ABSTRACT

Identification of genetic risk factors for albuminuria may alter strategies for early prevention of CKD progression, particularly among patients with diabetes. Little is known about the influence of common genetic variants on albuminuria in both general and diabetic populations. We performed a meta-analysis of data from 63,153 individuals of European ancestry with genotype information from genome-wide association studies (CKDGen Consortium) and from a large candidate gene study (CARe Consortium) to identify susceptibility loci for the quantitative trait urinary albumin-to-creatinine ratio (UACR) and the clinical diagnosis microalbuminuria. We identified an association between a missense variant (I2984V) in the CUBN gene, which encodes cubilin, and both UACR and microalbuminuria. We observed similar associations among 6981 African Americans in the CARe Consortium. The associations between this variant and both UACR and microalbuminuria were significant in individuals of European ancestry regardless of diabetes status. Finally, this variant associated with a 41% increased risk for the development of persistent microalbuminuria during 20 years of follow-up among 1304 participants with type 1 diabetes in the prospective DCCT/EDIC Study. In summary, we identified a missense CUBN variant that associates with levels of albuminuria in both the general population and in individuals with diabetes.
Elevated levels of urinary albumin (albuminuria) are a cardinal manifestation of chronic kidney disease (CKD) and affect as many as 8% of adults from the United States\(^1\) and 6% of adults from Germany.\(^2\) Higher levels of albuminuria, even within the low normal range, are associated with not only increased risks of ESRD but also cardiovascular disease and mortality.\(^3\)–\(^6\) Moreover, the presence of albuminuria offers key prognostic information at each stage of decline in GFR.\(^7\) However, the pathophysiologic basis of albuminuria remains incompletely understood, and as a result, interventions for the prevention and treatment of albuminuria are limited.

Diabetes mellitus and hypertension are key risk factors for albuminuria, but neither of these factors fully account for the high prevalence of albuminuria nor its association with adverse health outcomes.\(^8\) Heritability of albuminuria ranges from 0.16 to 0.49 in families enriched with hypertension or diabetes.\(^9\),\(^10\) Rare genetic variants are known to cause monogenic diseases featuring severe, nephrotic range proteinuria.\(^11\) However, linkage or candidate gene studies have not reproducibly identified common genetic variants in association with lower levels of albuminuria.\(^9\),\(^10\) Given recent successes in the use of genome-wide association studies (GWAS) of quantitative traits that can lead to the identification of relevant variants for a disease phenotype,\(^12\),\(^13\) we conducted a genome-wide association (GWA) analysis of albuminuria in 31,580 participants of European ancestry from the CKDGen Consortium, with follow-up in 27,746 additional participants. Albuminuria was analyzed as the quantitative trait urinary albumin-to-creatinine ratio (UACR) and as the dichotomous trait microalbuminuria (MA). Concurrently, we performed an analysis of albuminuria in the CARe Consortium using the ITMAT/Broad/CARE Vascular Disease 50k (IBC) single-nucleotide polymorphism (SNP) chip array\(^14\) in 19,499 Europeans and 6981 African Americans. Here, we report the results of our combined findings.

RESULTS

Study Samples
Basic characteristics of the participants from the studies in CKDGen and CARe are shown in Table 1. Studies in these consortia are primarily population-based, with mean age ranging from 42 to 74 years. Details regarding study-specific genotyping information can be found in Supplemental Table 1, A and B, and in the Supplemental Text (Methods).

CKDGen Stage 1
Figure 1A and Supplemental Figure 1A show the Manhattan plots of the meta-analysis \(P\) values for UACR and microalbuminuria (UACR \(>25\) mg/g [women], \(>17\) mg/g [men]), respectively. Meta-analysis of GWAs from CKDGen stage 1 showed that no locus achieved genome-wide significance \((P \leq 5.0 \times 10^{-8})\) for either UACR or microalbuminuria in both the overall and the nondiabetic analyses. Supplemental Figure 2, A and B, shows the quantile-quantile plot of the UACR and microalbuminuria meta-analysis results.

CKDGen Stage 2 Follow-up
In CKDGen, the 16 top independent SNPs (\(P\) value range \(1.1 \times 10^{-7}\) to \(5.7 \times 10^{-6}\)) were moved into stage 2 follow-up in 15 additional studies \((n = 27,746\) individuals of European descent). These SNPs and their study-specific imputation scores are displayed in Supplemental Tables 2 and 3, respectively. Supplemental Table 4 shows the results of these 16 SNPs for all analyzed traits.

