Epigenome-wide association study of kidney function identifies trans-ethnic and ethnic-specific loci

Charles E. Breeze1,2,3*, Anna Batorsky4, Mi Kyeong Lee5, Mindy D. Szeto6, Xiaoguang Xu7, Daniel L. McCartney8, Rong Jiang9, Amit Patki10, Holly J. Kramer11,12, James M. Eales7, Laura Raffield13, Leslie Lange6, Ethan Lange6, Peter Durda14, Yongmei Liu15, Russ P. Tracy14,16, David Van Den Berg17, NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, TOPMed MESA Multi-Omics Working Group, Kathryn L. Evans5, William E. Kraus15,18, Svati Shah15,18, Hermant K. Tiwan10, Lifang Hou19,20, Eric A. Whitser21,22, Xiao Jiang7, Fadi J. Charchar23,24,25, Andrea A. Baccarelli26, Stephen S. Rich27, Andrew P. Morris28, Marguerite R. Irvin29, Donna K. Arnett30, Elizabeth R. Hauser15,31, Jerome I. Rotter32, Adolfo Correa33, Caroline Hayward34, Steve Horvath35,36, Riccardo E. Marioni8, Maciej Tomaszewski37, Stephan Beck2, Sonja I. Berndt1, Stephanie J. London5, Josyf C. Mychaleckyj27 and Nora Franceschini21*

Abstract

Background: DNA methylation (DNAm) is associated with gene regulation and estimated glomerular filtration rate (eGFR), a measure of kidney function. Decreased eGFR is more common among US Hispanics and African Americans. The causes for this are poorly understood. We aimed to identify trans-ethnic and ethnic-specific differentially methylated positions (DMPs) associated with eGFR using an agnostic, genome-wide approach.

Methods: The study included up to 5428 participants from multi-ethnic studies for discovery and 8109 participants for replication. We tested the associations between whole blood DNAm and eGFR using beta values from Illumina 450K or EPIC arrays. Ethnicity-stratified analyses were performed using linear mixed models adjusting for age, sex, smoking, and study-specific and technical variables. Summary results were meta-analyzed within and across ethnicities. Findings were assessed using integrative epigenomics methods and pathway analyses.

(Continued on next page)
Background
The kidney has a central role in body homeostasis through the regulation of blood pressure, fluid, and electrolytes and by removing endogenous and exogenous toxins. Reduced kidney function measured using estimated glomerular filtration rate (eGFR) defines chronic kidney disease (CKD). CKD affects 14.5% of the adult US population and is a leading cause of death and disability [1, 2]. CKD has a high burden among non-European US ethnic groups, but mechanisms for this health disparity are poorly understood [3]. A better understanding of the mechanisms influencing kidney function may provide insights into CKD occurrence and risk.

Complex interactions between genetic, lifestyle, and environmental exposures likely contribute to the observed eGFR variation across populations. DNA sequence variation accounts for 7.6% of the estimated heritability of eGFR in trans-ethnic genome-wide association studies (GWAS) [4]. Epigenetic modifications of the genome such as DNA methylation (DNAm) are heritable and contribute to gene regulation. DNAm consists of the addition of a methyl group to cytosines, typically at cytosine-guanine dinucleotides (CpG sites). DNAm is influenced by lifetime exposures and may provide clues on ethnic-specific differences influencing eGFR. Differential DNAm at CpG sites can be studied using microarrays with reasonable genome-wide coverage through epigenome-wide association studies (EWAS) [5, 6].

Recent EWAS in whole blood have identified differentially methylated positions (DMPs) associated with blood pressure and eGFR, and disease states such as CKD and rapid decline in eGFR [7–11]. Early studies had modest sample sizes (40 to 407 individuals) or were limited to a single ethnic group [7–9]. A large EWAS performed separate discovery analyses within the Atherosclerosis Risk in Communities (ARIC, 2264 African Americans) and the Framingham Heart Study (FHS, 2395 white participants) followed by cross-replication of findings [11]. The study identified 19 DMPs for eGFR or CKD. Overall, these studies support a role for DNAm in CKD-related traits. However, previous studies did not account for important potential confounders such as smoking status and cumulative exposure in the discovery group, which have widespread effects on DNAm patterns [12] and are risk factors for CKD, nor did they assess DNAm at CpG sites across multiple ethnic groups during discovery.

The main aim of this study is to identify DNAm patterns associated with eGFR in multi-ethnic studies using data from European/European American (EA), African American (AA), and Hispanic/Latino (H/L) participants. We performed both trans-ethnic and ethnic-specific EWAS using whole blood-based Illumina DNAm data assayed in participants of the Women’s Health Initiative (WHI), the Multi-Ethnic Atherosclerosis Study (MESA), and the Jackson Heart Study (JHS). We replicated our findings in the HyperGEN, Generation Scotland, and CATHGEN studies, in addition to analyzing results via look-ups in a published study [11]. We identified DMPs associated with eGFR in trans-ethnic and ethnic-specific analyses, and provided supporting evidence for the contribution of identified DMPs to kidney function and development using in silico approaches and human kidney tissue-specific data.

Methods
Study design and populations
Our study design included a discovery step comprising three population-based studies (WHI, MESA, and JHS) and three replication studies (HyperGEN, Generation Scotland, and CATHGEN) (Additional file 1: Fig. S1). WHI is a study of postmenopausal women (aged 50–79 years), comprising 161,808 women recruited from 40 US clinical centers who participated in an observational study or in clinical trials during 1993–1998 as previously described [13–16]. MESA is a multi-ethnic study of subclinical cardiovascular disease and risk factors for cardiovascular disease [17], consisting of 6814 asymptomatic men and women aged 45–84 (38% EA, 28% AA, 22% H/L, and 12% Asian) recruited from six field centers across the USA and examined in 2000–2002, followed by four
subsequent examination periods. JHS is a study of cardiovascular disease and its risk factors in AA, comprising 5306 African Americans aged 21 to 94 years recruited from the Jackson, MS, metropolitan area from 2000 to 2004, with four follow-up exams [18]. HyperGEN is a family-based study with a sib-pair design. Hypertensive African American sibships were recruited from Forsyth County, NC, and from the community-at-large in Birmingham, AL, from 1995 to 2000 [19]. Generation Scotland is a family-based and population-based study consisting of 23,690 European participants recruited via general medical practices across Scotland between the years 2006 and 2011 [20]. CATHGEN is a biorepository of clinical samples from a prospectively collected clinical cohort of individuals undergoing cardiac catheterization at Duke University [21]. Both discovery and replication included multi-ethnic studies. Study descriptions are shown in Additional file 2: Supplementary Methods.

