17 | 2.28E-06 | 0.980 | 0.992 | 0.37 | 0.072 | 0.969 | 0.992 | 0.22 | 8.60E-07 |
| 6 | 80 | rs404977 | *CD44* | 11 | 35110720 | A | G | 0.444 | 0.519 | 0.75 | 0.0060 | 0.473 | 0.592 | 0.59 | 1.92E-05 | 0.458 | 0.546 | 0.68 | 1.15E-06 |
| 7 | 89 | rs181626154 | *RBFOX1* | 16 | 7328233 | G | A | 0.045 | 0.064 | 0.53 | 0.0104 | 0.043 | 0.093 | 0.30 | 1.06E-05 | 0.044 | 0.075 | 0.41 | 1.19E-06 |
| 8 | 121 | rs73098216 | *MACROD2* | 20 | 15522484 | G | A | 0.057 | 0.018 | 3.08 | 2.96E-05 | 0.056 | 0.023 | 2.07 | 0.011 | 0.057 | 0.020 | 2.55 | 1.75E-06 |
| 9 | 132 | rs7277766 | *BACE2* | 21 | 41066772 | G | T | 0.294 | 0.380 | 0.63 | 3.82E-05 | 0.300 | 0.361 | 0.71 | 0.012 | 0.297 | 0.373 | 0.66 | 1.89E-06 |
| 10 | 151 | rs521635 | *TBC1D4* | 13 | 75263362 | C | T | 0.007 | 0.027 | 0.23 | 0.00050 | 0.008 | 0.034 | 0.24 | 0.0013 | 0.008 | 0.029 | 0.23 | 2.14E-06 |

The top SNP at each of the top 10 genomic loci are shown. The top locus was an intronic region of OPCML with 3 SNPs in near-perfect linkage disequilibrium. The combined dataset includes 844 cases and 615 controls. Chromosome (chr), GRCh38/hg38 coordinates (position), Reference Allele (Ref), Alternate Allele (Alt), Odds ratio (OR). The discovery and replication sets were jointly analyzed with METAL (Combined).

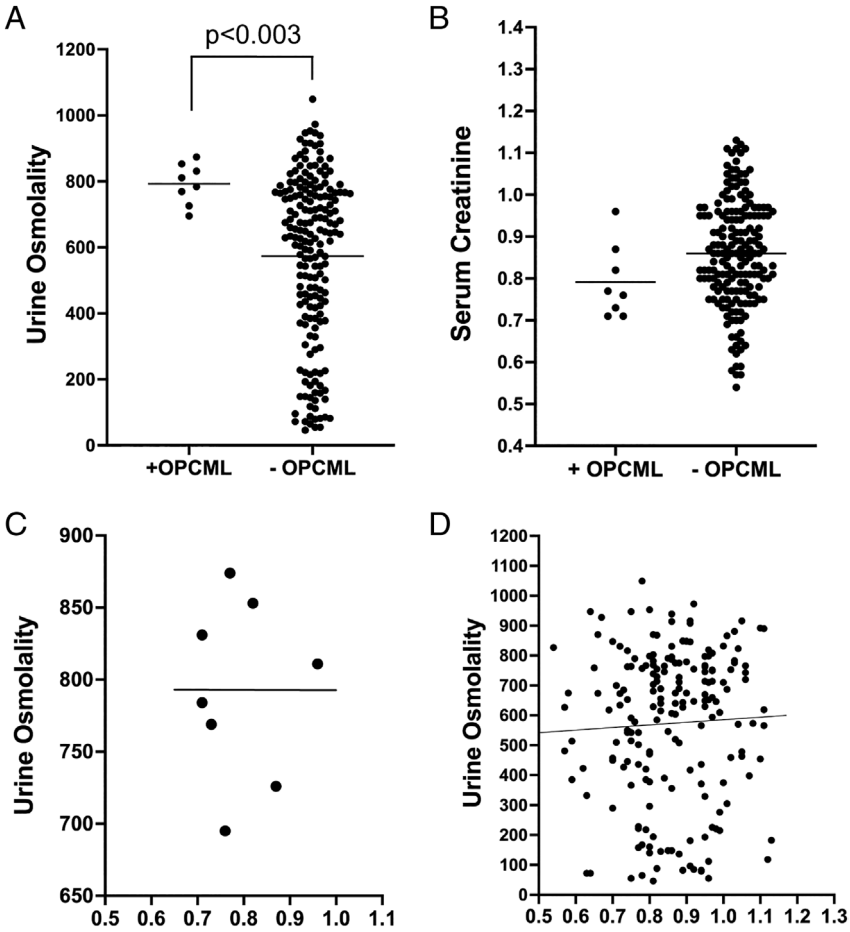


**Fig. 2.** Urine osmolality and kidney function in active sugarcane workers (controls) with and without the *OPCML* protective variant. (*A*) Urine osmolality was higher in apparently healthy, active (control) participants with the *OPCML* protective allele (mean 793 ± 62 mOsm, n = 8) than in noncarriers (573 ± 253 mOsm, n = 188) ($P = 0.003$, Mann–Whitney test). All individuals tested had either zero or 1 protective allele. (*B*) Serum creatinine, the standard biochemical measure of kidney function, did not differ statistically between control participants with (0.79 ± 0.09 mg/dL, n = 8) and without (0.86 ± 0.13 mg/dL, n = 188) the *OPCML* protective allele ($P > 0.05$). (*C* and *D*) There was no significant relationship between serum creatinine and urine osmolality in either workers with (*C*) or without (*D*) the *OPCML* protective allele (trendlines shown). These data suggest that the increased urinary concentration in carriers of the *OPCML* protective allele is not due to better preserved renal function.

groups of mice but peak CBT was lower in $Opcml^{-/-}$ mice than in $Opcml^{+/+}$ mice, with intermediate peak temperatures in $Opcml^{+/-}$ mice (Fig. 3*D*; statistical analysis in *SI Appendix*, Fig. S7*A*). We then heat-stressed the groups of mice and found markedly lower peak temperatures in $Opcml^{-/-}$ mice than in the other two groups (Fig. 3*E*; statistical analysis in *SI Appendix*, Fig. S7*B*); $Opcml^{+/-}$ mice did not differ significantly from $Opcml^{-/-}$ mice, potentially as a result of variability in brain OPCML expression levels in heterozygotes (*SI Appendix*, Fig. S5).

