Genetic Association Studies Identify *CUBN* as a Gene Locus for Albuminuria

SUPPLEMENTARY MATERIAL

Overview

Supplementary Table 1a-Genotyping and Imputation Information in CKDGen Consortium								
Study Name	**Array type**	**Genotype calling**	**QC filters for genotyped SNPs used for imputation**	**No of SNPs used for imputation**	**Imputation**	**Imputation Backbone (NCBI build)**	**Filtering of imputed genotypes**	**Data management and statistical analysis**
AGES	Illumina 370 HuCNV	Illumina Beadstudio	call rate <95%, pHWE<10E-6	325094	MACH version 1.0.16	HapMap release 22 (build 36)	none	R
Amish Studies	Affymetrix 500K	BRLMM	MAF < 1%, non-HapMap,call rate <95%, pHWE<10E-6	338,598	MACH version 1.0.15	HapMap release 22 (build 36)	none	Measured genotype accounting for polygenic component
ARIC (Stage 1)	Affymetrix 6.0	Birdseed	call rate <95%, MAF<1%, pHWE<10E-5	602,642	MACH v1.0.16	HapMap release 21 (build 35)	none	R, probABEL, PLINK
ARIC (Stage 2)	Affymetrix 6.0	Birdseed	MAF <1%, call rate <95%, HWE<10E-5	669450	MACH Imputation (v1.0.16)	HapMap release 22 (build 36)	none	R, probABEL, PLINK
Baltimore Longitudinal Study on Aging	Illumina 550K version 3	Illumina Beadstudio	call rate <99%, MAF <1%, pHWE <10E-4	501,764	MACH version 1.0.15	HapMap release 21 (build 35)	r2hat<0.3	MERLIN
Cardiovascular Health Study	Illumina 370CNV	Illumina BeadStudio	call rate <97%, heterozygotes=0, pHWE<10E-5, SNP not in HapMap	306,655	BimBam	HapMap release 21A (build 36)	dosage variance <0.01	R, robust variance option
CoLaus Study	Affymetrix 500K	BRLMM (Affymetrix)	call rate <70%, pHWE<10E-7	390631	IMPUTE version 0.2.0	HapMap release 21 (build 35)	none	Matlab
EPIC/cohort	Affymetrix 500K	BRLMM	Call Rate<90%, pHWE<10e-6	382,037	IMPUTE version 0.3.1	HapMap release 21 (Build 35)	none	SAS, Stata, Linux scripts
Fenland Study	Affymetrix 500K	BRLMM	call rate <90%, pHWE<10E-6, MAF<1%	362,055	IMPUTE v0.4.2	HapMap release 22 (build 36)	proper_info<0.4	LINUX Stata 10.1 SNPTEST v1.1.5
Framingham Heart Study	Affymetrix 500K Affymetrix 50K supplemental	Affymetrix	call rate <95%, pHWE<10E-6	503526	MACH version 1.0.15	HapMap release 22 (build 36)	none	R, GWAF
GENOA	Affymetrix 6.0	Birdseed-v2	call rate < 95%, pHWE <10E-6	+/-50kb of target snps	MACH version	HapMap release 22	none	PLINK, R

Supplementary Table 1a-Genotyping and Imputation Information in CKDGen Consortium								
					1.0.16	(build 36)		
HABC	Illumina Human1M-Duo	Illumina Beadstudio	call rate < 97%,	914263	MACH version 1.0.16	Hapmap CEU release 22 build 36	none	R
KORA F3	Affymetrix 500K	BRLMM	per chip call rate <93%, MAF <5%, discrepancy for one of the 50 SNPs common on both chips	380407	MACH version 1.0.15	HapMap release 22 (build 35)	none	ProbABEL, R, SAS, Visual Basic
KORA F4	Affymetrix 6.0	BRLMM	per-chip call rate <93%, per-SNP call rate <93%, MAF<1%, gender checks	629,893	MACH version 1.0.16	HapMap release 22 (build 36)	none	ProbABEL, R, SAS, Visual Basic
KORCULA	HAP370CNV	Illumina Beadstudio	call rate <98%, pHWE<10E-10	316730	MACH version 1.0.15	HapMap release 22 (build 36)	none	R GenABEL,ProbABEL
MICROS (South Tyrol)	Illumina, 300K	Illumina Beadstudio	call rate <98%, MAF < 1%, pHWE<10E-6	292,917	MACH v1.0.16	HapMap release 22 (build 36)	none	R, GenABEL, ProbABEL; details in Study-Specific Methods
Sorbs (Slavonic population from Eastern germany)	Affymetrix 500K; Affymetrix 6.0	500K:BRLMM algorithm Affymetrix 6.0:Birdseed	call rate <=94%, pHWE<10E-4	378513	IMPUTE version 1	HapMap release 21 (build 35)	MAF>1%, HWE<10-4	SNPTEST
SPLIT	HAP370CNV	Illumina Beadstudio	call rate <98%, pHWE<10E-10	330997	MACH version 1.0.15	HapMap release 22 (build 36)	none	R(GenABEL, ProABEL)
Study of Health in Pomerania	Affymetrix 6.0	Birdsuite-V2	none	869224	Impute v0.5.0	HapMap release 22 (build 36)	none	SNPTEST v1.1.5, Quicktest 0.94, R, InforSense, InterSystems Caché
MAF = minor allele frequency, pHWE = p value testing for deviation from Hardy-Weinberg equilibrium								

Supplementary Table 1b-Genotyping information from the IBC Array in the CARe Consortium

Study Name	Overall number Genotyped	Total Number of Samples Genotyped that Pass QC	Data management and statistical analysis
ARIC	15103	3029 African American; 9588 Caucasians	PLINK
CARDIA	3490	1295 African American; 1443 Caucasians	PLINK
CHS	5362	752 African American; 3952 Caucasians	PLINK
FHS	8016	7556 Caucasians	FHS R-LME/GEE
JHS	2157	2036 African American	PLINK
MESA	6482	1612 African American; 2298 Caucasian; 637 Asians; 1302 Hispanics	PLINK

Supplementary Table 2 - Lead SNP Association with Renal Traits in CKDGen Stage 1 and Stage 2 follow-up Meta-Analyses											
Trait	**SNP ID**	**Chr**	**position (b36)**	**Genes Within 60 kb or closest gene***	**SNP function**	**Coded Allele Frequency**	**Coded Allele**	**beta coefficent discovery****	**Trait Discovery P-value#**	**Trait Replication P-value$**	**Trait Meta-analysis P-value$**
UACR											
Overall	rs7576149	2	20789943	**C2orf43**, GDF7	intronic	0.1847	t	-0.0594	2.2E-06	7.1E-01	2.1E-04
Overall	rs3734316	6	37705107	MDGA1	intergenic	0.1947	t	0.0543	2.3E-06	1.4E-01	5.2E-06
Overall	rs1801239	10	16959058	**CUBN**, RSU1	nonsynonymous missense	0.9015	t	-0.0835	3.0E-07	2.0E-02	4.0E-08
No DM	rs9372871	6	127849645	**C6orf174**	intronic	0.8907	t	-0.0684	5.7E-06	6.8E-01	4.5E-04
No DM	rs2838520	21	44440324	ICOSLG;C21orf33; DNMT3L	intergenic	0.3861	t	0.0462	1.2E-06	4.3E-02	5.0E-02
MA											
Overall	rs6829211	4	4601533	STX18	intergenic	0.6994	a	-0.1401	5.5E-06	7.5E-01	6.9E-02
Overall	rs17319721	4	77587871	**SHROOM3**	intronic	0.4301	a	-0.1279	1.9E-06	8.4E-01	1.1E-01
Overall	rs2635165	8	98365385	TSPYL5	intergenic	0.9184	t	-0.2343	2.3E-06	6.1E-02	2.3E-04
No DM	rs850898	2	185921053	*ZNF804A, FSIP2*	intergenic	0.2635	a	0.1585	2.2E-06	2.1E-01	6.5E-01
No DM	rs7649382	3	34071517	*PDCD6IP*	intergenic	0.6886	a	-0.159	6.8E-07	6.3E-01	4.0E-02
No DM	rs13104825	4	186237102	HELT	intergenic	0.6569	c	-0.1465	2.4E-06	4.7E-01	1.9E-02
No DM	rs13160548	5	38814607	*OSMR*	intergenic	0.6963	t	-0.1811	1.1E-07	3.9E-01	3.4E-01
No DM	rs13215455	6	52847613	GSTA5;GSTA3	intergenic	0.6624	t	0.16	1.2E-06	1.0E+00	8.6E-02
No DM	rs4935985	11	126086026	**KIRREL3**	intronic	0.7782	a	-0.1683	1.6E-06	4.3E-01	3.9E-01
No DM	rs13337289	16	75895552	**ADAMTS18**	intronic	0.1077	t	0.2441	3.2E-06	9.7E-01	1.9E-01
No DM	rs159782	20	4295347	*ADRA1D*	intergenic	0.6971	a	0.1547	2.4E-06	4.6E-01	3.9E-01

*Bold means the SNP is located within the gene, italics means closest genes if no gene with 60kb of SNP

**Betas for MA from logistic regression and for UACR from general linear regression

#p values from fixed effects model meta analysis

$ p values for MA from fixed effects models; p values for UACR from sample size weighted analyses