Phenotypes

Serum creatinine-based eGFR was estimated using the Chronic Kidney Disease Epidemiology equation which includes age, sex, and a constant for AA [22].

Epigenetic data and quality control

Briefly, preprocessing included removal of probes with detection $p$ values > 0.01 in > 10% of samples and samples with detection $p$ values > 0.01 in > 1% of probes. Beta values were normalized using beta-mixture quantile (BMIQ) normalization method (WHI-BAA and WHI-EMPC) [23], the normal-exponential out-of-band (NOOB) preprocessing method (JHS, MESA, CATHGEN) [24], Subset Quantile Normalization (SQN) (CATHGEN) [25], or dasen (Generation Scotland) [26]. Batch effect correction was performed using ComBat [27] or adjusting batch as a covariate. CpGs overlapping with the list of potentially polymorphic sites in the relevant ethnic group and cross-reactive probes were removed [28]. Cell proportions were estimated using the reference-based Houseman method for whole blood [29]. To adjust for population structure, principal components (PCs) or ethnic-informative markers were obtained from the genome-wide genotype data available using standard methods [30]. DNAm sites were annotated to include chromosome, position, UCSC gene names, relationship to CpG islands, location in gene enhancer regions, and DNase I hypersensitive sites (DHSs) using Illumina’s annotation file [31]. Detailed methods for MESA, which have not been previously published, are included in Additional file 2: Supplementary Methods.

EWAS

Methylation betas were used as predictors for eGFR in ethnic-stratified analyses, adjusting for age, sex, smoking history (current, past, and no smoking, and pack-years), 4 to 10 principal components, cell type composition, and study-specific covariates. DMPs were modeled by fitting robust linear models (or linear mixed models in JHS to account for family relationships) and performing robust standard error calculations via the ‘sandwich’ package. These analyses were performed using R version 3.5.3. Because of the small samples within each ethnicity in MESA, we fitted linear models in that study without robust estimation. EWAS results were meta-analyzed across all samples and within each ethnicity using fixed-effect inverse-variance weighted methods implemented in METAL. We required a minimum of two studies for meta-analyses. We used an FDR < 0.05 (Benjamini-Hochberg).

EWAS analyses were subsequently performed in HyperGEN (AA), Generation Scotland (EA), and CATHGEN (AA and EA) using the same statistical protocols as in discovery analyses. For trans-ethnic replication, we combined all the replication samples and used a Bonferroni-adjusted $p$-value cutoff for the 78 tests that were performed ($p = 6.4E−04$). We also considered if the direction of effects between discovery and replication samples was concordant. For ethnic-specific replication, we used EA replication samples for EA or H/L discovery meta-analyses, while for AA replication, we used AA replication samples, a Bonferroni-corrected $p < 2.1E−03$, and consistency in direction of effects. We also attempted to replicate DMPs from a published study [11]. This published study used eGFR instead of DMPs as a predictor in models and therefore the estimates were not comparable to our study.

In silico annotation using eFORGE and eFORGE-TF, and pathway analyses

We performed functional overlap analysis of DMPs with eFORGE version 2.0, analyzing the default top 1000 probes from EA, AA, H/L, and all-ethnic probe sets for overlap enrichment across DNase I hotspots from the Roadmap Epigenomics Consortium [32, 33]. To ascertain whether the observed enrichment was robust and associated with top probe sets below the EWAS significance threshold, we applied additional eFORGE analyses across the 5 top EA CpG sets (comprising CpGs 1–1000, 1000–2000, 2000–3000, 3000–4000, 4000–5000, ordered by $p$-value), to detect overlap enrichment across DNase I hotspots from the Roadmap Epigenomics Consortium [32, 33]. We thus performed integrative epigenomics analyses on data from the Roadmap Epigenomics consortium [34] using the eFORGE framework (https://eforge.altiusinstitute.org/) [32, 33]. To further understand eFORGE enrichment results, we performed TF motif analysis on the probes underlying eFORGE tissue-specific enrichment signal for kidney. We used the eFORGE-TF module, seeking to uncover...
the main TF motifs associated with our DNase I hotspot enrichment for top sites [33]. We used PANTHER analysis via the AmiGO framework to uncover pathways associated with the identified TF motifs [35].

meQTL data in the whole blood

We used publicly available data from the biobank-based integrative omics study (BIOS) QTL database [36, 37] and cis-meQTL data from the FHS (n = 4170 participants, 450K DNAm panel, 1000G imputed data, defined within 2-Mb window) [38]. We also used meQTL data from mQTLdb [39].

meQTL in normal kidney tissue

We analyzed a total of 211 individuals with matching kidney genome-epigenome information from TRAN Scrip tome of renaL humAn TissuE Study (TRANSLATE), an extension of the TRANSLATE study (TRANSLATE E-T). Renal gEne expresSion and PredisPotion to cardiovascular and kidNey Disease (RESPOND), and molecular analysis of mechanisms regulating gene expression in post-ischemic injury to renal allograft (REPAIR) (n = 192), in addition to normal samples with available genotype and kidney DNAm profiles from the National Institutes of Health (NIH) Tissue Cancer Genome Atlas (TCGA) (n = 19) [40–43]. Kidney tissue samples from TRANSLATE, RESPOND, and TCGA were taken from the healthy, unaffected part of the organ after elective nephrectomies, and renal specimens in TRANSLATE-T and REPAIR studies were collected as pre-implantation biopsies from deceased donors’ kidneys before transplantation [40, 42, 43]. Information on local recruitment teams, genotyping, and DNAm methods are included in Additional file 2: Supplementary Methods. In total, 374,826 CpG sites were available for further analyses after quality control filters. For post-EWAS DMP analysis, we conducted kidney cis-meQTL analysis on 62 DMPs (out of 374,826) from the study. For GWAS SNP analysis, we conducted kidney cis-meQTL analysis on a set of published eGFR GWAS SNPs [44]. The cis-meQTL analysis was conducted on 195 kidney DNA samples that passed all quality control criteria. For analysis, we used the FastQTL pipeline [45]. We used normalized M-values and genotype information for all genotyped and imputed variants passing the quality control filters under an additive mode of inheritance. Regression models included age, sex, genotyping array, source of tissue indicator (nephrectomy/kidney biopsy), the top three PCs derived from genotyped autosomal variants (genotype PCs), and six PCs derived from methylation array control probes (methylation PCs). The FastQTL cis-region was defined as ±1Mb from each tested CpG/SNP position. For EWAS DMP blood-kidney meQTL comparison analysis, we compared DMP+kidney meQTL CpGs to blood meQTL CpGs from mQTLdb and calculated the percentage of overlap for data from both sources.