## Discussion

The causes of Mesoamerican nephropathy have not yet been identified, despite evaluation of multiple environmental risk factors. Our findings support an important role for genetic susceptibility in MeN, likely interacting with one or more nongenetic factors. Native American ancestry is strongly associated with MeN risk. Since case status is associated with Native American ancestry, correction for ancestry may have decreased the power to detect causal, ancestry-dependent variants. Nonetheless, we detected one set of candidate protective variants in the *OPCML* gene (also known as OBCAM, or IgLON1), encoding a plasma membrane opiate receptor-binding protein enriched in the brain (20, 22, 23). These intronic *OPCML* variants replicated in an independent cohort and achieved genome-wide significance in meta-analysis. Protective alleles in this variant were associated with higher urine osmolality during heat stress. Physiologic data from *Opcml* knockout mice support a role for OPCML in fluid balance and temperature regulation.

Although not yet definitive, there are plausible mechanisms by which OPCML relates to MeN susceptibility. Both natural and synthetic opioids act as modulators of AVP secretion. Hypothalamic-pituitary AVP secretion by the brain responds to physical stimuli such as pain (and opioids) in addition to its primary roles of maintaining serum tonicity and defending against hypovolemia (24, 25). Cell-based experiments, animal models, and human physiology studies all demonstrate that opioids regulate salt and water balance, although the regulatory mechanisms are complex and context-dependent (21, 26–31). The strong OPCML expression in AVP-releasing magnocellular neurons of the hypothalamus suggests OPCML may help regulate extracellular fluid tonicity or volume at least in part through action on the brain–kidney axis (20, 32). We found that the protective variants correlated with higher urine osmolality during the work shift under high heat stress conditions, potentially indicating better-preserved urinary concentrating ability by the kidney. We detected a trend ($P < 0.08$) toward better glomerular filtration indices in carriers than in noncarriers of *OPCML* variants among