Supplementary Table 3- Genome-wide Significant Loci: SNP Imputation Quality in CKDGen Stage 1 Cohorts (bold: genotyped SNPs used for imputation)																
	rs7576149	rs850898	rs7649382	rs6829211	rs17319721	rs13104825	rs13160548	rs3734316	rs13215455	rs9372871	rs2635165	rs1801239	rs4935985	rs13337289	rs159782	rs2838520
AMISH	0.58	0.96	0.89	0.44	0.99	**1.00**	0.81	**1.00**	0.91	0.80	0.07	0.36	0.59	0.46	**1.00**	**1.00**
ARIC	0.92	1.00	1.00	0.85	**1.00**	**1.00**	0.89	**1.00**	0.97	**1.00**	0.97	0.82	**1.00**	0.84	**1.00**	**1.00**
BLSA	**1.00**	0.99	**1.00**	**1.00**	0.99	0.88	0.91	0.62	0.96	0.99	**1.00**	**1.00**	**1.00**	0.90	1.00	1.00
CHS	**1.00**	0.97	**1.00**	**1.00**	0.72	0.72	0.92	0.37	0.68	0.91	**1.00**	0.64	**1.00**	0.49	0.98	0.97
COLAUS	0.76	0.99	0.97	0.66	1.00	**1.00**	0.89	**1.00**	0.97	0.97	0.74	0.87	0.93	0.83	**1.00**	**1.00**
EPIC	0.62	0.99	0.96	0.65	1.00	**1.00**	0.88	**1.00**	0.97	0.86	0.75	0.85	0.91	0.78	**1.00**	**1.00**
FENLAND	0.64	0.99	0.96	0.73	1.00	**1.00**	0.88	**1.00**	0.95	0.90	0.75	0.88	0.93	0.79	**1.00**	**1.00**
FHS	0.92	0.98	0.96	0.79	1.01	**1.00**	0.89	**1.00**	0.88	1.01	0.71	**1.00**	0.89	0.84	**1.00**	1.02
KORAF3	0.69	0.97	0.99	0.63	0.96	**1.00**	0.90	**1.00**	0.92	0.96	0.69	0.58	0.91	0.74	**1.00**	**1.00**
KORAF4	0.89	0.98	0.95	0.83	**1.00**	**1.00**	0.84	**1.00**	0.96	**1.00**	0.94	0.92	**1.00**	0.77	**1.00**	**1.00**
MICROS	**1.00**	0.99	0.94	**1.00**	0.97	0.88	0.91	0.64	0.82	1.00	**1.00**	0.87	**1.00**	0.92	1.00	0.98
SHIP	0.94	0.99	0.97	0.86	**1.00**	**1.00**	0.87	**1.00**	0.98	**1.00**	0.96	0.95	**1.00**	0.81	**1.00**	**1.00**

Supplementary Table 4 - Meta-analysis results of lead SNPs across traits in CKDGen Stage 1[#], CKDGen Stage 2, and CKDGen Stage 1 and 2 combined meta-analysis[$]

					UACR Overall			**UACR No Diabetes**			**Microalbuminuria Overall**			**Microalbuminuria No Diabetes**		
SNPID	Chr	Position	Coded allele	**Genes Within 60 kb or closest gene***	**Stage 1 P-value**	**Stage 2 P-value**	**Combined P-value**	**Stage 1 P-value**	**Stage 2 P-value**	**Combined P-value**	**Stage 1 P-value**	**Stage 2 P-value**	**Combined P-value**	**Stage 1 P-value**	**Stage 2 P-value**	**Combined P-value**
rs7576149	2	20789943	t	C2orf43, GDF7	**2.2E-06**	7.1E-01	2.1E-04	4.0E-06	5.4E-01	1.6E-04	6.5E-02	2.9E-01	9.7E-02	1.2E-01	4.4E-01	2.2E-01
rs850898	2	185921053	a	*ZNF804A, FSIP2*	4.0E-02	9.0E-01	1.3E-01	1.6E-02	9.0E-01	1.0E-01	1.7E-04	4.1E-01	4.7E-01	**2.2E-06**	2.1E-01	6.5E-01
rs7649382	3	34071517	a	*PDCD6IP*	9.5E-02	6.2E-01	4.0E-01	4.6E-03	3.7E-01	1.6E-01	3.6E-05	4.4E-01	2.5E-02	**6.8E-07**	6.3E-01	4.0E-02
rs6829211	4	4601533	a	STX18	1.4E-02	6.1E-01	2.6E-02	1.4E-01	9.7E-01	2.0E-01	**5.5E-06**	7.5E-01	6.9E-02	1.4E-04	8.4E-01	2.0E-01
rs17319721	4	77587871	a	**SHROOM3**	1.7E-04	3.9E-03	7.0E-07	1.4E-04	4.1E-03	6.8E-07	**1.9E-06**	8.4E-01	1.1E-01	1.2E-05	7.9E-01	2.4E-01
rs13104825	4	186237102	c	HELT	1.2E-03	1.6E-01	1.4E-01	2.4E-04	1.7E-01	6.8E-02	1.4E-04	9.7E-01	1.3E-01	**2.4E-06**	4.7E-01	1.9E-02
rs13160548	5	38814607	t	*OSMR*	3.7E-03	8.3E-01	2.1E-02	1.2E-03	8.0E-01	1.3E-02	3.2E-07	2.0E-01	4.5E-01	**1.1E-07**	3.9E-01	3.4E-01
rs3734316	6	37705107	t	MDGA1	**2.3E-06**	1.4E-01	5.2E-06	1.9E-05	1.5E-01	1.5E-05	1.7E-04	8.3E-02	1.5E-03	1.3E-03	7.7E-02	4.4E-03
rs13215455	6	52847613	t	GSTA5;GSTA3	2.2E-03	2.8E-01	1.6E-01	1.0E-04	4.2E-01	3.3E-02	2.5E-04	8.1E-01	8.8E-02	**1.2E-06**	1.0E+00	8.6E-02
rs9372871	6	127849645	t	**C6orf174**	9.2E-03	5.1E-01	2.0E-02	**5.7E-06**	6.8E-01	4.5E-04	9.4E-03	1.5E-01	2.3E-02	5.3E-05	5.7E-01	6.8E-02
rs2635165	8	98365385	t	TSPYL5	2.6E-03	6.7E-01	3.7E-03	9.3E-03	7.9E-01	1.5E-02	**2.3E-06**	6.1E-02	2.3E-04	2.0E-05	2.6E-01	8.8E-03
rs1801239	10	16959058	t	**CUBN**, RSU1	**3.0E-07**	2.0E-02	4.0E-08	3.6E-04	2.4E-02	1.7E-05	8.7E-07	4.3E-01	1.0E-02	6.8E-04	7.9E-01	1.8E-01
rs4935985	11	126086026	a	**KIRREL3**	4.6E-02	9.4E-01	1.1E-01	6.8E-03	9.2E-01	3.7E-02	4.1E-04	7.9E-01	2.8E-01	**1.6E-06**	4.3E-01	3.9E-01
rs13337289	16	75895552	t	**ADAMTS18**	1.2E-04	2.1E-02	2.1E-01	5.9E-05	4.4E-04	6.0E-01	1.5E-05	4.5E-01	3.7E-02	**3.2E-06**	9.7E-01	1.9E-01
rs159782	20	4295347	a	*ADRA1D*	7.0E-02	6.7E-02	1.1E-02	1.5E-02	2.9E-02	1.5E-03	8.2E-06	7.1E-01	1.8E-01	**2.4E-06**	4.6E-01	3.9E-01
rs2838520	21	44440324	t	ICOSLG;C21orf33; DNMT3L	3.8E-06	1.6E-01	2.2E-02	**1.2E-06**	4.3E-02	5.0E-02	2.6E-02	2.6E-01	8.4E-01	2.9E-02	3.3E-01	8.4E-01

bold indicates the trait and stratum with the lowest p-value in Stage 1

*__Bold__ means the SNP is located within the gene, *italics* means closest genes if no gene with 60kb of SNP

$ p values from inverse variance weighted fixed effects model for microalbuminuria and all discovery analyses; p values from a sample size weighted Z score method for UACR Stage 2 and UACR combined

Statistical Methods

An additive genetic model was used in all study-specific analyses. In single center studies, we adjusted for age and sex, whereas in multicenter studies further adjustment for study-center was performed to account for possible minor differences between recruiting centers.

CKDGen and CARe Study-specific Genetic Association Analysis For cohorts consisting of only unrelated individuals, we performed linear regression for the quantitative trait UACR and logistic regression for the dichotomous trait microalbuminuria. For family-based studies, we applied linear mixed effect (LME) models for UACR and logistic regression via generalized estimating equations (GEE) for MA to account for the familial relatedness.

Due to different ethnic groups involved in CARe, principal components within each ethnicity in each study were created for use as covariates to adjust for population stratification, and used as additional covariates in regression analyses.

Meta-analysis For the meta-analyses of UACR among CKDGen Stage 1 discovery or CARe discovery studies and for all meta-analyses with MA, we used the inverse-variance weighted fixed effect model for combining beta estimates. Due to different scaling of UACR in CKDGen Stage 2 studies, we used the weighted Z-score method for the meta-analysis of these association results and for the joint meta-analysis of CKDGen Stage 1 and Stage 2 studies combined with the CARe studies. The weighted Z-score method combines the p-values weighted by study sample size and is therefore independent of the scaling of the quantitative parameter[42] (METAL software, http://www.sph.umich.edu/csg/abecasis/Metal/index.html). Genomic control correction to guard against potential underlying population stratification was applied at the study-specific level and, again, for the meta-analysis results (see lambda values in Supplementary Figure 2).[43] For quality control, study specific files were checked for plausible ranges of values and all analyses were performed in duplicate by two independent analysts. Only SNPs with minor allele frequency (MAF) >1% were considered for further analysis.