eFORAGE analysis for kidney meQTL CpGs

Standard eFORAGE analyses were performed using default settings. For kidney meQTL CpGs, we analyzed eGFR SNPs overlapping kidney DNase I hotspots (top category from https://forge2.altiusinstitute.org/files/0x45A546E8E9C11EA8198793F5BFE3F98/index.html). We searched for these SNPs in the kidney meQTL file, obtaining preliminary associated meQTL CpGs (nominal p-value < 0.05) and preliminary non-associated meQTL CpGs (nominal p-value > 0.95), for the same set of SNPs. Standard eFORAGE analyses were performed on both of these sets.

Results

Overview of EWAS results

The study design is shown in Fig. 1 and Additional file 1: Fig. S1. Discovery EWAS meta-analyses included up to 5428 individuals, with 2879 AA, 1737 EA, and 812 H/L participants (Additional file 3: Table S1). Quality control of DNAm data and protocol analyses for each study are shown in Additional file 3: Table S2. We performed analyses within each study and ethnic group using standard statistical protocols followed by meta-analyses of the overall samples and within each ethnicity. The quantile-quantile and Manhattan plots for meta-analyses are shown in Additional file 1: Fig. S2. Lambdas for meta-analyses were 0.987 (AA), 1.001 (EA), 1.065 (H/L), and 1.194 (trans-ethnic meta-analysis). Across our discovery analyses, we identified a total of 93 DMPs associated with eGFR at an FDR of 0.05. Of these, 78 DMPs were identified in trans-ethnic meta-analyses, 23 DMPs in the meta-analysis of AA, 5 in the meta-analysis of H/L, and 5 in the meta-analysis of EA, with some overlap in DMP findings between trans-ethnic and ethnic-specific results.

Indeed, most of the DMPs identified in trans-ethnic analyses were also present in one or more ethnic groups (Fig. 1a). However, the overlap between DMPs of AA and EA was small. Trans-ethnic replication included up to 8109 participants from three studies (Generation Scotland, CATHGEN, and HyperGEN), composed of participants of EA and AA (9% of replication samples) (Additional file 1: Fig. S1) [11]. Replication of EA or H/L findings included Generation Scotland and CATHGEN EA samples (n = 7349) and AA included CATHGEN and HyperGEN samples (n = 760). Among the significantly identified trans-ethnic DMPs (Bonferroni corrected p-value and consistent direction of effects), 13 of 78 replicated (Table 1), and among the significantly identified DMPs in ethnic-specific meta-analyses, 1 of 5 AA DMPs replicated (cg14871770 at CYP2C9) (Additional file 3:
Table S3). Despite independent replication in AA, cg14871770 overlapped with trans-ethnic replicated DMPs (Fig. 1b). Twelve additional DMPs had replication (with consistent direction of effects) in trans-ethnic meta-analyses (Additional file 3: Table S3a). The DMP cg14871770 overlapped between trans-ethnic and AA meta-analyses, and the cg17944885 at the ZNF20-ZNF788P locus was previously described [11]. Several of the replicated DMPs were expression quantitative trait methylation loci (eQTM) or cis-meQTL CpGs in whole blood in FHS and BIOS QTL (Table 1) [36–38]. Replication results for all DMPs are shown in Additional file 3: Table S3a.
To understand the regulatory context of our DMPs, we annotated the 13 replicated significant DMPs with the closest gene and other information including epigenomic annotations. One of our DMPs peaks, tissue-specific gene expression via RNA-seq, and closest gene and other information including epigenomic annotations from the ENCODE consortium including LDB2, ZNF788P.*

Overlap of eGFR-associated DMPs with genes and regulatory elements

To understand the regulatory context of our DMPs, we annotated the 13 replicated significant DMPs with the closest gene and other information including epigenomic peaks, tissue-specific gene expression via RNA-seq, and chromatin interaction annotations. One of our DMPs (cg11789371) is located in an intron of HSP90AA1, a gene involved in protecting kidney tissue from inflammation, ischemia, and oxidative damage, and assisting with cellular repair (Fig. 2) [46]. HSP90 (heat shock protein 90) has a physiological role in eGFR regulation through the nitric oxide pathway and is a drug target candidate for kidney diseases [46]. Indeed, inhibition of HSP90 has been shown to reduce eGFR in animal models [46].

DMPs for eGFR are enriched for kidney regulatory function related to kidney development

To further understand the regulatory potential and chromatin context of our EWAS findings across different tissues, we performed integrative epigenomics analyses on data from the Roadmap Epigenomics consortium [34] using the eFORGE framework (https://eforge.altiusinstitute.org/) [32, 33]. We found enrichment for kidney-specific DNase I hotspots, which has also been described in GWAS of eGFR (Fig. 3a, d) [4, 47].