Overall, rs1801239, a missense SNP (T\(\rightarrow\)C) located in CUBN on chromosome 10 (minor allele frequency = 0.10), demonstrated direction-consistent association in stage 2 \((P = 0.02,\) Table 2\), with a genome-wide significant \(P\) value of \(4.0 \times 10^{-8}\) for UACR in the combined stage 1 and stage 2 analysis (Supplemental Tables 2 and 4). The regional association plot for CUBN is shown in Figure 2A. The minor C allele of rs1801239 in CUBN leads to an isoleucine-to-valine substitution (I2984V) in the encoded protein cubilin, which is predicted to be “probably benign” (SIFT,\(^15\) FastSNP,\(^16\) and PolyPhen\(^17\)). No additional nonsynonymous coding variants in high LD to rs1801239 \((r^2 > 0.2)\) were observed in dbSNP. Because albuminuria is a risk factor for progressive CKD, we assessed whether rs1801239 is associated with estimated GFR (eGFR) and CKD in the CKDGen eGFR data set. Among 67,093 individuals with available data, we observed no association with eGFR \((P = 0.53)\) or CKD \((P = 0.33)\).\(^18\)

The second highest ranking SNP for UACR in the combined stage 1 and stage 2 analysis was rs17319721, an intronic SNP in SHROOM3 (Supplemental Table 4). The minor allele (A) of rs17319721 is associated with lower albuminuria levels, and we have previously identified this same allele in association with lower eGFR.\(^19\)

CARe IBC SNP Array Stage 1
Concurrently, the CARe Consortium carried out a large densely tagged candidate gene screen using the IBC SNP ar-
Table 1. Study sample characteristics in the CKDGen and CARe Consortia