DCCT/EDIC For the time to event outcome (see outcome section in Methods), two Cox proportional hazards models were run: 1) a simple model that included the following covariates: cohort status (primary vs. secondary), treatment (intensive vs. conventional) and cohort*treatment interaction (stratified by DCCT year of entry); 2) an extended model that included the covariates in the first model plus additional covariates measured at DCCT baseline: age of diagnosis squared, sex, diabetes duration squared, body mass index, blood pressure, triglyceride, HDL-C, total cholesterol, baseline smoking, as well as time-dependent updated mean hemoglobin A1C, and time-dependent indicators for hypertension diagnosis and treatment. The SNP effect was tested using a likelihood ratio test to compare models with and without the SNP included.

Proportion of variance explained by locus The proportion of variance explained by a replicated locus was calculated using the formula "explained variance = 2*MAF*(1-MAF)*((beta/SD)^2)", where beta is from the CARe meta-analysis, and the SD was obtained from the large ARIC cohort, separately for samples of European descent and for African Americans.

Statistical software In CKDGen and CARe, statistical software used included Mixed Model Analysis for Pedigrees (MAPP) software, developed by Dr. J. R. O'Connell at University of Maryland, Baltimore, the R based GenABEL and ProbABEL[44] http://mga.bionet.nsc.ru/~yurii/ABEL/) software, SAS statistical package (version 9.1 for UNIX; SAS Institute, Cary, NC), MERLIN,[45] Mach2 QTL (http://www.sph.umich.edu/csg/abecasis/MACH/), QUICKTEST v0.94 (http://toby.freeshell.org/software/quicktest.shtml), GWAF (http://cran.r-project.org/web/packages/GWAF/)[46] and SNPTEST v1.1.5 (http://www.stats.ox.ac.uk/~marchini/software/gwas/snptest.html). Imputation software included MACH (http://www.sph.umich.edu/csg/abecasis/MACH/), BIMBAM, or IMPUTE.

Study-Specific Methods

CKDGen Stage 1 Cohorts

Amish The Old Order Amish individuals included in this study were participants of several ongoing studies of cardiovascular health carried out at the University of Maryland. Participants were relatively healthy volunteers from the Old Order Amish community of Lancaster County, Pennsylvania and their family members,[47,48] Examinations were conducted at the Amish Research Clinic in Strasburg, PA. Study participants were enrolled within the 2000-2008 time period. Using urine samples stored at – 80℃, and urinary albumin concentration was measured with a quantitative immunoturbimetric assay (Roche Diagnostics, Indianapolis, Indiana). Urinary creatinine was measured using a modified Jaffe method.

ARIC The Atherosclerosis Risk in Communities (ARIC) Study started in 1987-89, when 15,792 mostly African American and white participants were enrolled from four US communities and attended a baseline visit.[49] The participants returned on average every three years for three additional study visits. For the current study in the CKDGen Consortium, participants were included if they were of Caucasian origin, successfully genotyped on the Affymetrix 6.0 array and had measurements of the imputed SNP rs1801239 as well as urinary albumin and creatinine available. There was no association between UACR and the first 10 principal components obtained using Eigenstrat;[50] therefore, only adjustment for genomic control was used to account for potential population stratification. For the current study in the CARe Consortium, 7687 self-reported white participants were genotyped on the IBC chip and data on rs1801239 genotypes and UACR available. In the self-reported African American participants, 2010 were genotyped on the IBC chip and had data on rs1801239 genotypes and UACR available, and 1910 were genotype on the Affy 6.0 array and had imputed data on rs1801239 as well as the phenotype available. Urine samples were obtained at the fourth study visit (1996-98), frozen within 12 hours and stored at -70 ˚C. Urinary albumin was measured by a

nephelometric method either on the Dade Behring BN100 or on the Beckman Image Nephelometer, and urinary creatinine was measured using the Jaffe method.[51] The correlation coefficient between duplicate measurements of ln(UACR) on 516 samples was 0.95. The covariates used in the analyses were ascertained at the time of urine collection. Diabetes was defined as fasting glucose of ≥126 mg/dl (7.0 mmol/L), non-fasting glucose of ≥200 mg/dl (11.1 mmol/L), self-reported physician-diagnosis of diabetes, or intake of diabetes medication.

Baltimore Longitudinal Study on Aging The Baltimore longitudinal study on Aging (BLSA) study is a population-based study aimed to evaluate contributors of healthy aging in the older population residing predominantly in the Baltimore-Washington DC area.[52] Starting in 1958, participants are examined every one to four years depending on their age. Currently there are approximately 1100 active participants enrolled in the study. Blood samples were collected for DNA extraction, and genome-wide genotyping was completed for 1231 subjects using Illumina 550K.[53] The analysis was restricted to subjects with European ancestry with urinary albumin excretion data (N=354) and each analysis was adjusted for the top two principal components derived from an EIGENSTRAT analysis utilizing ~10,000 randomly selected SNPs from the 550K SNP panel.[50] Urinary albumin and creatinine was measured in an aliquot of a 24-hour urine sample collection. Urinary albumin excretion was measured with nephelometry (Beckman Array System). Urinary creatinine was measured using a Vitros enzymatic assay (Johnson & Johnson Co., Rochester, NY).

Cardiovascular Health Study The Cardiovascular Health Study (CHS) is a community-based longitudinal study of risk factors for cardiovascular disease and stroke in adults 65 years of age or older, recruited at four field centers (Forsyth County, NC; Sacramento County, CA; Washington County, MD; Pittsburgh, PA).[54] 5201 predominantly Caucasian individuals were recruited in 1989-1990 from random samples of Medicare eligibility lists, followed by an

additional 687 African-Americans recruited in 1992-1993 (total n=5888). A total of 1908 persons were excluded from the GWAS study sample due to the presence at study baseline of coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke or transient ischemic attack or lack of available DNA. African American participants were excluded from the CKDGen component of analysis since the other cohorts were predominantly Caucasian. Of remaining participants, 1865 underwent measurement of urine albumin and creatinine at the 1996-1997 CHS study visit and were included in this study. Urine was collected as a single morning void. Urine albumin was measured by rate nephelometry using the Array 360 CE Protein Analyzer (Beckman Instruments, Fullerton, CA). Urine creatinine was measured using a Kodak Ektachem 700 Analyzer (Eastman Kodak company, Rochester, NY).

CoLaus The Cohorte Lausannoise (CoLaus) study is a population-based study aimed at assessing the prevalence and molecular determinants of cardiovascular risk factors in the population of Lausanne, Switzerland.[55] Study participants (5,311) were randomly selected from the population register of Lausanne in 2003 (n=56,694, aged 35-75 years). All individuals were of Caucasian origin, defined as having both parents and grandparents born in a defined list of European countries. Baseline examination occurred in 2003-2006. Urinary albumin concentration was measured with a Bromocresol green assay (inter- and intra-assay CV of 2.5% and 0.4%, Roche Diagnostics, Basel, Switzerland). Urinary creatinine concentration was measured using a Jaffe kinetic compensated method, (inter- and intra-assay CV of 2.9% and 0.7%, respectively).

EPIC The EPIC Norfolk GWA cohort includes 2,566 participants randomly selected from the EPIC-Norfolk Study, a population-based cohort study of 25,663 men and women of European descent aged 39-79 years recruited in Norfolk, UK between 1993 and 1997.[56] All participants attended a health check during which urine and blood samples were taken and health and

lifestyle questionnaires were completed. The measurement of urinary albumin and creatinine concentrations has been described in detail before.[57] A random spot urine sample was obtained during the health check. Urinary albumin concentration was measured by means of immunonephelometry using the Nephelometer II analyzer (Dade Behring, Marburg, Germany) (intra-assay CV [n=10] 2.91%). Urinary creatinine was measured by means of colorimetry using the Dimension AR Analyser (Dade Behring Marburg, Germany).

Fenland The Fenland Study is an ongoing population-based cohort study (started in 2005) designed to investigate the association between genetic and lifestyle environmental factors and the risk of obesity, insulin sensitivity, hyperglycaemia and related metabolic traits in men and women aged 30 to 55 yrs. Potential volunteers were recruited from General Practice sampling frames in the Fenland, Ely and Cambridge areas of the Cambridgeshire Primary Care Trust in the U.K. Exclusion criteria for the study were: prevalent diabetes, pregnant and lactating women, inability to participate including terminal illness, psychotic illness, or inability to walk unaided. Currently, the study comprises more than 3,000 participants of whom the first 1,500 volunteers with complete anthropometric data were genotyped and included in the current analyses. All participants were measured at the MRC Epidemiology Unit Clinical Research Facilities in Ely, Wisbech and Cambridge. Participants attended after an overnight fast for a detailed clinical examination and blood samples were collected. Using stored urine samples, urinary albumin concentration was measured by means of immunonephelometry using the Nephelometer II analyzer (Dade Behring, Marburg, Germany) (intra-assay CV [n=10]: 2.91%). Urinary creatinine was measured by means of colorimetry using the Dimension AR Analyser (Dade Behring Marburg, Germany).