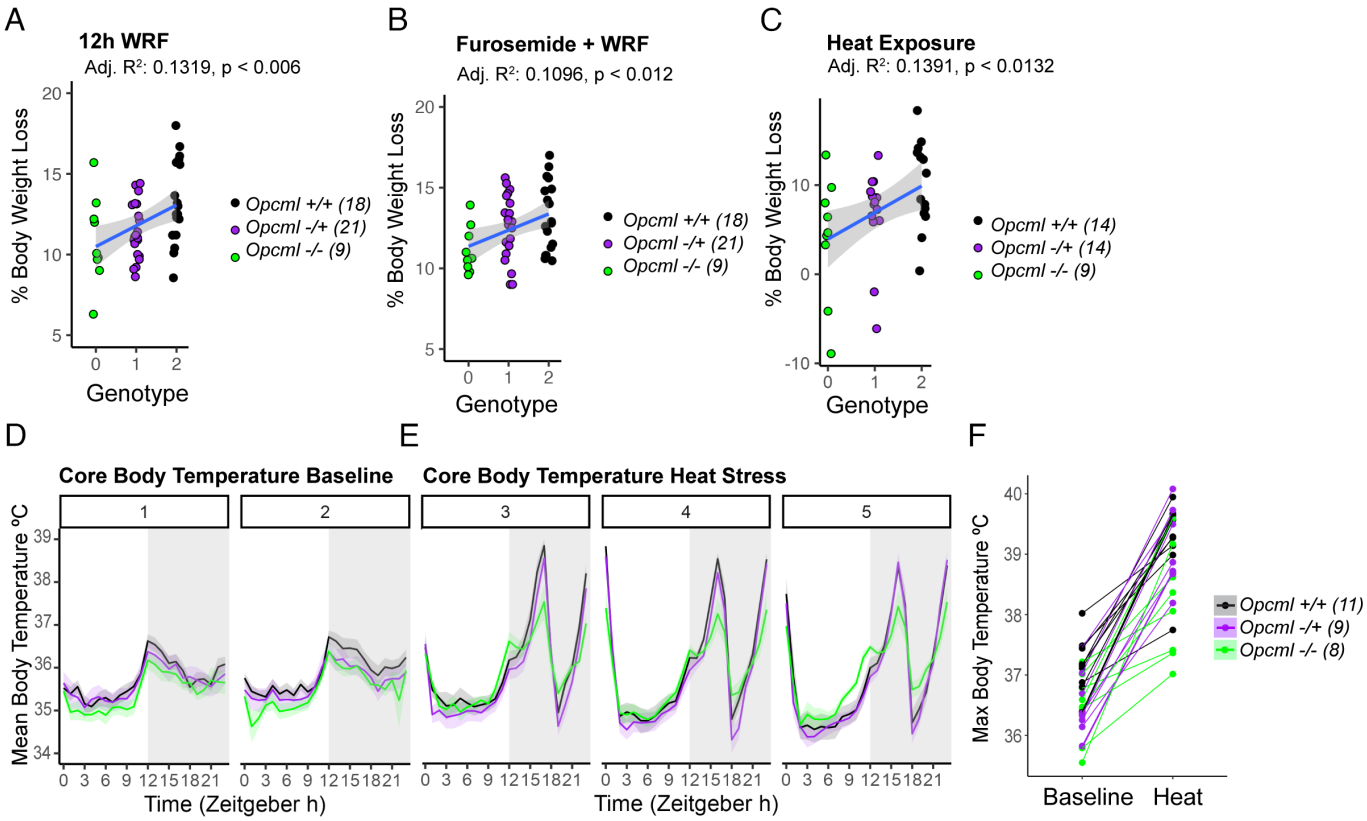


**Fig. 3.** *Opcml* knockout mice exhibit disease-relevant physiological differences in response to volume challenges and heat stress. (*A*) *Opcml* deletion protects against body weight loss after water restriction and fasting (WRF) for 12 h. (*B*) *Opcml* deletion protects against body weight loss after furosemide (50 mg/kg) challenge at the beginning of an 8 h water restriction and fast. (*C*) *Opcml* deletion protects against body weight loss after a 3 d intermittent heat challenge. (*D*) Longitudinal core body temperature measurements collected over 2 d by continuous telemetry under baseline conditions at 23 °C followed by (*E*). 3 d of intermittent heat exposure. In Zeitgeber time (ZT), ZT 0-12 is the light/low activity phase and ZT 12-23 is the dark/high activity phase. (*F*) Baseline and maximum core body temperature during intermittent heat stress. $Opcml^{-/-}$ peak temperature was 1.1 °C/2°F lower than $Opcml^{+/+}$ mice. Summary statistical data for heat stress experiments are located in *SI Appendix*, Fig. S7. Peak temperature is lowest in $Opcml^{-/-}$mice. $Opcml^{+/+}$: Wild-type; $Opcml^{+/-}$: heterozygote; $Opcml^{-/-}$: *Opcml* homozygous null.

participants without apparent disease, so it is possible that the better-preserved concentrating ability in *OPCML* protective variant carriers is in part a reflection of better-preserved renal function within the normal range.

The pathway from variants to altered function of the *OPCML* intronic variants remains unclear. The very low frequency of the protective *OPCML* variants in Europeans makes testing correlations with levels of brain or kidney gene expression difficult at present because non-Europeans are typically underrepresented in eQTL databases. The only other protein-coding gene within 1 million base pairs of the lead *OPCML* SNPs is Neurotrimin (*NTM*), a related member of the IgLON gene family. Although we cannot rule out distant effects of intronic *OPCML* SNPs on NTM function, we used $Opcml^{-/-}$ mice to test OPCML impact on physiological processes that are relevant to heat-related kidney injury. We found that $Opcml^{-/-}$ mice appear resistant to fluid losses and maintain lower body temperatures as compared to $Opcml^{+/+}$ mice. The relative contribution of OPCML to salt and water handling, maintenance of core body temperature, and other processes regulating fluid balance and response to physiologic stress remain to be determined. Expression patterns of OPCML in the human and mouse suggest a central effect in the brain, rather than the kidney, as the anatomical locus of the protective phenotype.

Investigators have postulated that MeN may result from repeated episodes of acute kidney injury caused by volume depletion and kidney hypoperfusion (10, 33). Our data are consistent with this speculative model in which variants that promote enhanced urinary concentrating ability also protect against repetitive episodes of low renal blood flow causing tissue injury and subsequent longer-term decline in renal function. Lower core body temperature would similarly reduce renal hypoperfusion by lessening evaporative fluid loss and preventing temperature-mediated redistribution of blood flow away from the kidney in order to dissipate heat through the skin.

The three *OPCML* protective variants in high linkage disequilibrium are most common in African populations and are rare in European and Asian populations, with intermediate frequencies in admixed populations (*SI Appendix*, Table S7). Although the total contribution of the *OPCML* locus on our study population is in itself probably relatively modest due to its low frequency (~3% in controls), its effect is quite large in individuals who carry the minor allele, with a ~sixfold to sevenfold decrease in the odds ratio for disease. We suspect that multiple genetic variants may act in aggregate to determine adaptation to temperature extremes. While the large effect of ancestry in our study suggests that many variants contribute to MeN susceptibility, we have considered the possibility that genetics may also influence the exposure of individuals to some high-risk occupations and continue to explore this hypothesis.

Reports of the MeN kidney phenotype in Central America appear confined to the Pacific Coast in populations composed of mixed Native American, European, and African ancestry. Although data are too limited to draw a definitive conclusion, the disease

has not been observed widely in Central Americans from the Atlantic coast, where ancestry is predominantly African, or among sugarcane workers of primarily African ancestry from Cuba, Jamaica, Barbados, and other Caribbean countries (12). Ancestors of the Native American population journeyed to the Americas after residing for millennia in Arctic-like conditions. It is thus tempting to speculate that the ancestors of Native Americans faced selection pressures in extremely cold climates that favored genetic variants promoting heat conservation or alternatively lost variants advantageous in the heat of tropical climates (34). Consequent changes in variant frequencies driven by historical exposure to very cold climates may present physiologic challenges to present-day inhabitants of tropical regions performing heavy labor under conditions of high heat and humidity (35). This hypothesis requires additional evidence from population genetics and physiological studies for confirmation. In the meantime, MeN researchers have implemented "hydration, rest, and shade" intervention programs for sugarcane workers and report improved acute kidney injury outcomes compared to historical controls, though the long-term impact on CKD is not yet known (36). These studies suggest that limiting exposure to heat stress may reduce the burden of Mesoamerican nephropathy in at-risk communities in Central America.