Table 1 Main findings from trans-ethnic EWAS meta-analyses for 13 replicated DMPs

| DMP | Chr | Position (hg38) | Gene | Effect | p        | 450K EPIC 850K | Effect | p        | Combined discovery and replication |
|-----|-----|----------------|------|--------|----------|---------------|--------|----------|------------------------------------|
| cg13235761^ | 1 | 203,592,452 |  | −37.61 | 4.33E−06 | 3048 2378 | −25.28 | 5.08E−05 | −29.81 1.88E−09 |
| cg26099045%^ | 2 | 64,064,666 |  | 14.82 | 5.74E−06 | 3050 2378 | 11.24 | 1.01E−04 | 12.81 3.26E−09 |
| cg04428662% | 4 | 2,932,461 | MFS10 | −30.84 | 3.21E−06 | 3049 2378 | −34.94 | 3.43E−07 | −32.82 5.51E−12 |
| cg23174201% | 5 | 151,674,695 | SPARC | −35.83 | 5.10E−06 | 3050 2378 | −34.87 | 1.72E−07 | −35.27 4.02E−12 |
| cg17170437 | 6 | 44,229,461 | SLC29A1 | −47.62 | 1.96E−06 | N/A 2378 | −18.62 | 1.49E−04 | −24.24 3.80E−08 |
| cg14871770 | 10 | 96,658,622 | CYP2C9, CYP19 | −53.12 | 1.98E−06 | N/A 2378 | −25.87 | 1.70E−04 | −33.36 1.23E−08 |
| cg02157636 | 11 | 68,709,367 | TESMIN | −53.21 | 2.15E−06 | N/A 2378 | −27.25 | 1.14E−05 | −33.33 8.58E−10 |
| cg26059141% | 11 | 75,402,116 | RPS3 | −39.14 | 3.40E−06 | 3050 2378 | −33.42 | 1.78E−05 | −36.06 2.90E−10 |
| cg22593432% | 13 | 32,001,768 |  | −29.91 | 1.54E−07 | 3050 2378 | −17.66 | 1.79E−04 | −22.64 4.61E−10 |
| cg11789371% | 14 | 102,085,048 | HSP90AA1 | −35.40 | 4.44E−06 | 3050 2378 | −22.66 | 3.58E−04 | −27.80 1.41E−08 |
| cg05796561 | 18 | 57,128,273 |  | −45.24 | 2.96E−06 | N/A 2378 | −22.91 | 1.69E−04 | −29.24 1.41E−08 |
| cg17944885% | 19 | 12,114,920 | ZNF20, ZNF788P | −32.71 | 1.41E−09 | 3048 2378 | −16.34 | 5.67E−07 | −20.72 1.24E−13 |
| cg15787712% | 19 | 13,837,429 | LOC284454, MR23 | −37.60 | 6.00E−09 | 3050 2378 | −28.65 | 2.06E−05 | −33.30 9.01E−13 |

Table S3a (trans-ethnic analyses) and Table S3b (ethnic-specific analyses). Forest plots for each of the replicated DMPs are shown in Additional file 1: Fig. S3. We also replicated 6 DMPs identified in a prior publication (DMPs located at genes DAZAP1, KIAA1549L, TUBGCP4/ZSCAN29, JAZF1, ZNF20-ZNF788P, LDB2) (Additional file 3: Table S4) [11].

p-value for replication < 6.4E−04 (Bonferroni correction for 78 DMPs tested). Four of the 13 DMPs that replicated were non-450K probes (cg17170437, cg14871770, cg02157636, cg05796561). Results from combined discovery and replication are also shown.

GS Generation Scotland, N/A not applicable

^cis-meQTL from the BIOS QTL [36, 37]

 cis-Expression quantitative trait methylation (eQTM) data from the BIOS QTL (enrichment p-value < 0.01) [36, 37]

* cis-meQTL in the Framingham Heart Study [38]
transcription factor (TF) motifs (including motifs for OSR1, OSR2, TBX1, and PAX2). OSR1, OSR2, TBX1, and PAX2 have all been shown to have roles in kidney development (Fig. 3c) [48–50]. To further understand the pathways underlying TF motif enrichment, we performed PANTHER pathway analysis using significant TF motifs [35]. This analysis uncovered pathways associated with kidney development, including metanephros development ($8.1 \times 10^{-3}$), mesonephros development ($9.7 \times 10^{-3}$), and retinoic acid receptor signaling pathway ($6.4 \times 10^{-3}$) (Additional file 3: Table S5). Overall, these findings suggest that epigenetic changes related to eGFR are enriched in kidney regulatory regions and pathways related to kidney development.
Fig. 3 (See legend on next page.)

(a) DMPs analyzed across samples for eGFR EWAS data

(b) SNPs in DNase I hotspots across samples for eGFR GWAS data

(c) Probeset TF motif associations

(d) Aggregate eFORGE analysis

Fig. 3
eGFR GWAS variants, meQTL CpGs, and tissue-specific DNase I hotspots

Previous reports and our analyses confirm that eGFR GWAS variants are enriched for kidney DNase I hotspots (Fig. 3b, Additional file 1: Fig. S4) [4, 47], indicating that genotypes might function through kidney-specific regulatory pathways, potentially converging with identified EWAS DMPs through mechanisms which are not fully understood. To explore this further, we used mQTLdb [39] to identify meQTL CpG targets of eGFR-associated GWAS SNPs obtained from the GWAS catalog [44, 51]. These significant meQTL CpGs linked to GWAS SNPs revealed significant overlap with eGFR EWAS sites in our study (p < 0.002, Fig. 4a) and presented significant enrichment for kidney, renal cortex, and renal pelvis, among other tissues (Fig. 4b, d, Additional file 1: Fig. S5). These results support a model in which meQTL CpGs linked to eGFR GWAS SNPs overlap with EWAS DMPs and tend to localize to kidney DNase I hotspots, with potential involvement in their regulatory action (Fig. 4c). To assess these findings in human kidney tissue, we analyzed meQTL CpGs of eGFR SNPs localizing to kidney DNase I hotspots. The meQTL CpGs were identified from DNAm data from 195 normal kidney tissue samples (acquired from elective nephrectomies—taking the non-cancer affected segment—or from kidney donors) as reported previously [40–42]. While kidney meQTL CpGs not associated with eGFR SNPs (nominal p-value > 0.95) showed no enrichment in kidney DNase I hotspots (Additional file 1: Fig. S6), the kidney meQTL CpGs with a nominal p-value < 0.05 were enriched in kidney DNase I hotspots (Additional file 1: Fig. S7) for the same GWAS SNPs. These results support the aforementioned findings using eGFR SNP-associated meQTL CpGs from mQTLdb [39]. Additionally, we evaluated the overlap between kidney meQTL CpGs and blood meQTL CpGs (mQTLdb) for our top EWAS DMPs. The majority (58.3%) of these kidney meQTL CpGs were also mQTL CpGs in blood.