<table>
<thead>
<tr>
<th>Study</th>
<th>n(^a)</th>
<th>Women %</th>
<th>Age, years</th>
<th>UACR(^b)</th>
<th>MA %</th>
<th>CKD %</th>
<th>HTN %</th>
<th>DM %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage 1: GWAS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amish</td>
<td>744</td>
<td>45.0</td>
<td>42.8</td>
<td>3.8 (2.2, 6.1)</td>
<td>NA</td>
<td>1.2</td>
<td>11.6</td>
<td>0.7</td>
</tr>
<tr>
<td>ARIC</td>
<td>6525</td>
<td>53.1</td>
<td>63.1</td>
<td>3.9 (2.0, 7.7)</td>
<td>9.4</td>
<td>9.1</td>
<td>42.9</td>
<td>13.6</td>
</tr>
<tr>
<td>BLSA</td>
<td>354</td>
<td>46.9</td>
<td>69.8</td>
<td>7 (4.4, 11.0)</td>
<td>NA</td>
<td>15.5</td>
<td>53.7</td>
<td>14.4</td>
</tr>
<tr>
<td>CHS</td>
<td>1865</td>
<td>60.8</td>
<td>71.2</td>
<td>9.3 (5.3, 19.9)</td>
<td>23.0</td>
<td>16.0</td>
<td>61.2</td>
<td>13.7</td>
</tr>
<tr>
<td>Colaus</td>
<td>5311</td>
<td>53.2</td>
<td>53.4</td>
<td>5.1 (3.4, 9.1)</td>
<td>9.5</td>
<td>3.8</td>
<td>36.1</td>
<td>9.6</td>
</tr>
<tr>
<td>EPIC</td>
<td>2371</td>
<td>53.3</td>
<td>59.2</td>
<td>3.6 (1.5, 8.3)</td>
<td>8.1</td>
<td>NA</td>
<td>49.3</td>
<td>3.0</td>
</tr>
<tr>
<td>Fenland</td>
<td>1398</td>
<td>56.2</td>
<td>44.9</td>
<td>4.5 (3.2, 7.1)</td>
<td>5.5</td>
<td>0.9</td>
<td>18.9</td>
<td>1.4</td>
</tr>
<tr>
<td>Framingham Heart Study</td>
<td>6523</td>
<td>53.2</td>
<td>47.6</td>
<td>4.6 (2.6, 9.9)</td>
<td>9.6</td>
<td>3.4</td>
<td>27.4</td>
<td>5.3</td>
</tr>
<tr>
<td>KORA F3</td>
<td>1530</td>
<td>50.5</td>
<td>62.5</td>
<td>4.9 (2.1, 11.1)</td>
<td>11.7</td>
<td>10.8</td>
<td>41.1</td>
<td>10.3</td>
</tr>
<tr>
<td>KORA F4</td>
<td>1803</td>
<td>51.3</td>
<td>60.9</td>
<td>10.7 (7.0, 18.4)</td>
<td>20.9</td>
<td>7.0</td>
<td>20.9</td>
<td>9.2</td>
</tr>
<tr>
<td>Micros</td>
<td>503</td>
<td>56.3</td>
<td>46.0</td>
<td>NA</td>
<td>5.4</td>
<td>3.6</td>
<td>40.6</td>
<td>3.7</td>
</tr>
<tr>
<td>SHIP</td>
<td>2653</td>
<td>48.9</td>
<td>55.3</td>
<td>8.9 (5.0, 20.6)</td>
<td>25.3</td>
<td>8.5</td>
<td>54.2</td>
<td>12.4</td>
</tr>
<tr>
<td><strong>Stage 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGES</td>
<td>3196</td>
<td>57.9</td>
<td>76.4</td>
<td>2.66 (1.2, 7.0)</td>
<td>11.9</td>
<td>24.2</td>
<td>80.6</td>
<td>11.6</td>
</tr>
<tr>
<td>Amish</td>
<td>695</td>
<td>54.4</td>
<td>47.2</td>
<td>7 (4.3, 13.5)</td>
<td>NA</td>
<td>NA</td>
<td>21.0</td>
<td>11.5</td>
</tr>
<tr>
<td>ARIC</td>
<td>759</td>
<td>56.0</td>
<td>63.0</td>
<td>3.9 (2.1, 7.5)</td>
<td>7.6</td>
<td>7.7</td>
<td>39.3</td>
<td>13.1</td>
</tr>
<tr>
<td>GENOA</td>
<td>987</td>
<td>43.0</td>
<td>59.0</td>
<td>3.1 (1.7, 5.9)</td>
<td>6.4</td>
<td>10.7</td>
<td>73.5</td>
<td>15.8</td>
</tr>
<tr>
<td>Health ABC</td>
<td>211</td>
<td>27.0</td>
<td>73.9</td>
<td>13.9 (6.1, 49.6)</td>
<td>42.2</td>
<td>25.0</td>
<td>64.1</td>
<td>100.0</td>
</tr>
<tr>
<td>HUNT</td>
<td>2386</td>
<td>54.5</td>
<td>51.4</td>
<td>6.1 (4.4, 9.2)</td>
<td>8.6</td>
<td>4.7</td>
<td>43.4</td>
<td>4.0</td>
</tr>
<tr>
<td>KORA F3</td>
<td>1392</td>
<td>52.4</td>
<td>50.8</td>
<td>4.4 (1.9, 9.6)</td>
<td>10.9</td>
<td>5.5</td>
<td>28.6</td>
<td>4.5</td>
</tr>
<tr>
<td>KORA F4</td>
<td>1197</td>
<td>52.3</td>
<td>49.2</td>
<td>10.1 (6.6, 20.1)</td>
<td>23.7</td>
<td>5.9</td>
<td>13.4</td>
<td>4.0</td>
</tr>
<tr>
<td>KORCULA</td>
<td>863</td>
<td>63.9</td>
<td>56.3</td>
<td>4.6 (2.5, 10.4)</td>
<td>12.4</td>
<td>7.5</td>
<td>57.3</td>
<td>10.4</td>
</tr>
<tr>
<td>Nurses Health Study</td>
<td>4138</td>
<td>100.0</td>
<td>58.2</td>
<td>2.8 (1.7, 4.5)</td>
<td>3.6</td>
<td>10.4</td>
<td>36.1</td>
<td>4.4</td>
</tr>
<tr>
<td>PREVEND</td>
<td>7803</td>
<td>50.5</td>
<td>49.5</td>
<td>7.1 (4.8, 13.3)</td>
<td>16.7</td>
<td>6.4</td>
<td>33.5</td>
<td>3.7</td>
</tr>
<tr>
<td>SAPHIR Study</td>
<td>1707</td>
<td>37.3</td>
<td>51.4</td>
<td>3.8 (2.3, 8.2)</td>
<td>9.8</td>
<td>1.4</td>
<td>56.5</td>
<td>3.2</td>
</tr>
<tr>
<td>SORBS</td>
<td>877</td>
<td>58.7</td>
<td>48.0</td>
<td>6.8 (4.6, 13.5)</td>
<td>13.6</td>
<td>4.0</td>
<td>52.1</td>
<td>9.2</td>
</tr>
<tr>
<td>SPLIT</td>
<td>472</td>
<td>56.5</td>
<td>49.3</td>
<td>2.5 (1.3, 5.8)</td>
<td>7.8</td>
<td>3.6</td>
<td>39.4</td>
<td>5.1</td>
</tr>
<tr>
<td>ULSAM</td>
<td>1063</td>
<td>0.0</td>
<td>71.0</td>
<td>7.5 (4.9, 16.1)</td>
<td>15.9</td>
<td>15.2</td>
<td>74.0</td>
<td>11.0</td>
</tr>
<tr>
<td><strong>CARe Consortium</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>European Americans(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CARDIA</td>
<td>1242</td>
<td>50.4</td>
<td>40.7</td>
<td>4.4 (3.3, 6.9)</td>
<td>4.5</td>
<td>0.53</td>
<td>9.3</td>
<td>5.0</td>
</tr>
<tr>
<td>MESA</td>
<td>2289</td>
<td>52.1</td>
<td>62.7</td>
<td>4.7 (3.1, 8.6)</td>
<td>9.6</td>
<td>9.7</td>
<td>38.7</td>
<td>6.3</td>
</tr>
<tr>
<td>African Americans</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIC</td>
<td>2256</td>
<td>64.5</td>
<td>61.7</td>
<td>2.7 (0.9, 9.6)</td>
<td>16.3</td>
<td>7.3</td>
<td>70.9</td>
<td>27.7</td>
</tr>
<tr>
<td>CARDIA</td>
<td>1089</td>
<td>56.3</td>
<td>39.5</td>
<td>4.3 (1.3, 7.4)</td>
<td>9.6</td>
<td>0.9</td>
<td>24.7</td>
<td>6.4</td>
</tr>
<tr>
<td>CHS</td>
<td>426</td>
<td>64.3</td>
<td>72.1</td>
<td>10.9 (5.0, 26.9)</td>
<td>31.9</td>
<td>18.4</td>
<td>68.2</td>
<td>24.4</td>
</tr>
<tr>
<td>Jackson Heart Study</td>
<td>1530</td>
<td>58.5</td>
<td>51.8</td>
<td>6 (4.0, 13.0)</td>
<td>17.8</td>
<td>5.9</td>
<td>58.8</td>
<td>16.1</td>
</tr>
<tr>
<td>MESA</td>
<td>1680</td>
<td>54.6</td>
<td>62.2</td>
<td>5.5 (3.1, 13.2)</td>
<td>16.6</td>
<td>8.5</td>
<td>59.8</td>
<td>17.8</td>
</tr>
</tbody>
</table>