In addition to MeN, other areas with extremely high rates of CKD of unknown cause (CKDu) have been identified in Sri Lanka and India. It remains unclear whether CKDu in these diverse locales shares a common etiology. Clinical presentation and high prevalence in agricultural communities are shared features across regions with CKDu. Recent reports of CKDu among returning Nepalese workers after laboring in the Middle East do support heat as a potential etiology since those workers do not share the many other correlated exposures of agricultural work. Important differences are also observed between CKDu regions, such as large differences in male:female ratios, roughly 3:2 in Sri Lanka but likely on the order of 10:1 in the highest prevalence regions with MeN. While manual labor in high heat is a shared feature of CKDu across these regions, in some regions such as Sri Lanka CKDu rates do not track closely with heat, and heat is typically considered a potentially important modifier rather than the primary driver of disease. We did not observe overlap in genetic loci between our study and CKDu genetic studies in Sri Lanka, where the most compelling locus is for a dicarboxylate transporter (*SLC13A3*) with a potential role in excretion of xenobiotics (37). We also did not observe any association between MeN and variants in *APOL1*, *UMOD*, or other known kidney disease genes.

Our study does have some limitations. While our extreme phenotype study design does help minimize misclassification, it does not eliminate it. We expect some of our controls will go on to develop CKD given the high prevalence of MeN in our study population, though we pursue periodic follow-up visits that will identify at least some of these individuals. Because kidney biopsy is not performed in this population, a small number of our cases may have other forms of kidney disease. We expect that both types of misclassification would bias our findings toward the null rather than toward false positives. In addition, we do not yet know how applicable the genetic findings of our study population will be to the wider Central American population. Our urinary concentration studies in field workers are based on a modest number of individuals with the protective OPCML variant, so we will need to expand these observations to more individuals with studies performed under rigorous experimental conditions. Last, though our global OPCML mouse phenotypes related to volume maintenance and temperature regulation appear clear, a more detailed mechanistic evaluation will ultimately be needed to fully understand the specific molecular mechanisms leading to these broad phenotypic differences.

While our findings provide strong evidence for genetic susceptibility to MeN, environmental exposures also clearly play a major role and may ultimately be the focus of preventative measures. Many of the suspected environmental factors are highly correlated and result in large part from poverty. This high correlation between risk factors such as heat stress, agrichemical exposure, environmental toxins, and repetitive pathogen infections make it exceptionally hard to disentangle the causal relationships of each to MeN. Identifying genetic factors that underlie susceptibility may help better understand which specific modifiable risk factors drive the disease.

## Methods

**Study Population.** Study participants were recruited in the Chinandega and León departments of Nicaragua between 2014 and 2020. The study protocol was approved by both US-based (Boston University) and Nicaraguan Ministry of Health institutional review boards. Each participant provided written informed consent. Study investigators conducted the study according to Declaration of Helsinki principles.

Participants in both the discovery and replication cohorts were drawn from the same basic four population groups: 1) Active apparently healthy sugarcane workers were recruited as controls from two sugarcane agribusiness companies separated by ~35 to 40 km. About half the participants were recruited at work and half at their homes or in community centers. Former sugarcane workers with CKD who had previously worked at the same two sugarcane companies were recruited as cases at local recruitment centers set up by our study team. Recruitment efforts were aided by a large association of former sugarcane workers with CKD. 2) Individuals from a mining community originally recruited for an infectious disease study (38). 3) Brickmakers originally recruited as part of an occupational study (8). 4) Residents in a diverse occupational community originally recruited to participate in a longitudinal study of CKD (39). Starting in 2018, participants were paid a small stipend ($6) to participate, based on feedback from community groups and approved by local and US-based IRBs.

The study used an extreme phenotype design in order to minimize misclassification given the very high rates of disease in this MeN hotspot: We sought controls who were highly resistant to kidney injury and cases who appeared highly susceptible to kidney injury. For controls, we prioritized low serum creatinine, older age, and long duration of manual labor involving heat exposure; for cases, we prioritized high serum creatinine, younger age, and short duration of manual labor involving heat exposure. We prioritized cases with age of CKD diagnosis <35 y; we did include some older nonsugarcane cases with longstanding CKD where the age of diagnosis was uncertain. Most sugarcane cases also had renal ultrasounds performed by a renal clinic to exclude cystic kidney disease, congenital urinary tract abnormalities, or obstruction due to kidney stones. We used serum creatinine values instead of estimated GFR because there were no validated estimating equations in admixed Central American populations during the recruitment period.

**Biosampling and Laboratory Tests.** Biosamples were kept at 4 °C in the field and processed the same day. Samples were stored locally for up to 48 h at −20 °C prior to transport to −80 °C freezers in Managua. Serum creatinine was measured either at the Centro Nacional de Diagnóstico y Referencia (CNDR), the national laboratory in Nicaragua located at the Ministry of Health in Managua, or by iSTAT autoanalyzer (Abbott) in the field. Correlation between laboratory and iSTAT creatinine measurements was high ($r^2 = 0.99$) although iSTAT tended to read ~0.1 mg/dL higher (15).