Discussion

Main findings

This study used trans-ethnic and ethnic-specific analyses of multi-ethnic cohorts to identify and replicate DMPs associated with eGFR. Our main findings include replicated DMPs at 13 sites from trans-ethnic analyses and 1 DMP from AA-specific analyses, which may reflect the larger AA discovery sample compared to other ethnic groups. All associations were newly identified in this study, except for the ZNF20-ZNF788P locus, which was previously described in two separate eGFR EWAS, including a population-based cohort and an HIV-infected cohort [11, 52]. Several DMPs were associated with accessible chromatin sites in kidney tissue, suggesting a regulatory role for these sites. Overall, our study identified 12 previously unreported DMPs that replicated, in addition to 6 previously published DMPs for eGFR among our 93 discovery DMPs [11]. These findings support an association between DNAm and eGFR, reflecting a convergence of lifetime influences of genetic effects, lifestyle, behaviors, and environmental exposures [53–55].

Differences between ethnicities

Our approach of studying multi-ethnic groups identified DMPs from trans-ethnic and ethnic-specific analyses. This approach contrasts with a prior study that performed separate EA and AA analyses with cross-replication across these groups [11]. Across our two largest ethnic-specific samples (AA and EA), there was little overlap between DMPs from discovery analysis, although several discovery ethnic-identified DMPs did overlap with trans-ethnic findings (Fig. 1a). The DMP cg14871770, which is located between CYP2C9 and CYP2C19 (cytochrome P450 family 2 subfamily C members 9 and 19), was identified both in ethnic-specific and trans-ethnic analysis. The closest GWAS SNP to this DMP is rs4110517, a SNP associated with blood pressure identified in a multi-ethnic cohort [56]. CYP2C9 and CYP2C19 encode members of the cytochrome P450 superfamily of enzymes. Cytochrome P450...
**Fig. 4** (See legend on next page.)

(a.) Histogram of 1000 random background simulations

- **Background eGFR GWAS**
  - Mean = 0.912
  - SD = 1.998
  - P = 0.002

(b.) Histogram of 1000 random background simulations

- **eGFR GWAS**
  - Mean = 67.021
  - SD = 24.754
  - P < 0.001

(c.) Kidney DNase I hotspots

(d.) FDR q-value

- q < 0.01
- q < 0.05
- non-sig

- -log10 binomial p-value

**Legend:**

- Fetal Kidney
- Fetal Kidney Left
- Fetal Kidney Right
- Fetal Renal Cortex
- Fetal Renal Cortex Left
- Fetal Renal Cortex Right
- Fetal Renal Pelvis
- Fetal Renal Pelvis Left
- Fetal Renal Pelvis Right
proteins are monoxygenases that catalyze many reactions involved in the synthesis of cholesterol, other steroids, and other lipids, and in drug metabolism. CYP2C19 genotype is associated with the metabolism of compounds influencing both renal function and hypertension [57]. The relevance of ethnic-specific findings to clinical phenotypes will need further evaluation in studies with larger samples inclusive of multiple ethnicities to define improved DNAm signatures for eGFR that may be unique to one single ancestry. Our findings suggest reduced utility of DNAm biomarkers for eGFR in diverse populations if discovery EWAS is performed in a single homogenous population.

Relevance of HSP90AA1 locus
Among other findings, we report a replicated trans-ethnic DMP (cg11789371) associated with eGFR that localizes to an intron of HSP90AA1 (heat shock protein 90 alpha family class A member 1), a gene expressed in podocytes, parietal epithelial cells, proximal tubular cells, endothelium, and mesangial cells in normal kidney tissue, with gene expression increasing in glomerulonephritis and acute kidney injury [46]. The gene product, HSP90, regulates renal blood flow and eGFR through nitric oxide metabolism and plays a role in protein folding [46]. HSP90 and other heat shock proteins are candidate drug targets for a variety of kidney diseases [46, 58]. Treatment with radicicol, an inhibitor of HSP90, has been shown to reduce eGFR in animal models [58]. DMP cg11789371 overlaps an accessible chromatin region in kidney cells, and a GeneHancer interacting site contacting the promoters of WDR20 and HSP90AA1. These and other different annotations point to a potential regulatory role in kidney tissue for this eGFR-associated DMP.

Integrative epigenomics and pathway analysis
We detected a significant overlap of eGFR GWAS SNP-associated meQTL CpGs with kidney DNase I hotspots. Additionally, our eGFR EWAS DMPs were enriched for kidney DNase I hotspots. These findings suggest potential links between the regulatory action of both genotypes and epigenetic DNAm elements. Importantly, these integrative epigenomic analyses considered all tissues available instead of only kidney tissue.

While some regions of the methylome show tissue specificity [59], EWAS have also shown that some DMPs are shared across different tissues, e.g., the AHRR locus DNAm patterns in response to smoking are shared across multiple tissues [60]. Pan-tissue findings for AHRR suggest similar underlying pathways in response to the same environmental stimulus [60]. Regarding discrepancies in trans-ethnic results, both genetic and environmental differences could be at play, potentially interacting with each other. In this context, our findings of both a kidney-specific DNase I hotspot and GWAS meQTL enrichment for a whole blood-based EWAS could be due to both genetic and environmental origins and warrant further research of kidney tissue DNAm in association with eGFR. Indeed, such results raise the intriguing hypothesis that tissue-specific enrichments observed separately in GWAS and EWAS might be related by the same genomic variants, thus aiding the integration of both approaches (Fig. 4a). It is important to highlight that both whole blood and kidney tissue eQTLs were obtained from individuals of European ethnicity [40–43].