Framingham Heart Study In 1948, the Framingham Heart Study was initiated with the enrollment of the Original Cohort.[58] In 1971, enrollment of the Offspring Cohort was started

(5,124 participants); the design and methodology has been described.[59,60]. In 2002, enrollment began for the Third Generation cohort (n=4095).[61] For the current study, participants are derived from the offspring cohort who attended the sixth examination (1995 - 1998), and individuals from the Third Generation first examination. We observed no association with UACR or MA with the 10 principal components estimated using Eigenstrat;[50] therefore, we only used genomic control adjustment to account for potential population stratification. Using stored urine samples, urinary albumin concentration was measured with a Tina-quant immunoturbimetric assay (intra-assay CV 7.2% for the Offspring cohort, Third Generation CV= 2.1%; Roche Diagnostics, Indianapolis, Indiana). Using a modified Jaffe method, urinary creatinine concentration was measured (intra-assay CV=2.3% for the Offspring cohort and 1.0% for the Third Generation cohort).

KORA F3 and F4 The KORA surveys for genetic research have been described in detail previously[62,63] and have been initiated as part of the MONICA (Monitoring of Trends of Cardiovascular Diseases) multi-center study. The third KORA survey (KORA S3) is a population-based sample from the general population of the South-German city of Augsburg and surrounding counties, recruited 1994/1995. A subsample consisting of 1530 individuals from this survey with 10-year follow-up (KORA F3) information and urine available was successfully genotyped. The fourth KORA survey (KORA S4) is a sample recruited 1999-2001 independent from KORA S3 using the same platform with the same standard operating procedures and based on the same population. From the sample with a 7-year follow-up (KORA F4), 1803 subjects with complete urine information were available for the GWA analysis.All participants had a German passport and were of European origin. Using stored urine samples, urinary albumin concentration in KORA F3 and KORA F4 was measured with a latex enhanced nephelometric assay (Siemens Healthcare Diagnostics) on a Dade Behring BN2 apparatus.

Urinary creatinine concentration was measured using an enzymatic method in KORA F3 and a kinetic Jaffe method in KORA F4.

MICROS The *MICROS* study is part of the genomic health care program 'GenNova' and was carried out in three villages of the Val Venosta, South Tyrol (Italy), in 2001-2003. It comprised members of the populations of Stelvio, Vallelunga and Martello. A detailed description of the *MICROS* study is available elsewhere[64]. Briefly, study participants were volunteers from three isolated villages located in the Italian Alps, in a German-speaking region bordering with Austria and Switzerland. Owing to geographical, historical and political reasons, the entire region experienced a prolonged period of isolation from surrounding populations. Information on the participant's health status was collected through a standardized questionnaire. Laboratory data were obtained from standard blood analyses. Urinary albumin-to-creatinine ratio (ACR) was measured on a point-of-care diabetes management platform (Bayer DCA 2000+ analyzer). The limits of detection (LODs) of the assay for albumin and creatinine assessment were of 5.0 mg/l and 12.0 mg/dl, respectively, with a consequent, wide range of left-censored values when calculating the ratio. In this situation, the assessment of microalbuminuria based on the thresholds defined in the Methods section was straightforward. However, a consistent analysis of individual ACR values was difficult to achieve and, given the very small sample size of the study, was not performed. Covariates were obtained during the interview phase. Diabetes was defined as fasting serum glucose of at least 126 mg/dl (7.0 mmol/L) or self reported diabetes.

The Study of Health in Pomerania The Study of Health in Pomerania (SHIP) is a longitudinal population-based cohort study conducted in West Pomerania, the north-east area of Germany. For the baseline examinations, a sample of 6267 eligible subjects aged 20 to 79 years was drawn from population registries. Only individuals with German citizenship and main residency in the study area were included. Selected persons received a maximum of three written

invitations. In case of non-response, letters were followed by a phone call or by home visits if contact by phone was not possible. The SHIP population finally comprised 4310 participants (response 68.8%). Baseline examinations were conducted between 1997 and 2001. Between 2002 and 2006 all participants were re-invited for an examination follow-up, in which 3300 subjects (83.5% of eligible persons) took part.[65] A blood sample was drawn from the cubital vein in the supine position (the participants were non-fasting due to the duration of the cumulative examinations, 4-6 hours in total). A urine sample was taken from spontaneous urine. The urinary albumin concentration was determined on a Behring Nephelometer (Siemens BN albumin; Siemens Healthcare, Marburg, Germany). The creatinine concentration in urine was determined using the Jaffe method on a Hitachi 717 (intra-assay CV 2.2%; Roche Diagnostics, Germany).

CARe Consortium

The CARe consortium (http://www.broadinstitute.org/gen_analysis/care/index.php/Main_Page) is a consortium of 9 studies (ARIC, CARDIA, CHS, Cleveland Family Study (CFS), Cooperative Study of Sickle Cell Disease (CSSCD), Framingham Heart Study, Jackson Heart Study, MESA, the Sleep Heart Health Study) funded by the National Heart, Lung, and Blood Institutes of the NHLBI. CARe was instituted in 2006 in order to study candidate genes for cardiovascular disease and its risk factors, and to perform genome-wide association in African Americans. Methods for ARIC, CHS, and FHS can be found above.

CARDIA: Study design details of Coronary Artery Risk Development In Young Adults (CARDIA) have been previously published.[66] Briefly, CARDIA recruited a cohort of young black and white adults, age 18-30 at the time of enrollment. From1985-1986 CARDIA recruited participants from four sites: Birmingham, AL, Chicago, IL, Minneapolis, MN, and Oakland, CA.

Follow-up examinations occurred at Years 2, 5, 7, 10, 15, and 20 years. In the present analyses, we included whites and African Americans with measured urinary albumin and creatinine at the year 15 exam. All study protocols were approved by the appropriate institutional review boards. Albuminuria was measured from a spot sample. Albumin was assessed by nephelometry and creatinine was assessed using the Jaffe method. Urinary albumin and creatinine from Year 15 were reported as albumin to creatinine ratio in mg/g. The CV for urinary albumin was 2.9%.

Jackson Heart Study (JHS) The Jackson Heart Study is a single site longitudinal population-based study. The sample consists of 5,302 African-American women and men selected between 2000 and 2004 (first visit) from the tri-county area encompassing Jackson MS.[67,68] DNA and consent for data sharing consistent with NIH guidelines were available for 3,443 participants for the CARe consortium. Using stored urine samples, urinary albumin concentration was measured using the human albumin kit (Dade, Behring,Newark, DE) on the Dade Behring BN II nephelometer. Biochemical testing for urine creatinine was performed at the University of Mississippi Medical Center Laboratory Reading Center by using a multipoint enzymatic spectrophotometric assay (Vitros CREA dry reaction slides on a Vitros 950 Otho-Clinical Diagnositics analyzer, Raritan, NJ). CVs were 1.9-2.1% for urinary creatinine and 4.2-5.4% for urinary albumin.

MESA The Multi-Ethnic Study of Atherosclerosis (MESA) is a community-based cohort study designed to investigate the prevalence, correlates, and progression of subclinical cardiovascular disease.[69] The MESA cohort is comprised of 6,814 adults of diverse race/ethnicity (38% Caucasian, 28% African American, 22% Hispanic, and 12% Chinese) from six U.S. regions: Forsyth County, North Carolina, Northern Manhattan and the Bronx, New York, Baltimore City

and Baltimore County, Maryland, St. Paul, Minnesota, Chicago, Illinois, and Los Angeles County, California. Participants ranged between 45 and 84 years of age and were free of clinical cardiovascular disease at the baseline examination, which was conducted from 2000 through 2002. All data used for the analyses reported herein were collected at this baseline examination. Urine was collected from single voided specimens. Urine albumin concentration was determined by nephelometry using the Array 360 CE Protein Analyzer (Beckman Instruments, Inc., Drea, CA). Urine creatinine was measured using the Vitros 950IRC instrument (Johnson & Johnson Clinical Diagnostics, Inc., Rochester, NY). Urinary creatinine was measured using the Vitros 950IRC instrument (Johnson & Johnson Clinical Diagnostics, Inc., Rochester, NY) at the Clinical Chemistry Laboratory at Fletcher Allen Health Care (Burlington, VT). Thin film technology is used to quantitatively measure creatinine via a colorimetric reaction. The CV range is 2.5 – 2.9%. Urinary albumin was determined using the Array 360 CE Protein Analyzer (Beckman Instruments, Inc., Drea, CA) at the Clinical Chemistry Laboratory at Fletcher Allen Health Care (Burlington, VT). This system utilizes a nephelometer to measure the rate of light scatter formation resulting from an immunoprecipitation reaction.

CKDGen Stage 2 In Silico Cohorts

The AGES Study The Reykjavik Study cohort originally comprised a random sample of 30,795 men and women born in 1907-1935 and living in Reykjavik in 1967. A total of 19,381 people attended, resulting in 71% recruitment rate. The study sample was divided into six groups by birth year and birth date within month. One group was designated for longitudinal follow-up and was examined in all stages. One group was designated a control group and was not included in examinations until 1991. Other groups were invited to participate in specific stages of the study. Between 2002 and 2006, the AGES-Reykjavik study re-examined 5,764 survivors of the original cohort who had participated before in the Reykjavik Study.[70] At the baseline visit, participants

were asked to bring in morning urine samples and albumin was measured on a fresh sample by Tina-quant® immunoturbimetric assay (intra-assay CV 7.2%); Roche Diagnostics, Mannheim. In these same samples Urine creatinine was measured using HiCo Creatinine Jaffe method (intra-assay CV 4.2%) Roche Diagnostics Mannheim.