**GWAS Analysis.**

***Genotyping and quality control of MeN samples.*** DNA was isolated from saliva (DNA Genotek; Ontario, Canada) or blood (PAXgene; Preanalytix, Switzerland). Genotyping was performed on Illumina MEGA (discovery set) or GDA (replication set) arrays by Illumina (San Diego, CA). For discovery, 815 participants were

genotyped using the Multi-Ethnic Global Array Beadchip (1,534,476 variants); replication included 644 individuals using the Global Diversity Array Beadchip (1,896,998 variants) that replaced the Multi-Ethnic Global Array Beadchip on the Illumina platform. Quality control (QC) of individual samples checked for genotyping consistency, concordance between annotated and genetically determined sex, unexpected duplicates, contamination, and missing genotypes > 5%; samples failing QC were repeated in a subsequent batch. Genotypes were called from Genome Reference Consortium Human Build 37 (GRCh37), and coordinates were changed to Build 38 using Picard Tools. High-quality, informative genotypes were selected from autosomal biallelic SNPs, excluding: ambiguous (A/T and G/C) calls, duplicate variants, loci of call-rate < 95%, minor allele frequency (MAF) < 1%, heterozygosity > 0.45, or Hardy–Weinberg equilibrium $P$-value < $10^{-6}$ among controls. Additionally, ~100 SNPs with large differences ($P < 5 \times 10^{-5}$) in allele frequency across replication and discovery groups were removed.

***Population structure and genetic relatedness adjustments.*** We calculated PCs to capture PS from continental ancestry and genetically determined kinship coefficients for known and unknown relatedness (40). PC plots were examined using 1,000 Genomes phase 3 (1kGp3) as anchor populations separately for discovery and replication sets to further QC genotyping and guide selection of iteratively created PCs to reflect PS after accounting for cryptic relatedness. We used two rounds of PCAiR (41) to calculate all PCs. In each iteration, we used approximately 100,000 LD-pruned SNPs ($r^2 < 0.3$ within a 50 kb sliding window and step size of 5 kb), common to both MeN and 1kG, with call-rate > 98% and MAF > 5%, and among individuals estimated to less than 3rd-degree relatives (pi-hat > 0.1875). No outliers were found in PC plots. PCs weights from MeN samples were projected on to coordinates from 1kGp3 samples for interpretation of PC values. Additionally, PCs and genetic relatedness between participants were calculated from our study population to capture structure in our data. Based on PCA and the strength of the association between PCs and case-status in bivariate analysis, we parsimoniously chose to include PC1 as a fixed effect in our statistical models. We also included random effects with correlation structure specified by pairwise kinship coefficients. Additionally, we assessed the robustness of our results by repeating the analysis while adjusting for PCs 1–10.

***Discovery, replication, and meta-analysis.*** Discovery and replication data were nonoverlapping samples from the same regional population.

The top 1,000 SNPs from the discovery cohort were carried forward for testing in the replication cohort, adjusting the significance threshold for multiple testing while accounting for linkage disequilibrium between variants (42), leading to a threshold of $P < 1.3 \times 10^{-4}$ for replication (0.05/384 independent tests). We used inverse-variance weighted meta-analysis, implemented in METAL (17), to jointly analyze variants shared between discovery and replication (threshold $P < 5 \times 10^{-8}$). Summary statistics for variants with MAF > 0.01 in the combined samples of cases and controls were examined, and beta coefficients checked to ensure consistent direction of effect between datasets.

***Imputation.*** Genetic variants not covered by the arrays, removed for QC, and not genotyped in both discovery and replication were imputed separately for discovery and replication sets using a cosmopolitan (43) reference panel of available populations. The TOPMed Imputation server (https://imputation.biodatacatalyst.nhlbi.nih.gov) was used to prephase [Eagle v2.4 (44)] and impute (Minimac4) (45) MeN genotypes to TOPmed freeze 5 as a reference panel (46). The TOPMed server removed any variants not present in the 125,568 reference haplotypes and only returned results with a minimum imputation $R^2 > 0.3$.

***Genome-wide ancestry inference.*** We utilized RFMix (47) v.2.0 to classify loci by ancestral origin using the Human Genome Diversity Project (HGDP) (48, 49) as the reference panel. Based on the population history of Nicaragua, we selected HGDP to best represent the Amerindian (AMR), European (EUR), and African (AFR) founder populations' genotypes, and only used samples from these populations. High-quality SNPs from MeN were further filtered to remove genotypes within the lactase gene, the human leukocyte antigen (HLA) genes, or polymorphic inversion regions on chromosomes 8 and 17. RFmix requires phased data without missing genotypes: We implemented Beagle (v.4) to phase and impute sporadically missing (50) genotypes in a combined dataset of MeN and HGDP. We balanced the sample sizes for AMR, EUR, and AFR to prevent bias in our ancestry inferences (minimum n = 61). RFMix classified 625,984 SNPs in discovery and 540,239 in replication based on 12 generations since admixture, a minimum of 5 reference haplotypes/tree, 10,000 trees, 0.2 cM windows, and 3 iterations of expectation maximization. To estimate individuals' proportion of each ancestry, we calculated means weighted to the number of SNPs in each haploblock.

***Ancestry analysis.*** We used RFMix v2.0 to classify loci by ancestral origin using the HGDP as the reference panel. To assess the effect of a 1% increase in Native American ancestry on odds of MeN, we fit a logistic mixed model for the fixed effect of ancestry accounting for relatedness as a random effect. We performed a second analysis to allow for nonlinear effects of the percent Native American ancestry on the risk of MeN by binning all participants into five evenly spaced intervals based on percent Native American ancestry from lowest to highest in our population and performing contingency testing using the low-Native American ancestry interval as the referent.

***Statistical methods.*** In discovery and replication, we utilized a logistic mixed model to test for genetic associations between estimated SNP dosages and odds of MeN. The effect of SNPs (Gβ) in our model, unless stated otherwise, was adjusted for PC1 (Xα), as a fixed effect from population stratification, and for polygenic effects as a random effect (b):

$$logit(y) = X\alpha + G\beta + b. \quad [1]$$

Polygenic effects of SNPs arise from recent, often cryptic, genetic relatedness in a population. We assume these individual-level random effects (b) are normally distributed with correlation structure specified by pairwise kinship coefficients. We used GENESIS (51) to perform association tests for each SNP (G) via a score test, comparing a model without genetic variants to a model that includes each SNP (G) separately. We relied on a genome-wide significance of $5 \times 10^{-8}$ to control for familywise error, due to multiple testing, for stage 1 (discovery) and meta-analysis.