The identification of eGFR DNAm signature-associated pathways is an important step towards characterizing epigenetic mechanisms for this physiologic trait and may provide clues to underlying mechanisms for CKD. Identified DMPs highlight an association with pathways of kidney development, which can influence nephron endowment at birth and subsequent CKD risk [48–50, 61]. Our in silico results were influenced by our DNAm findings in healthy adult kidney. Therefore, pathway results from this study support epigenetic effects during developmental windows with long-term influence on eGFR, which warrant further investigation.
Limitations
This study is limited by the use of whole blood as the main tissue (chosen due to its availability). The discovery datasets included both the Illumina 450K and the EPIC 850K arrays, which contributed to differences in sample sizes and power to detect associations for some CpGs (Table 1). Post hoc power analyses suggest adequate power to detect DNAm differences of the range observed in the study (Additional file 2: Supplementary Methods). The methods for normalization of the DNAm beta values and study-specific quality control varied (Additional file 3: Table S2). However, we applied standardized protocols for data harmonization and statistical analyses in addition to stringent quality control as part of our meta-analyses. Our findings showed no heterogeneity of effects across studies (Additional file 1: Fig. S3), suggesting that results are robust to study-specific quality control and normalization procedures. Additionally, our reported DMPs replicated in independent samples, further validating our results. It is important to consider whether these DNAm sites associated with eGFR in blood are also applicable to effects in kidney tissue. We attempted to answer this by examining our identified meQTLs in normal kidney tissue in the TRANSLATE study. However, kidney tissue studies still have small sample sizes and lack ethnic diversity. While epigenomic mapping consortia such as Roadmap Epigenomics and ENCODE have made important steps to increase the free availability of a wide range of tissue and cell type-specific datasets, the important issue of including additional epigenomics mapping data sets for other ancestries remains to be addressed. It is important to highlight that we only observe eFORGE kidney enrichment for top EA probes. While this kidney-specific DNase I hotspot enrichment has been further confirmed by analyzing ranked DMPs in study p-value order (Fig. 3d), it is not apparent for DMPs from other ethnicities, or for DMPs for analysis comprising all ethnicities. DNase-seq datasets for these samples originate from the Roadmap Epigenomics consortium [34], which focused mainly on tissue samples obtained from EA individuals. Without datasets from diverse ethnic groups, it will be difficult to conclusively study inter-ethnic epigenomic variability or perform tissue-specific analyses for loci from GWAS and EWAS performed on individuals of non-European origin.

Conclusions
We identified trans-ethnic and ethnic-specific differential DNAm positions, validated prior published associations, and showed that several eGFR DMPs identified in this study replicated in independent samples. We have also shown that some of the DMPs are meQTL CpGs, many of which are associated with pathways relevant for kidney tissue regulation and development. Our findings include a DMP at HSP90AA1, a gene involved in the regulation of eGFR in kidney tissue. Identification of trans-ethnic and ethnic-specific DMPs and elucidation of their potential functional impact are preliminary steps towards identifying disease-associated epigenetic mechanisms that are specific to a particular population or shared across different populations.

Supplementary Information
Supplementary information accompanies this paper at https://doi.org/10.1186/s13073-021-00877-z.

Acknowledgements
The authors thank the CATHGEN Steering Committee, the faculty and staff of the Duke cardiac catheterization lab, and the CATHGEN participants for making this work possible. We are grateful to all the families who took part, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team that includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants, and nurses. The authors also wish to thank the staff and participants of the JHS. The authors thank the WHI investigators and staff for their dedication and the study participants for making the program possible.

Authors’ contributions
Advised or performed statistical analyses (CEB, AB, MKL, JCM, MDS, RJ, DLM, XX, AP, JME, APM, REM, SB, SJL). Study design (CEB, SJL, NF). Generated data (YL, RPT, DVDB, EAW, LH, HJK, LR, LL, EL, PD, KE, WKE, SS, HKT, XJ, FJC, AAB, SSR, MI, DKA, ERH, JIR, AC, CH, SH, MT). CEB and NF wrote the paper with contributions from all other authors. All authors read and approved the final manuscript.

Funding
This study was supported by the National Institutes of Health R01-M012765, R01-DK117445, and R21-HL140385 to NF. The TRANSLATE studies and MT are supported by British Heart Foundation project grants PG/17/35/33301 and PG/19/16/34270, Kidney Research UK grant RP_017_20180302, and Medical University of Silesia grants KNW-1-152/N/7/K and KNW-1-171/N/6/K, WT026194. DLM and REM are supported by Alzheimer’s Research UK major project grant ARUK-PG20178-10. CH is supported by an MRC University Unit Programme Grant MC_UU_000710 (QTL in Health and Disease). SJL and MKL are supported by the Intramural Program of the NIH, National Institute of Environmental Health Sciences (Z01 ES043012), SB acknowledges funding from the Wellcome Trust (218274/Z/19/Z). CEB and SJL are supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH. The CATHGEN study was supported by NIH grants HL059587 and HL036587. Generation Scotland (GS) received core support from the Chief Scientist Office of the Scottish Government Health Directors (CZD/16/6) and the Scottish Funding Council (HR03006). Genotyping and DNA methylation profiling of the GS samples was carried out by the Genetics Core Laboratory
at the Clinical Research Facility, University of Edinburgh, Edinburgh, Scotland, and was funded by the Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award “Stratifying Resilience and Depression Longitudinally” (STRADL), Reference 104036/2/14/Z).

Support for the Multi-Ethnic Study of Atherosclerosis (MESA) projects are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts 5N02HD020001, HHSN268201500003, N01-HC-95159, 75N02HD020005, N01-HC-95160, 75N02HD020006, N01-HC-95161, 75N02HD020002, N01-HC-95162, 75N02HD020003, N01-HC-95163, 75N02HD020004, N01-HC-95164, 75N02HD020007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-000040, UL1-TR-000179, and UL1-TR-000140. Also supported in part by the National Center for Advancing Translational Sciences, CTSA grant UL1TR001881, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The TOPMed MESA Multi-Omics project was conducted by the University of Washington and LABioMed (HHSN2682015000031/HHSN2682000004).