ARIC Because additional genotype data became available during the course of this study, an additional 759 self-reported white ARIC participants were included in the Stage 2 CKDGen analyses. These individuals were not part of the discovery samples, did not have a first-degree relationship with any individual in the discovery sample, nor would they have been classified as an outlier based on allele sharing measures applied in data cleaning of the discovery sample. Albuminuria and covariate assessment and statistical methods are described above.

DCCT/EDIC The Diabetes Control and Complications Trial (DCCT) was a clinical trial to compare conventional as compared to intensive insulin treatment with respect to the development and progression of complications related to type 1 diabetes. After closeout of DCCT, the majority of subjects have been followed-up in the Epidemiology of Diabetes Interventions and Complications (EDIC) study. The current study consists of 1,304 white participants who underwent genotyping on the Illumina 1M SNP chip. Renal function was assessed at yearly visits in the DCCT and on alternate years (based on randomization in the DCCT) during EDIC. The assessment included measurement of urinary albumin excretion based on 4 hour timed urine collections. Antihypertensive medications including ACE inhibitors or angiotensin II receptor blockers were not discontinued. Urine albumin was measured by fluoroimmunoassay[71] with coefficients of variation and coefficients of reliability of 14% and 95% respectively.[72] Renal outcomes were as follows; time from DCCT baseline until: 1) Persistent microalbuminuria, defined as time to two consecutive albumin excretion rates (AERs) >30 mg/24 h (>20.8 µg/min); 2) severe nephropathy, defined as the time to AER >300 mg/24 h

(>208 µg/min) with prior persistent microalbuminuria or end-stage renal disease. For each outcome, two Cox Proportional hazards models were run: 1) a simple model that included the following covariates: cohort status (primary vs. secondary, treatment (intensive vs. conventional) and cohort*treatment interaction (stratified by DCCT year of entry) 2) an extended model that included the covariates in the first model plus additional covariates measured at DCCT baseline: age of diagnosis squared, sex, diabetes duration squared, body mass index, blood pressure, triglyceride, HDL-C, total cholesterol, baseline smoking, as well as time-dependent updated mean A1C, and time-dependent indicators for hypertension diagnosis and treatment. The SNP effect was tested using an additive model using a likelihood ratio test between models with and without the SNP included.

GENOA The Family Blood Pressure Program (FBPP), established by the National Heart Lung and Blood Institute in 1996, joined existing research networks that were investigating hypertension and cardiovascular diseases (http://public.nhlbi.nih.gov/GeneticsGenomics/home/fbpp.aspx). One of the four FBPP networks is the Genetic Epidemiology Network of Arteriopathy (GENOA), which recruited hypertensive, Caucasian sibships for linkage and association studies to investigate genetic contributions to hypertension and hypertension-related target organ damage.[73] Sibships containing at least two individuals with clinically-diagnosed essential hypertension before age 60 years were recruited from Rochester, Minnesota. After identifying each hypertensive sibship, all members of the sibship were invited to participate regardless of their hypertension status. Using stored urine samples, urine albumin and creatinine concentrations were measured by standard method on a Hitachi 911 Chemistry Analyzer (Roche Diagnostics, Indianapolis, IN) for GENOA participants.

Health ABC The Health ABC study is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in

older adults. Health ABC enrolled well-functioning, community-dwelling black (n=1281) and white (n=1794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. Genotyping was performed by the Center for Inherited Disease Research (CIDR) using the Illumina Human1M-Duo BeadChip system. Stored urine was assayed by Synarc Inc., Paris, France. Urinary albumin concentration was measured with a nephelometric method using a Nephelometer-Analyzer (intra-assay CV 17.8%; BEHRING BNA, SIEMENS). Urinary creatinine concentration was measured using a modified Jaffe method (Colorimetric method) on a KONE 20 analyzer (intra-assay CV 15.4%; KONELAB).

KORCULA The Korcula Study is a family-based, cross-sectional study on the Dalmatian island of Korcula.[74] Fasting urine and blood samples and were collected from 969 healthy volunteers aged 18 and over from the villages of Lumbarda, Žrnovo, and Račišće on the Island of Korcula, Croatia in 2007. Over 200 health-related phenotypes and environmental exposures were measured in each individual. Urinary albumin excretion was measured, in stored urine samples, by an utomated assay based on a turbimetric method with automatic calibration and quality control (Synchron CX System, Beckman Coulter).The analytical range is between 0.2 and 30 mg/dL, and from 24 to 97mg/dL in the Manual ORDAC mode.

SORBS All subjects are part of a sample from an extensively phenotyped self-contained population from Eastern Germany, the Sorbs.[75] The Sorbs are of Slavonic origin, and lived in ethnic isolation among the Germanic majority during the past 1100 years. Today, the Sorbian-speaking, Catholic minority comprises approximately 15,000 full-blooded Sorbs resident in about 10 villages in rural Upper Lusatia (Oberlausitz), Eastern Saxony. At present, about 1000

Sorbian individuals are enrolled in the study. Sampling comprised unrelated subjects as well as families. Mean IBD sharing in the pairwise comparison was 0.008, median < 10^{-6} (25% percentile <10^{-6}, 75% percentile: 0.012). Extensive phenotyping included standardised questionnaires for past medical history and family history, collection of anthropometric data and a 75g-Glucose-tolerance-test. 877 subjects were available for the present study. Urinary creatinine was measured using an enzymatic method (Roche Inc), urinary albumin was measured by immunological turbidity test (Roche Inc). Genomic control adjustment was used to account for potential population stratification and cryptic relatedness.

SPLIT The Split Study is an ongoing population-based study in the city of Split in Croatia. Fasting urine and blood samples were collected from 535 healthy volunteers aged 18 and over from Split on the Dalmation coast in Croatia in 2008/2009. Over 200 health-related phenotypes and environmental exposures were measured in each individual. Urinary albumin excretion was measured, in stored urine samples, by an automated assay based on a turbimetric method with automatic calibration and quality control (Synchron CX System, Beckman Coulter).

CKDGen Stage 2 De novo Genotyped Cohorts

De novo genotyping was performed in 8 cohorts, either on a MassARRAY system using Assay Design v.3.1.2 and the iPLEX™ chemistry (Sequenom, San Diego, USA) at the Helmholtz Zentrum in Munich, Germany (KORA F3, KORA F4, SAPHIR); or by using 5' nuclease allelic discrimination assays on 7900HT Fast Real-Time Taqman PCR or OpenArray SNP genotyping Systems (Applied Biosystems, Foster City, CA, USA) at the Innsbruck Medical University (KORA F3, KORA F4, SAPHIR), the University of Maryland (AMISH), the St Olav University Hospital of Trondheim (HUNT-2), the Brigham and Women's Hospital (NHS), and KBiosciences, UK (PREVEND); or as part of a larger panel of SNPs at the SNP technology platform at Uppsala University using the Golden Gate assay from Illumina Inc, USA (ULSAM).

For the obtained duplicate genotypes (2-18% of the subjects in each study), concordance was 99.7-100%. Call rates ranged from 94-100% (mean 98.6%) across all studies and SNPs (mean 99.2% for rs1801239 across all studies). The distribution of genotypes in all studies did not deviate from Hardy-Weinberg-Equilibrium ($p>0.05$), except for rs13104825 in the AMISH sub-study ($p=0.0007$).

Amish: The participants from the Amish who were not part of the genome-wide association analyses served as replication samples. Methods for albuminuria assessment, covariates, and the statistical analysis are described above.

KORA F3/F4: The subjects of KORA F3 and KORA F4 with available spot urine albumin und creatinine data who had not been genotyped genome wide as described under “discovery cohorts” were used as replication samples including 1392 and 1197 subjects from KORA F3 and F4, respectively. Covariates were obtained at the index examination.

HUNT 2: The HUNT 2 study is a Norwegian large-scale general health study. From 1995-1997, all individuals residing in Nord-Trøndelag county aged 20-years were invited. The population is stable (net out migration of 0.3% per year) and ethnically homogenous (97 % Caucasians).The objectives, methods and participation in the HUNT 2 Study are described in detail elsewhere.[76] All participants gave an informed consent before the examination, and the study was approved by the Regional Committee of Ethics in Medical Research. Of 92 939 subjects invited, 27 350 did not respond. Thus, 70.6% of the entire adult population participated. Blood was obtained from all participants and frozen for later DNA extraction. A random 5% sample was also asked to deliver 3 urine samples; 75.6% returned all requested urine samples. Urine-albumin concentration in refrigerated urine samples was measured within 5-days using an immunoturbidimetric method (Dako A/S, Denmark; lower detection level 1mg/L). We defined

diabetes as non-fasting plasma glucose of at least 200 mg/dl (11mmol/L) or treatment with a hypoglycemic agent.

Nurses Health Study (NHS)- NHS I began in 1976, when 121,700 female nurses aged 30-55 years completed a detailed questionnaire pertaining to health-related information such as illnesses, medications, and lifestyle. NHS II started in 1989, when 116,430 female nurses aged 25-42 completed a similar questionnaire. Since the start of the studies, these women have completed questionnaires to update health-related information every two years and detailed dietary questionnaires every four years.