For replication, we adjusted our significance level for the effective number of independent genetic variants among the top 1,000 discovery (stage 1) variants. We anticipated strong linkage disequilibrium among loci from our study population due to relatively recent admixture, which we observed in our local ancestry estimates. Therefore, we considered 384 independent loci and a $P$-value of $1.3 \times 10^{-4}$ as statistically significant for stage 2 based on the method of Gao et al. (42).

To assess the effect of a 1% increase in Native American ancestry on odds of MeN, we fit a null logistic mixed model of Eq. **1** in GENESIS for the fixed effect of ancestry adjusted for relatedness as a random effect. Proportions of Native American ancestry were transformed to a 100% scale.

***Mitochondrial and Y haplotype analyses.*** To assign mitochondrial haplogroups, we used HaploGrep2 software (52) with the Phylotree database version 17 (53), consolidating haplogroups to their oldest common root. Male participants were assigned a Y-haplogroup using yHaplo software application (BioRxiv DOI: 10.1101/088716) with the International Society of Genetic Geneology chromosome Y phylogenetic tree 2016 database (http://www.isogg.org/tree/). Y-Haplogroups consolidated the predicted haplogroups to their most recent common root node. The world distribution of haplogroups was determined with a heuristic approach using PhyloGeographer Heatmap tool (https://phylogeographer.com/). Differences in the frequency of haplogroups between cases and controls were tested using a chi-square test by Monte Carlo simulation.

***Evaluation of urinary concentration.*** A subset of the Discovery controls were active sugarcane workers who were studied during work shifts. We collected urine from these participants in the setting of exertion in high heat and humidity. For osmolality testing, we included control participants that we had recruited during or at the end of the workshift, excluding participants with serum sodium <130 mEq/L or serum potassium <3 mEq/L. Urine osmolality was measured using the freezing point depression method with an Osmette A osmometer (Precision Systems). Differences in the urine osmolality, serum creatinine, and proportion of African ancestry of participants with the protective variant rs73588969 and without the variant were compared using the Mann–Whitney test.

***Mouse methods.*** Mouse model: C57BL/6NJ-*Opcml*$^{em1(IMPC)J}$/Mmjax (MMRRC Stock#66831) mice were generated by the Knockout Mouse Phenotyping Program at The Jackson Laboratory using CRISPR/Cas9 mediated knock-out of the opioid binding protein/cell adhesion molecule-like gene (*Opcml*). Briefly, this allele was generated by electroporating Cas9 protein and 2 guide sequences TGGTTGGTTGACTAAGGGGA and GAGCTGAGAACTCAACAAAA resulting in a 441 bp deletion beginning at Chromosome 9 position 28,813,224 bp and ending after 28,813,664 bp (GRCm39/mm39). This mutation deletes ENSMUSE00000537413 (exon 6) and 320 bp of flanking intronic sequence including the splice acceptor and donor and causes a frameshift

at residue 213 and early truncation 1 amino acid later. OPCML expression levels were verified by real-time qPCR and protein using the whole mouse brain (*SI Appendix*, Fig. S4). Heterozygote siblings were bred to obtain male experimental mice with wild-type (*Opcml*$^{+/+}$), heterozygous (*Opcml*$^{+/-}$), and OPCML null (*Opcml*$^{-/-}$) genotypes. 8 to 12 h fasting/water restriction ± diuretic challenge was performed in 8- to 12-wk-old mice with body weight used as a surrogate for volume depletion. Core body temperature experiments were performed in 10- to 24-wk-old mice in Promethion metabolic cages (Sable Systems International) using implantable peritoneal body temperature probes. Body composition data were obtained using the EchoMRI 3-in-1 body composition analyzer (BIDMC Energy Balance Core). All mice experiments were approved by the IACUC at Beth Israel Deaconess Medical Center.

***Mouse physiology assays.***

***12 h water restriction and fasting.*** 8- to 12-wk-old male mice were weighed after urine voiding just before lights off (Zeitgeber Time 11–12). Mice were placed back in their home cages without access to food and water for 12 h during the activity phase. Mice were weighed at lights on (ZT0) after urine voiding at the end of the food and water restriction and then food and water were returned to the cage.

***Furosemide 50 mg/kg bodyweight with 8 h water restriction and fasting.*** Food and water were removed, urine was voided, and 8- to 12-wk-old male mice were weighed prior to receiving 50 mg/kg bodyweight of furosemide in isotonic saline (10 mg/mL, Sigma Millipore F4381) by oral gavage at ZT4. Bodyweights were measured just prior to lights off (ZT12) at 8 h post furosemide dosing. Water and food were returned to the cage at the end of the exposure.

***Core body temperature.*** 16- to 24-wk-old male mice were used for metabolic phenotyping and core body temperature data obtained through the Energy Balance Core at Beth Israel Deaconess Medical Center using Promethion Metabolic Cages. Mice were implanted with a body temperature probe placed in the peritoneal cavity. Mice recovered for 1-wk postimplantation surgery prior to beginning the experiment. Mice were housed at 23 C and acclimatized to the cages. Continuous data were collected for 48 h post acclimatization to obtain baseline metrics. Data were visualized in CalR (Cell Metabolism, 2018). All data were plotted in ggplot2 using R.