The Jackson Heart Study (JHS) is supported and conducted in collaboration with Jackson State University (HHSN268201800013), Tougaloo College (HHSN268201800014), the Mississippi State Department of Health (HHSN268201800015), and the University of Mississippi Medical Center (HHSN268201800010, HHSN268201800011, and HHSN268201800012) contracts from the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute for Minority Health and Health Disparities (NMHD). The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the US Department of Health and Human Services. The funders had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services, through contracts HHSN268201600018C, HHSN26820160001C, HHSN26820160002C, HHSN26820160003C, and HHSN26820160004C. A full listing of WHI investigators can be found at https://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHIRegistrant%20Long%20List.pdf.

The Hypertension Genetic Epidemiology Network (HyperGEN) Study is part of the National Heart, Lung, and Blood Institute (NHLBI) Family Blood Pressure Program; collection of the data represented here was supported by grants U01 HL054472, U01 HL054473, U01 HL054495, and U01 HL054509. The HyperGEN Genetics of Left Ventricular Hypertrophy Study was supported by NHLBI grant R01 HL055673 with whole-genome-sequencing made possible by supplement -851. The epigenetic data in HyperGEN was funded by an American Heart Association Cardiovascular Genome-Phenome Study Pathway Grant Award 15SPGPSP23890000.

Availability of data and materials

Investigators interested in retrieving the controlled-access data at dbGaP for WHI-BAAC2 and MESA should apply using identifiers phs002020.v10.p3 and phs001416.v1.p, respectively (available online at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002020.v10.p3 and https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001416.v1.p). The summary results from this study are placed in dbGaP with accession number phs000930.v8.p1. While awaiting data release via dbGaP, data are available at https://sph.unc.edu/wp-content/uploads/sites/11/2021/02/EWAS_COGENT.tar. Access to the WHI-EMPC DNA methylation dataset is available upon request to the corresponding author from Ammous et al. [32]. Currently, JHS and WHI datasets are available through a scientific review application process directed to each respective study participation and presentation committee. Data can be obtained from the coordinating center of WHI and JHS after signing a data use agreement with the study. For research projects that meet the rules for access, CATHGEN data are available from the CATHGEN Steering Committee. Relevant requests are to be sent to the contact person at the CATHGEN Steering Committee, Melissa Hurdle (melissa.hurdle@duke.edu). According to the terms of consent for GS participants, access to individual-level data (omics and phenotypes) must be reviewed by the GS Access Committee. Applications should be made to access@generationscotand.org. The source code used for the eFORCE analyses is publicly available at https://github.com/charlesbreeze/eFORCE [32] and https://github.com/charlesbreeze/eFORCE-TF [33].

Declarations

Ethics approval and consent to participate

All participants provided informed written consent to participate in these studies, and studies were approved by the Institutional Review Boards at each recruiting site.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA. 2UCL Cancer Institute, University College London, London WC1E 6BT, UK. 3Atlantis Institute for Biomedical Sciences, Seattle, WA 98121, USA. 4Department of Biostatistics, University of North Carolina, Chapel Hill, NC, 27516, USA. 5Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, NC 27709, USA. 6Division of Biomedical Informatics and Personalized Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, USA. 7Division of Cardiovascular Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. 8Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh,Crewe Road, Edinburgh EH4 2XU, UK. 9Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC 27701, USA. 10Department of Biostatistics, University of Alabama, Birmingham, AL, USA. 11Department of Public Health Sciences and Medicine, Loyola University Chicago, Maywood, IL, USA. 12Division of Nephrology and Hypertension, Loyola University Chicago, Maywood, IL, USA. 13Department of Genetics, National Cancer Institute, Chapel Hill, NC, USA. 14Department of Pathology & Laboratory Medicine, Larner College of Medicine, University of Vermont, Burlington, VT, USA. 15Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC, USA. 16Department of Biochemistry, Larner College of Medicine, University of Vermont, Burlington, VT, USA. 17Center for Genetic Epidemiology, Department of Preventive Medicine, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA, USA. 18Division of Cardiology, Department of Medicine, School of Medicine, Duke University, Durham, NC, USA. 19Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. 20Center for Global Oncology, Institute of Global Health, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. 21Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA. 22Department of Medicine, University of North Carolina, Chapel Hill, NC, 27599, USA. 23School of Health and Life Sciences, Federation University Australia, Ballarat, VIC, Australia. 24Department of Physiology, University of Melbourne, Parkville, VIC, Australia. 25Department of Cardiovascular Sciences, University of Leicester, Leicester, UK. 26Laboratory of Environmental Epigenetics, Departments of Environmental Health Sciences and Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA. 27Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA. 28Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, The University of Manchester, Manchester, UK. 29Department of Epidemiology, University of Alabama at Birmingham, Birmingham, AL, USA. 30College of Public Health, University of Kentucky, Lexington, KY, USA. 31Duke VA Health System, Durham, NC 27705, USA. 32The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, USA. 33Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA. 34MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Crewe Road, Edinburgh EH4 2XU, UK. 35Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA.
Smyth LJ, McKay GJ, Maxwell AP, McKnight AJ. DNA hypermethylation and...t: https://doi.org/10.1038/s41467-017-01297-5.

Sarhan R, Robinson B, Abbott KC, Agodoa LYC, Bhave N, Bragg-Gresham J, Taudt A, Colomé-Tatché M, Johannes F. Genetic sources of population...well-known for their potential for genetic research on health and illness. Int J Epidemiol. 2013;42(3):689–700. https://doi.org/10.1093/ije/dys084.

Breeze et al. Genome Medicine (2021) 13:74
Bioinformatics. 2017;33(4):558–60. https://doi.org/10.1093/bioinformatics/btw691.

25. Touleimat N, Tost J. Complete pipeline for Infinium® Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics. 2012;4(3):325–41. https://doi.org/10.2174/1878341012440318

26. Piddrew R, Y Wong CC, Volta M, Lunnun K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450k methylation array data. BMC Genomics. 2013;14:293, 1, doi: https://doi.org/10.1186/1471-2164-14-293.

27. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Bioinformatics. 2007;23(16):276–80. https://doi.org/10.1093/bioinformatics/btm229.

28. Chen Y, Lemire M, Choufani S, Butler DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymeric CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013;8(2):203–9. https://doi.org/10.4161/epi.23470.