The NHS participants in this analysis were part of a study of analgesic use and renal function, since we have urinary albumin and creatinine measured for this cohort.[77] This sub-study was approved by the Brigham and Women's Hospital Institutional Review Board (IRB), and the use of this cohort for the current study was also approved by the IRB. This sub-cohort included women who provided an initial blood sample in 1989 for NHS I (N=32,826) or a blood and urine sample in 1997 for NHS II (N=29,616) and returned a supplementary questionnaire about analgesic use in 1999 for NHS I (N=3876) and 1998 for NHS II (N=4024). We wanted to mail a supplementary questionnaire to a subset of these women who were likely to have a high lifetime intake of analgesics and women who were likely to have low intake. Therefore, we oversampled women who had reported high frequency of analgesic use ($\geq$15 days per month) on biennial questionnaires, and also women who reported no analgesic use on biennial questionnaires. NHS I participants included in this study also submitted blood and urine samples in 2000 (N=3123). This was the first urine sample for this cohort. Due to financial constraints, 2712 women from NHS I and 1643 women from NHS II were selected for these analyses, with oversampling of those with the highest levels of lifetime analgesic consumption but including women of all levels of lifetime intake including low levels. Women with a history of

cardiovascular disease or a history of cancer (except for non-melanoma skin cancer) in 1989 for NHS I and 1997 for NHS II were excluded from the initial blood collection. However, women who developed cardiovascular disease or cancer after these dates were not excluded. Urine albumin was measured by immunoassay using the Hitachi 911 analyzer and Roche Diagnostics reagents (Indianapolis, IN). Using blinded quality control samples, the coefficient of variation for this assay was 8%. Urine creatinine was measured using a modified Jaffe method (coefficient of variation 2%).[78] Hypertension and diabetes were self-reported by mailed questionnaire.

PREVEND- The Prevention of REnal and Vascular ENd stage Disease (PREVEND) study consists of residents ages 28 to 75 years (n=85,421) of Groningen, The Netherlands, who were invited to complete a questionnaire. Overall, nearly half (47%) responded; individuals were selected to participate if they had a urinary albumin concentration ≥10 mg/L (n = 7,768); a control group was randomly selected if participants had a urinary albumin concentration < 10 mg/L (n = 3,395). The urinary albumin excretion (UAE) was measured as the average of two 24-hour urine collections and was classified according to clinical classes. Diabetes was defined as a fasting glucose level ≥126 mg/dL, nonfasting plasma glucose level ≥ 200 mg/dL, or the use of oral hypoglycemic agents.

SAPHIR-The "Salzburg Atherosclerosis Prevention Program in subjects at High Individual Risk" (SAPHIR) is an observational study conducted in the years 1999-2002 involving healthy unrelated subjects: 641 females from 39 to 67 years of age and 1092 males from 39 to 66 years of age. Study participants were recruited by health screening programs in large companies in and around the city of Salzburg as described recently.[79] All individuals were of West-Eurasian origin. Subjects with established coronary artery, cerebrovascular or peripheral arterial disease, congestive heart failure, valvular heart disease, chronic alcohol (more than three drinks a day)

or drug abuse, severe obesity (BMI>40kg/m²) and pregnant women were excluded. Informed consent was obtained from each participant. At baseline all study participants were subjected to a comprehensive screening examination. A detailed personal and family history was assessed via standardized questionnaires. A physical examination included measurement of anthropometric parameters such as weight, height, waist circumference and percentage body fat. Blood samples were collected after an overnight fasting period. Urinary creatinine (mg/dl) was measured using a modified kinetic Jaffe reaction (CREA®, Roche Diagnostics GmbH, Mannheim, Germany); Urinary albumin concentration (mg/l) was determined using the Tina-quant® assay (Roche Diagnostics GmbH, Mannheim, Germany).

ULSAM The Uppsala Longitudinal Study of Adult Men (ULSAM) was initiated in 1970. All 50-year-old men who were born between 1920 and 1924 and were living in Uppsala, Sweden, were invited to participate in a health survey that focused on identifying cardiovascular risk factors (n =2322), as described in detail at www.pubcare. uu.se/ULSAM.[80] The present analyses are based on the third examination cycle of the ULSAM cohort, when participants were approximately 71 years of age (baseline 1991 to 1995).[80,81] Of the 1221 participants who attended this reinvestigation, 1027 had available DNA and valid measurements of urinary albumin excretion rate. Urinary albumin excretion rate (UAER) was calculated on the amount of albumin in the urine collected during the night. The subjects were instructed to void immediately before going to bed and to record the time. All samples during the night and the first sample of urine after rising were collected and used for the analysis. The assay employed a commercially available radioimmunoassay kit (Albumin RIA 100, Pharmacia, Uppsala, Sweden). Covariates were obtained at the time of albuminuria collection.

References (Supplement)

42. Whitlock MC: Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *J Evol Biol* 18**:** 1368-1373, 2005
43. Bacanu SA, Devlin B & Roeder K: The power of genomic control. *Am J Hum Genet* 66**:** 1933-1944, 2000
44. Aulchenko YS, Ripke S, Isaacs A & van Duijn CM: GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 23**:** 1294-1296, 2007
45. Abecasis GR, Cherny SS, Cookson WO & Cardon LR: Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30**:** 97-101, 2002
46. Chen MH & Yang Q: GWAF: an R package for genome-wide association analyses with family data. *Bioinformatics* 26**:** 580-581,
47. Mitchell BD, McArdle PF, Shen H, Rampersaud E, Pollin TI, Bielak LF, Jaquish C, Douglas JA, Roy-Gagnon MH, Sack P, Naglieri R, Hines S, Horenstein RB, Chang YP, Post W, Ryan KA, Brereton NH, Pakyz RE, Sorkin J, Damcott CM, O'Connell JR, Mangano C, Corretti M, Vogel R, Herzog W, Weir MR, Peyser PA & Shuldiner AR: The genetic response to short-term interventions affecting cardiovascular function: rationale and design of the Heredity and Phenotype Intervention (HAPI) Heart Study. *Am Heart J* 155**:** 823-828, 2008
48. Rampersaud E, Bielak LF, Parsa A, Shen H, Post W, Ryan KA, Donnelly P, Rumberger JA, Sheedy PF, 2nd, Peyser PA, Shuldiner AR & Mitchell BD: The association of coronary artery calcification and carotid artery intima-media thickness with distinct, traditional coronary artery disease risk factors in asymptomatic adults. *Am J Epidemiol* 168**:** 1016-1023, 2008
49. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* 129**:** 687-702, 1989
50. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA & Reich D: Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38**:** 904-909, 2006
51. Hsu CC, Brancati FL, Astor BC, Kao WH, Steffes MW, Folsom AR & Coresh J: Blood pressure, atherosclerosis, and albuminuria in 10,113 participants in the atherosclerosis risk in communities study. *J Hypertens* 27**:** 397-409, 2009
52. Shock NW GR, Andres R, Arenberg D, Costa. P, et al: Normal Human Aging: The Baltimore Study of Aging. *NIH publication* 84, 1984
53. Tanaka T, Scheet P, Giusti B, Bandinelli S, Piras MG, Usala G, Lai S, Mulas A, Corsi AM, Vestrini A, Sofi F, Gori AM, Abbate R, Guralnik J, Singleton A, Abecasis GR, Schlessinger D, Uda M & Ferrucci L: Genome-wide association study of vitamin B6, vitamin B12, folate, and homocysteine blood concentrations. *Am J Hum Genet* 84**:** 477-482, 2009
54. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A & et al.: The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1**:** 263-276, 1991
55. Firmann M, Mayor V, Vidal PM, Bochud M, Pecoud A, Hayoz D, Paccaud F, Preisig M, Song KS, Yuan X, Danoff TM, Stirnadel HA, Waterworth D, Mooser V, Waeber G & Vollenweider P: The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovasc Disord* 8**:** 6, 2008
56. Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A & Wareham N: EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* 80 Suppl 1**:** 95-103, 1999