***Heat stress.*** 10- to 18-wk-old male mice were acclimated in the metabolic cages for 24 h prior to the experimental start. Baseline temperature data were collected continuously for 2 full 24 h light cycles at 23 °C. In the heat stress phase, the cage temperature ramped over 4 h from 23 °C to 38 °C, was maintained at 38 °C for 2 h, and then cooled to 23 C for 1 h. This cycle was performed twice during the activity phase (lights off) for three consecutive days. Mice had ad libitum access to food and water throughout the baseline collection and the heat exposure. At the end of the heat exposure, mice were removed from the chamber and weighed.

**Data, Materials, and Software Availability.** The study includes participants recruited between 2014 and 2020. Individual data sharing was not approved until 2018. Aggregate GWAS data for the full study and individual-level genetic data for participants recruited between 2018 and 2020 will be available through dbGAP to qualified users (accession number: phs003813.v1.p1) (54). The use of this data is restricted to the study of kidney disease. All participants of this study were recruited from a single small geographic region; a large percentage are current of former employees of two large companies. To protect the privacy of these individuals, application to dbGAP will be required for access to data.

**ACKNOWLEDGMENTS.** We are very grateful to the individuals who participated in this study. We also appreciate the support of the National Laboratory at the Nicaraguan Ministry of Health (MINSA) for help with storage, testing, and shipment of samples. This work was funded by the Doris Duke Charitable Foundation (D.J.F.), NIH/NIDDK DK116021 (D.J.F. and D.R.B.), a Satellite Healthcare Foundation Coplon Award (D.J.F.), and Beth Israel Deaconess Medical Center (D.J.F.).

Author affiliations: [a]Division of Nephrology, Beth Israel Deaconess Medical Center, Boston, MA 02215; [b]Harvard Medical School, Boston, MA 02215; [c]Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA 02142; [d]Department of Epidemiology, Boston University School of Public Health, Boston, MA 02118; [e]Centro Medico del Pacifico, Masaya, Nicaragua 41000; [f]Especialistas en Medicina Interna, Chichigalpa, Nicaragua 26100; [g]Department of Population Medicine, Harvard Medical School, Boston, MA 02215; [h]Harvard Pilgrim Healthcare Institute, Boston, MA 02215; [i]Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118; [j]Barcelona Institute for Global Health, Barcelona 08036, Spain; [k]Endocrinology Division, Beth Israel Deaconess Medical Center, Boston, MA 02215; [l]Department of Environmental and Occupational Health, Milken Institute School of Public Health, George Washington University, Washington, DC 20052; [m]Department of Environmental Health, Boston University School of Public Health, Boston, MA 02118; and [n]Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, QC H3A 1G1, Canada