29. Houseman EA, Accomando WP, Koester DC, Christensen BC, Marist CJ, Nelson HH, Wieneke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13(1):186. https://doi.org/10.1186/1471-2105-13-186.

30. Wise I, Preest PR, Glyda M, Bogdanski P, Zukowska-Szczechowska E, Berzunzi C, Woolf AS, Samani NJ, Charchar FJ, Tomaszewski M. Molecular insights into genome-wide association studies of chronic kidney disease-defining traits. Nat Commun. 2018;9(1):4800. https://doi.org/10.1038/s41467-018-07260-4.

31. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature. 2013;499(7456):43–9. https://doi.org/10.1038/nature12222.

32. Rowland J, Akkarov A, Eales J, Xu X, Dormer JP, Guo H, Denniff M, Jiang X, Ranjaz P, Nazgiwicz A, Preest PR, Antczak A, Sulinska M, Wise IA, Zukowska-Szczechowska E, Bogdanski P, Woolf AS, Samani NJ, Charchar FJ, Tomaszewski M. Uncovering genetic mechanisms of kidney aging through transcriptomics, genomics, and epigenomics. Kidney Int. 2019;95(5):624–35. https://doi.org/10.1016/j.kint.2018.10.029.

33. Tomaszewski M, Eales J, Denniff M, Myers S, Cheew GS, Nelson CP, Christophidou P, Desai A, Büsct C, Wlonar J, Musialik K, Jozwiai D, Debric R, Dominiczak AF, Navis G, van Gilst WH, van der Harst P, Samani NJ, Hamra S, Bogdanski P, Zukowska-Szczechowska E, Charchar FJ. Renal mechanisms of association between fibroblast growth factor 1 and blood pressure. J Am Soc Nephrol. 2015;26(12):3151–60. https://doi.org/10.1681/ASN.2014121211.

34. Hellwegne JN, Velez Edwards DR, Giri A, Qiu C, Park J, Torstenson ES, Keaton JM, OWLD, Robinson-Cohen C, Chung CP, Roumie CL, Klarin D, Darrnauer SM, DuVall SL, Sieve E, Akwo EA, Wutke M, Gorski M, Li M, Li Y, Gaziwn JM, Piong PF, Tsao PS, O'Donnell CJ, Koveyse CP, Patiaro C, Kritten A, Suzak T, Edwards TL, Hung AM. Mapping eQTL loci to the renal transcription and phenotype in the VA million veteran program. Nat Commun. 2019;10(1):3842. https://doi.org/10.1038/s41467-019-11704-w.

35. Ongen H, Bull A, Brown AA, Demirzakis ET, Delaune A. Fast and efficient QTL mapper for thousands of molecular phenotypes. Bioinformatics. 2016;32(10):1749–55. https://doi.org/10.1093/bioinformatics/btv222.

36. Chebotareva N, Bobkova I, Shilov E. Heat shock proteins and kidney disease: perspectives of HSP therapy. Cell Stress Chaperones. 2017;22(2):319–43. https://doi.org/10.1007/s12192-017-0790-0.

37. Wutke M, Li Y, Li M, Sieber KS, Feitosa MF, Gorski M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. Nat Genet. 2019;51(6):957–72. https://doi.org/10.1038/s41588-019-0407-x.

38. Xu J, Liu H, Chai OH, Lan Y, Jiang R. Osr1 interacts synergistically with Wt1 to regulate kidney organogenesis. PLoS One. 2016;11(7):e0159597. https://doi.org/10.1371/journal.pone.0159597.

39. Jiang H, Li L, Yang H, Bai Y, Jiang H, Li Y. Pax2 may play a role in kidney development by regulating the expression of TBX1. Mol Biol Rep. 2014;41(1):2491–9. https://doi.org/10.1007/s11033-014-3639-y.

40. Alam-Faruque Y, Hill DP, Dimmer EC, Harris MA, Foulger RE, Tweedie S, Attrill H, Howe DG, Thomas SR, Davidson W, Aslake JA, Mungall CJ, Ongen H, Wellin D, Mark J, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. Nat Genet. 2019;51(6):957–72. https://doi.org/10.1038/s41588-019-0407-x.

41. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Morgan NB, Cookson M, Shelley M, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014;42(D1):D1001–6. https://doi.org/10.1093/nar/gkt1229.

42. Chen J, Huang Y, Qiu H, Mathur R, Gwinn M, So-Armah K, Freiberg MS, Justice AC, Xu K, Marconi VC, Sun Y. Epigenetic associations with estimated glomerular filtration rate among men with human immunodeficiency virus infection. Clin Infect Dis. 2020;70(4):667–71. https://doi.org/10.1093/cid/ciaa173.
57. Imamura CK, Furihata K, Okamoto S, Tanigawara Y. Impact of cytochrome P450 2C19 polymorphisms on the pharmacokinetics of tacrolimus when coadministered with voriconazole. J Clin Pharmacol. 2016;56(4):408–13. https://doi.org/10.1002/jcph.605.

58. Ramírez V, Mejía-Vilet JM, Hernández D, Gamba G, Bobadilla NA. Redicicol, a heat shock protein 90 inhibitor, reduces glomerular filtration rate. Am J Physiol Renal Physiol. 2008;295(4):F1044–51. https://doi.org/10.1152/ajprenal.90278.2008.

59. Lowe R, Siodkowicz G, Goldman N, Rakyan VK. The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. Epigenetics. 2015;0, 10, 4, 274, 281, doi: https://doi.org/10.1080/15592294.2014.1003744.

60. Tsai P-C, Glastonbury CA, Eliot MN, Bollepalli S, Yet I, Castillo-Fernandez JE, Camero-Montoro E, Hardiman T, Martin TC, Vickers A, Mangino M, Ward K, Pietiläinen KH, Deloukas P, Spector TD, Vickers A, Olsson I, Kelsey KT, Small KS, Bell JT. Smoking induces coordinated DNA methylation and gene expression changes in adipose tissue with consequences for metabolic health. Clin Epigenetics. 2018;10(1):126. https://doi.org/10.1186/s13148-018-0558-0.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.