57. Lee CT, Adler AI, Forouhi NG, Luben R, Welch A, Khaw KT, Bingham S & Wareham NJ: Cross-sectional association between fish consumption and albuminuria: the European Prospective Investigation of Cancer-Norfolk Study. *Am J Kidney Dis* 52**:** 876-886, 2008
58. Dawber TR, Kannel WB & Lyell LP: An approach to longitudinal studies in a community: the Framingham Study. *Ann N Y Acad Sci* 107**:** 539-556, 1963
59. Feinleib M, Kannel WB, Garrison RJ, McNamara PM & Castelli WP: The Framingham Offspring Study. Design and preliminary data. *Prev Med* 4**:** 518-525, 1975
60. Garrison RJ, Castelli WP, Feinleib M, Kannel WB, Havlik RJ, Padgett SJ & McNamara PM: The association of total cholesterol, triglycerides and plasma lipoprotein cholesterol levels in first degree relatives and spouse pairs. *Am J Epidemiol* 110**:** 313-321, 1979
61. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'Agostino RB, Sr., Fox CS, Larson MG, Murabito JM, O'Donnell CJ, Vasan RS, Wolf PA & Levy D: The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 165**:** 1328-1335, 2007
62. Wichmann HE, Gieger C & Illig T: KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* 67 Suppl 1**:** S26-30, 2005
63. Baumeister SE, Böger CA, Krämer BK, Döring A, Eheberg D, Fischer B, John J, Koenig W & Meisinger C: Effect of Chronic Kidney Disease and Comorbid Conditions on Health Care Costs: A 10-Year Observational Study in a General Population. *Am J Nephrol* 31**:** 222-229, 2009
64. Pattaro C, Marroni F, Riegler A, Mascalzoni D, Pichler I, Volpato CB, Dal Cero U, De Grandi A, Egger C, Eisendle A, Fuchsberger C, Gogele M, Pedrotti S, Pinggera GK, Stefanov SA, Vogl FD, Wiedermann CJ, Meitinger T & Pramstaller PP: The genetic study of three population microisolates in South Tyrol (MICROS): study design and epidemiological perspectives. *BMC Med Genet* 8**:** 29, 2007
65. John U, Greiner B, Hensel E, Ludemann J, Piek M, Sauer S, Adam C, Born G, Alte D, Greiser E, Haertel U, Hense HW, Haerting J, Willich S & Kessler C: Study of Health In Pomerania (SHIP): a health examination survey in an east German region: objectives and design. *Soz Praventivmed* 46**:** 186-194, 2001
66. Friedman GD, Cutter GR, Donahue RP, Hughes GH, Hulley SB, Jacobs DR, Jr., Liu K & Savage PJ: CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol* 41**:** 1105-1116, 1988
67. Taylor HA, Jr., Wilson JG, Jones DW, Sarpong DF, Srinivasan A, Garrison RJ, Nelson C & Wyatt SB: Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. *Ethn Dis* 15**:** S6-4-17, 2005
68. Fuqua SR, Wyatt SB, Andrew ME, Sarpong DF, Henderson FR, Cunningham MF & Taylor HA, Jr.: Recruiting African-American research participation in the Jackson Heart Study: methods, response rates, and sample description. *Ethn Dis* 15**:** S6-18-29, 2005
69. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Jr., Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M & Tracy RP: Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 156**:** 871-881, 2002
70. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF, Hoffman HJ & Gudnason V: Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* 165**:** 1076-1087, 2007
71. Effect of intensive therapy on the development and progression of diabetic nephropathy in the Diabetes Control and Complications Trial. The Diabetes Control and Complications (DCCT) Research Group. *Kidney Int* 47**:** 1703-1720, 1995

72. Sustained effect of intensive treatment of type 1 diabetes mellitus on development and progression of diabetic nephropathy: the Epidemiology of Diabetes Interventions and Complications (EDIC) study. *JAMA* 290**:** 2159-2167, 2003
73. Daniels PR, Kardia SL, Hanis CL, Brown CA, Hutchinson R, Boerwinkle E & Turner ST: Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med* 116**:** 676-681, 2004
74. Zemunik T, Boban M, Lauc G, Jankovic S, Rotim K, Vatavuk Z, Bencic G, Dogas Z, Boraska V, Torlak V, Susac J, Zobic I, Rudan D, Pulanic D, Modun D, Mudnic I, Gunjaca G, Budimir D, Hayward C, Vitart V, Wright AF, Campbell H & Rudan I: Genome-wide association study of biochemical traits in Korcula Island, Croatia. *Croat Med J* 50**:** 23-33, 2009
75. Tönjes A, Zeggini E, Kovacs P, Bottcher Y, Schleinitz D, Dietrich K, Morris AP, Enigk B, Rayner NW, Koriath M, Eszlinger M, Kemppinen A, Prokopenko I, Hoffmann K, Teupser D, Thiery J, Krohn K, McCarthy MI & Stumvoll M: Association of FTO variants with BMI and fat mass in the self-contained population of Sorbs in Germany. *Eur J Hum Genet* 18**:** 104-110,
76. Holmen J MK, Kruger O The Nord-Trøndelag health study 1995-97 (HUNT 2): objectives, methods and participation. *Norsk Epidemiologi* 13**:** 19-32, 2003
77. Curhan GC, Knight EL, Rosner B, Hankinson SE & Stampfer MJ: Lifetime nonnarcotic analgesic use and decline in renal function in women. *Arch Intern Med* 164**:** 1519-1524, 2004
78. Forman JP, Fisher ND, Schopick EL & Curhan GC: Higher levels of albuminuria within the normal range predict incident hypertension. *J Am Soc Nephrol* 19**:** 1983-1988, 2008
79. Heid IM, Wagner SA, Gohlke H, Iglseder B, Mueller JC, Cip P, Ladurner G, Reiter R, Stadlmayr A, Mackevics V, Illig T, Kronenberg F & Paulweber B: Genetic architecture of the APM1 gene and its influence on adiponectin plasma levels and parameters of the metabolic syndrome in 1,727 healthy Caucasians. *Diabetes* 55**:** 375-384, 2006
80. Byberg L, Zethelius B, McKeigue PM & Lithell HO: Changes in physical activity are associated with changes in metabolic cardiovascular risk factors. *Diabetologia* 44**:** 2134-2139, 2001
81. Ingelsson E, Sundstrom J, Lind L, Riserus U, Larsson A, Basu S & Arnlov J: Low-grade albuminuria and the incidence of heart failure in a community-based cohort of elderly men. *Eur Heart J* 28**:** 1739-1745, 2007

Supplementary Figure Legends

Supplementary Figure 1: Genome-wide -$\log_{10}$ p-value plot from Stage 1 analyses of microalbuminuria from participants of European ancestry in the CKDGen Consortium (Panel A) and the CARe Consortium (IBC chip analyses, Panel B)

Supplementary Figure 2: Quantile-quantile plots of observed vs. expected -$\log_{10}$ (p-values) from CKDGen discovery analyses of UACR (Panel A) and microalbuminuria (Panel B).

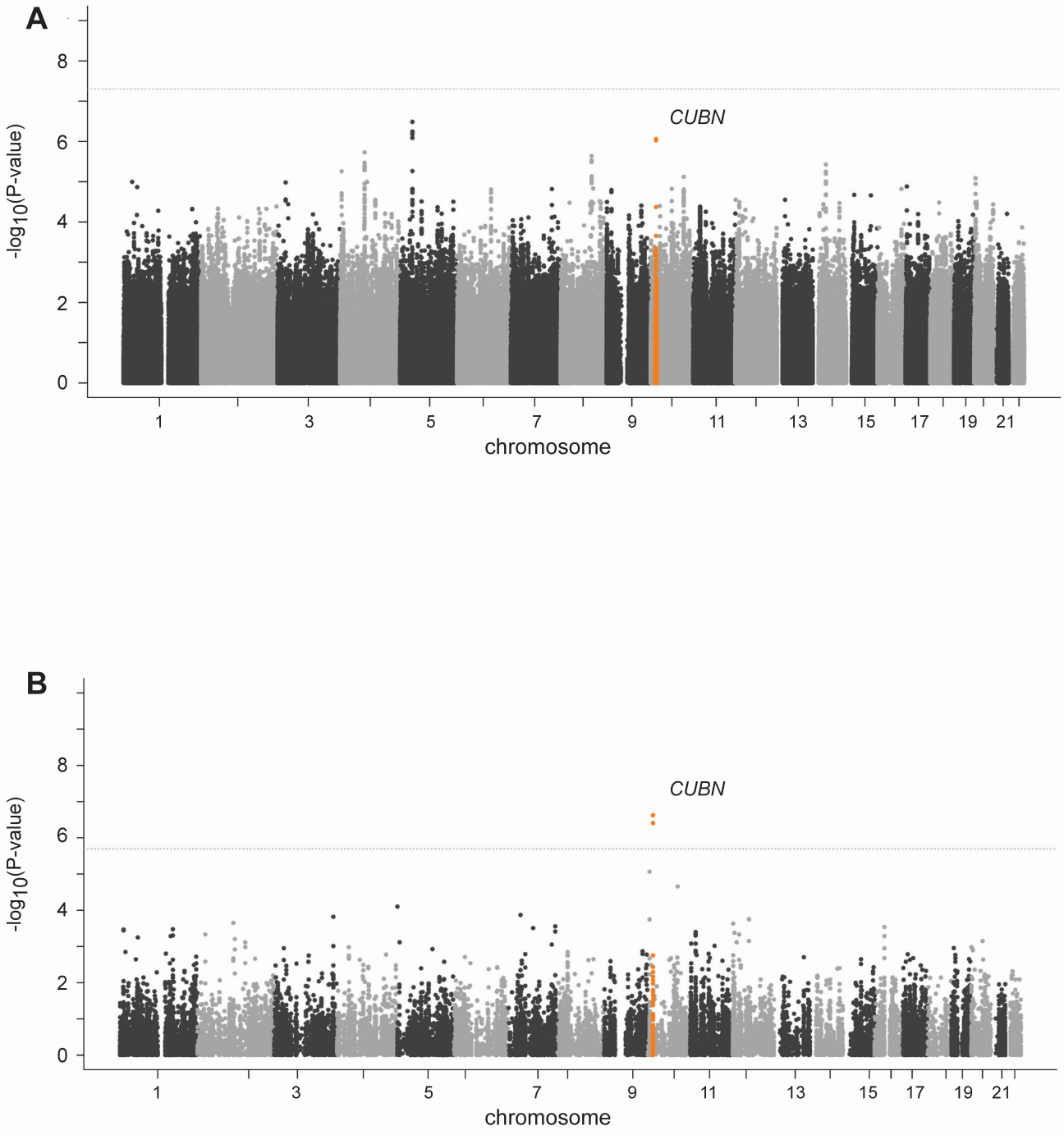
Footnote: λ_{meta} represents the genomic control parameters after discovery meta-analysis, and the λ for the individual studies is reported next to each trait. The graphs present p-values corrected for inflation at the study-specific level before meta-analysis as well as after meta-analysis for the meta-analysis genomic control parameter (if $\lambda_{meta}>1$). No correction was applied to data from studies with $\lambda<1$. NA denotes phenotype unavailability. Black: results from meta-analysis, orange: null hypothesis.