1. J. Cohen, Mesoamerica's mystery killer. *Science* **344**, 143–147 (2014).
2. P. Ordunez *et al.*, The epidemic of chronic kidney disease in Central America. *Lancet Glob. Health* **2**, e440–e441 (2014).
3. R. J. Johnson, C. Wesseling, L. S. Newman, Chronic kidney disease of unknown cause in agricultural communities. *N. Engl. J. Med.* **380**, 1843–1852 (2019).
4. R. Chatterjee, Occupational hazard. *Science* **352**, 24–27 (2016).
5. M. Gonzalez-Quiroz *et al.*, Population-level detection of early loss of kidney function: 7-year follow-up of a young adult cohort at risk of Mesoamerican nephropathy. *Int. J. Epidemiol.* **53**, dyad151 (2023), 10.1093/ije/dyad151.
6. R. Correa-Rotter, R. Garcia-Trabanino, Mesoamerican nephropathy. *Semin. Nephrol.* **39**, 263–271 (2019).
7. J. Wijkstrom *et al.*, Renal morphology, clinical findings, and progression rate in Mesoamerican nephropathy. *Am. J. Kidney Dis.* **69**, 626–636 (2017).
8. L. Gallo-Ruiz *et al.*, Prevalence and Risk Factors for CKD Among Brickmaking Workers in La Paz Centro, Nicaragua. *Am. J. Kidney Dis.* **74**, 239–247 (2019).
9. S. A. Keogh *et al.*, High prevalence of chronic kidney disease of unknown etiology among workers in the Mesoamerican Nephropathy Occupational Study. *BMC Nephrol.* **23**, 238 (2022).
10. J. Kupferman *et al.*, Acute kidney injury in sugarcane workers at risk for Mesoamerican nephropathy. *Am. J. Kidney Dis.* **72**, 475–482 (2018).
11. M. Gonzalez-Quiroz *et al.*, Decline in kidney function among apparently healthy young adults at risk of Mesoamerican nephropathy. *J. Am. Soc. Nephrol.* **29**, 2200–2212 (2018).
12. D. J. Friedman, Genes and environment in chronic kidney disease hotspots. *Curr. Opin. Nephrol. Hypertens* **28**, 87–96 (2019).
13. D. Friedman, V. A. Luyckx, Genetic and developmental factors in chronic kidney disease hotspots. *Semin. Nephrol.* **39**, 244–255 (2019).
14. M. Gonzalez-Quiroz, N. Pearce, B. Caplin, D. Nitsch, What do epidemiological studies tell us about chronic kidney disease of undetermined cause in Meso-America? A systematic review and meta-analysis. *Clin. Kidney J.* **11**, 496–506 (2018).
15. J. Kupferman *et al.*, Characterization of Mesoamerican nephropathy in a kidney failure hotspot in Nicaragua. *Am. J. Kidney Dis.* **68**, 716–725 (2016).
16. S. Das *et al.*, Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
17. C. J. Willer, Y. Li, G. R. Abecasis, METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
18. P. Govitrapong, X. Zhang, H. H. Loh, N. M. Lee, Transfection of NG108-15 cells with antisense opioid-binding cell adhesion molecule cDNA alters opioid receptor-G-protein interaction. *J. Biol. Chem.* **268**, 18280–18285 (1993).
19. C. M. Lane, R. Elde, H. H. Loh, N. M. Lee, Regulation of an opioid-binding protein in NG108-15 cells parallels regulation of delta-opioid receptors. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11234–11238 (1992).
20. S. Miyata *et al.*, Expression of the IgLON cell adhesion molecules Kilon and OBCAM in hypothalamic magnocellular neurons. *J. Comp. Neurol.* **424**, 74–85 (2000).
21. K. Gupta, M. L. Weber, Renal effects of opioid exposure: Considerations for therapeutic use. *J. Opioid. Manag.* **2**, 236–240 (2006).
22. M. Uhlen *et al.*, Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
23. P. J. Thul *et al.*, A subcellular map of the human proteome. *Science* **356**, eaal3321 (2017).
24. E. J. Hoorn, R. Zietse, Diagnosis and treatment of hyponatremia: Compilation of the guidelines. *J. Am. Soc. Nephrol.* **28**, 1340–1349 (2017).
25. G. Liamis, H. Milionis, M. Elisaf, A review of drug-induced hyponatremia. *Am. J. Kidney Dis.* **52**, 144–153 (2008).
26. S. Mercadante, E. Arcuri, Opioids and renal function. *J. Pain* **5**, 2–19 (2004).
27. D. R. Kapusta, S. Y. Jones, G. F. DiBona, Renal mu opioid receptor mechanisms in regulation of renal function in rats. *J. Pharmacol. Exp. Ther.* **258**, 111–117 (1991).
28. G. F. DiBona, S. Y. Jones, Role of endogenous peripheral opioid mechanisms in renal function. *J. Am. Soc. Nephrol.* **4**, 1792–1797 (1994).
29. K. A. Van Tilborg, T. J. Rabelink, H. A. Koomans, Naloxone inhibits renal hemodynamic effect of head-out water immersion in humans. *Kidney Int.* **48**, 860–865 (1995).
30. D. R. Kapusta, E. M. Dzialowski, Central mu opioids mediate differential control of urine flow rate and urinary sodium excretion in conscious rats. *Life Sci.* **56**, PL243-8 (1995).
31. L. A. Walker, J. C. Murphy, Antinatriuretic effect of acute morphine administration in conscious rats. *J. Pharmacol. Exp. Ther.* **229**, 404–408 (1984).
32. S. Miyata, K. Taguchi, S. Maekawa, Dendrite-associated opioid-binding cell adhesion molecule localizes at neurosecretory granules in the hypothalamic magnocellular neurons. *Neuroscience* **122**, 169–181 (2003).
33. R. S. B. Fischer *et al.*, Clinical evidence of acute Mesoamerican nephropathy. *Am. J. Trop. Med. Hyg.* **97**, 1247–1256 (2017).
34. E. Ruiz-Pesini, D. Mishmar, M. Brandon, V. Procaccio, D. C. Wallace, Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* **303**, 223–226 (2004).
35. D. Mishmar *et al.*, Natural selection shaped regional mtDNA variation in humans. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 171–176 (2003).
36. J. Glaser *et al.*, Preventing kidney injury among sugarcane workers: Promising evidence from enhanced workplace interventions. *Occup. Environ. Med.* **77**, 527–534 (2020).
37. S. Nanayakkara *et al.*, An integrative study of the genetic, social and environmental determinants of chronic kidney disease characterized by tubulointerstitial damages in the North Central Region of Sri Lanka. *J. Occup. Health* **56**, 28–38 (2014).
38. W. K. Yih *et al.*, Investigating possible infectious causes of chronic kidney disease of unknown etiology in a Nicaraguan mining community. *Am. J. Trop. Med. Hyg.* **101**, 676–683 (2019).

39. J. K. O'Donnell *et al.*, Prevalence of and risk factors for chronic kidney disease in rural Nicaragua. *Nephrol. Dial. Transplant.* **26**, 2798–2805 (2011).
40. A. L. Price *et al.*, Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
41. M. P. Conomos, M. B. Miller, T. A. Thornton, Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. *Genet. Epidemiol.* **39**, 276–293 (2015).
42. X. Gao, J. Starmer, E. R. Martin, A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet. Epidemiol.* **32**, 361–369 (2008).
43. B. Howie, J. Marchini, M. Stephens, Genotype imputation with thousands of genomes. *G3 (Bethesda)* **1**, 457–470 (2011).
44. P. R. Loh, P. F. Palamara, A. L. Price, Fast and accurate long-range phasing in a UK Biobank cohort. *Nat. Genet.* **48**, 811–816 (2016).
45. C. Fuchsberger, G. R. Abecasis, D. A. Hinds, minimac2: Faster genotype imputation. *Bioinformatics* **31**, 782–784 (2015).
46. D. Taliun *et al.*, Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. *Nature* **590**, 290–299 (2021).
47. B. K. Maples, S. Gravel, E. E. Kenny, C. D. Bustamante, RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. *Am. J. Hum. Genet.* **93**, 278–288 (2013).
48. L. L. Cavalli-Sforza, The human genome diversity project: Past, present and future. *Nat. Rev. Genet.* **6**, 333–340 (2005).
49. H. M. Cann *et al.*, A human genome diversity cell line panel. *Science* **296**, 261–262 (2002).
50. S. R. Browning, B. L. Browning, Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).
51. S. M. Gogarten *et al.*, Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics* **35**, 5346–5348 (2019).
52. H. Weissensteiner *et al.*, HaploGrep 2: Mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).
53. M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30**, E386–E394 (2009).
54. D. J. Friedman *et al.*, Genetic susceptibility to kidney disease in Nicaragua. dbGAP. Accession number phs003813v1.p1 (application required). Deposited 5 November 2024.