Supplementary Figure 3: Forest plot showing the effect estimates (beta) for the association of rs1801239 with UACR across CKDGen and CARe cohorts of European (Panel A) and African American (Panel B) descent.

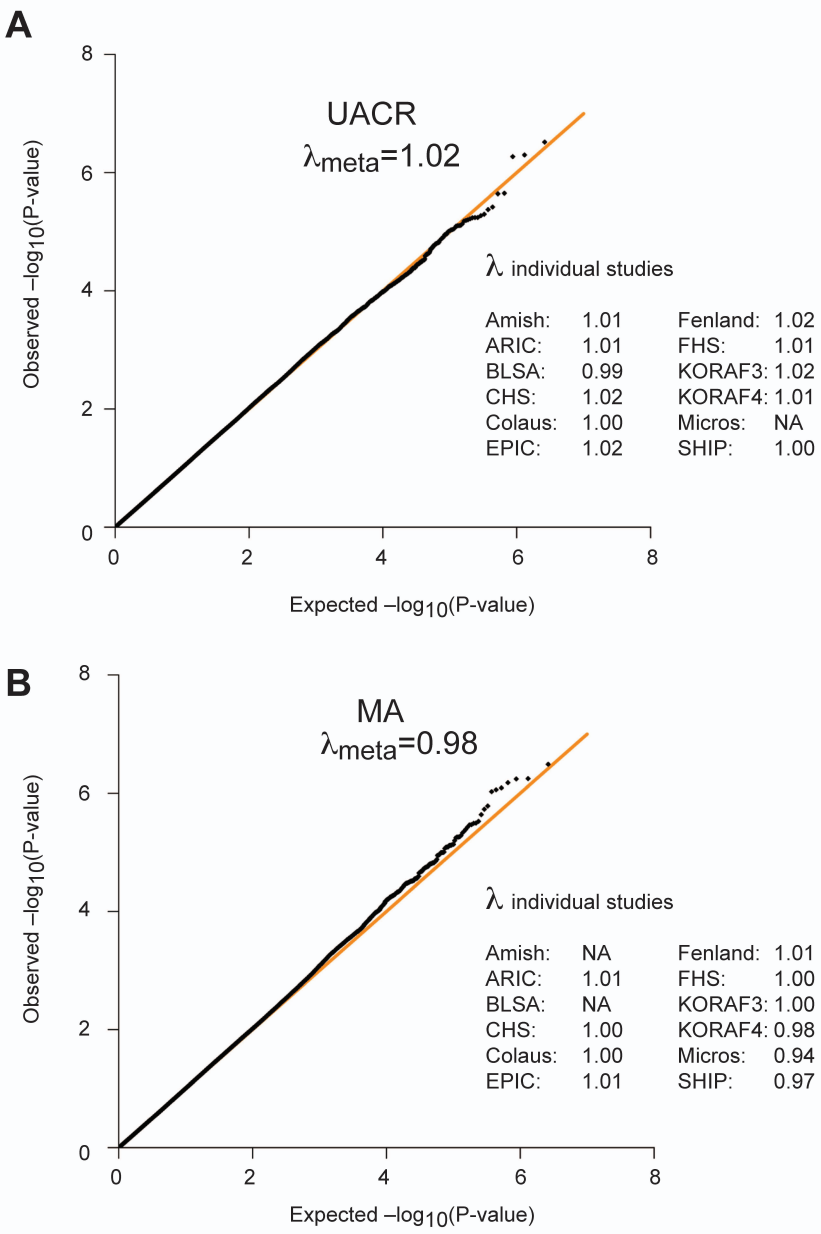
Footnote: The minor allele C is the effect allele. For ARIC, FHS and CHS (European descent) the effect estimates are shown using the directly genotyped data from the IBC chip. The square size reflects the relative cohort size. The error bars show the 95% confidence interval.

Supplementary Figure 4: Flow chart showing the overall study design.

Suppl. Figure 1: Genome-wide -log10 p-value plot from discovery analyses of microalbuminuria (MA) from participants of European ancestry in the CKDGen Consortium (Panel A) and the CARe Consortium (IBC chip analyses, Panel B)

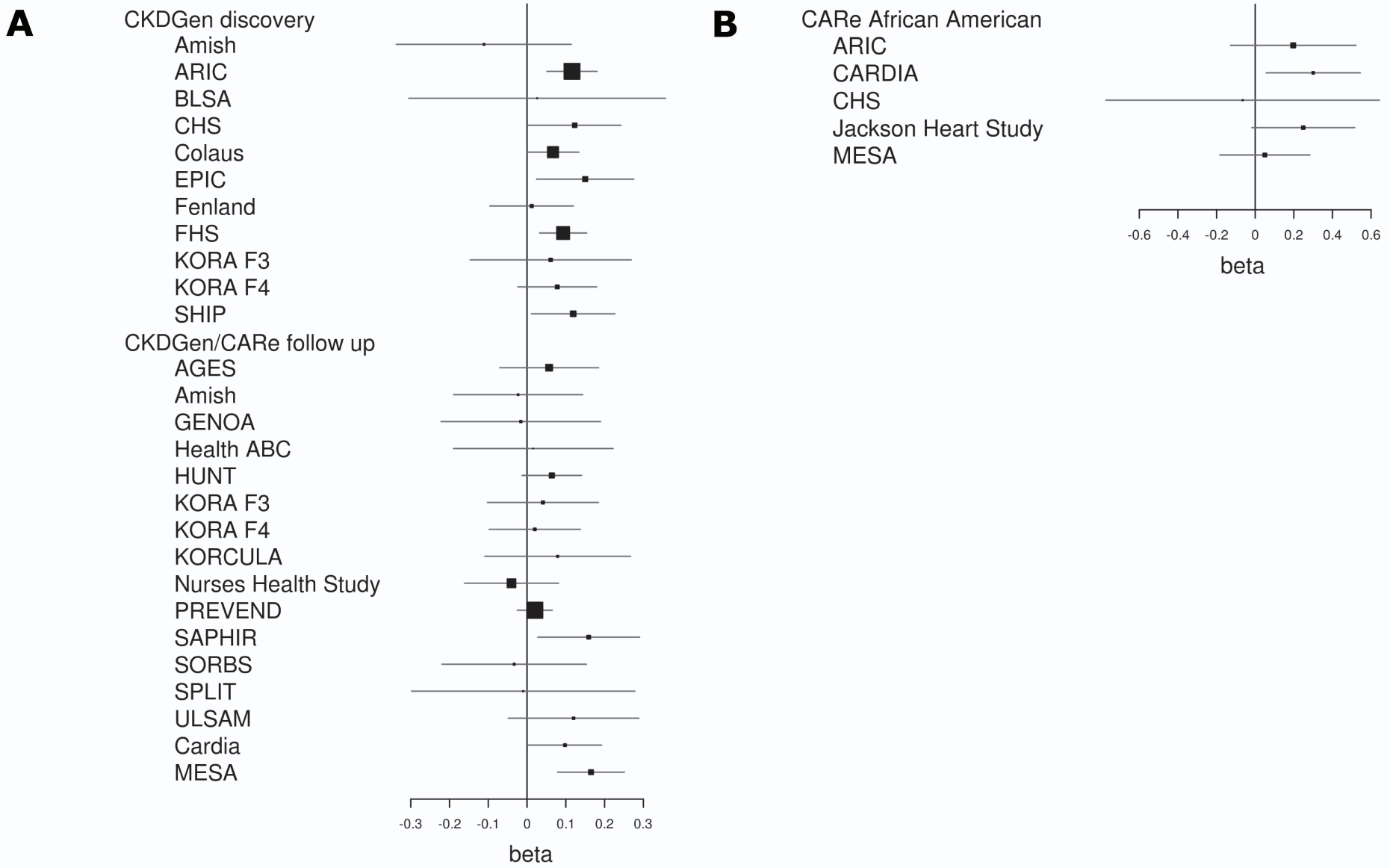


Suppl. Figure 2: Quantile-quantile plots of observed vs. expected -$\log_{10}$(p-values) from discovery analyses of UACR (A) and microalbuminuria (B) in the CKDGen Consortium.



λ_{meta} represents the genomic control parameters after discovery meta-analysis, and the λ for the individuals studies is reported next to each trait. The graphs present p-values corrected for inflation at the study-specific level before meta-analysis as well as after meta-analysis for the meta-analysis genomic control parameter. No correction was applied to data from studies with λ<1. NA denotes phenotype unavailability. Black: results from meta-analysis, orange: null hypothesis.

Supplementary Figure 3: Forest plot showing the effect estimates (beta) for the association of rs1801239 with UACR across CKDGen and CARe cohorts of European (Panel A) and African American (Panel B) descent.



The minor allele C is the effect allele. For ARIC, FHS and CHS (European descent) the effect estimates are shown using the directly genotyped data from the IBC chip. The square size reflects the relative cohort size. The error bars show the 95% confidence interval.

Supplementary Figure 4:
Flow chart showing the overall study design.