Data presented as means except where otherwise indicated. DM, diabetes; HTN, hypertension.

\(^a\)Refers to the successfully genotyped and analyzed sample with UACR or MA data, which may differ from the recruited sample described in Supplemental online Methods.

\(^b\)Median (interquartile range: 25th percentile, 75th percentile).

\(^c\)Refers to studies that are nonoverlapping with CKDGen; the sample size of ARIC, CHS, and FHS in CARe was 7687, 2073, and 6208, respectively.
SNPs in the region (0.05/158 = 0.0003), nor was it associated with UACR in individuals of European descent from CKDGen (P = 0.65) or CARe (P = 0.36).

Stratified Analysis
Because hypertension and diabetes are key risk factors for albuminuria, we performed stratified analyses for rs1801239 in the joint CKDGen/CARe analysis of populations of European descent. For UACR, we observed significant association among individuals with (n = 4915; P = 0.006) and without (P = 3.2 × 10^{-8}) diabetes as well as among individuals with (n = 13,097, P = 7.5 × 10^{-8}) and without (P = 1.3 × 10^{-6}) hypertension. Similar findings across these groups were also observed for microalbuminuria (Table 2). In particular, among individuals with diabetes, each copy of the C allele at rs1801239 was associated with an odds ratio of 1.27 for microalbuminuria (95% CI: 1.11 to 1.45).

To investigate if CKD modifies the association between rs1801239 and UACR, we performed an analysis of rs1801239 and UACR in 6 of the largest cohorts in CKDGen stratified by CKD status. We found that the association of the CUBN C allele and UACR was of similar magnitude among participants with CKD (n = 1808, β = 0.09, P = 0.26) as compared with those without CKD (n = 2129, β = 0.10, P = 9.6 × 10^{-11}), although the power was low in the CKD stratum because of its smaller sample size.

Independent Replication in the Diabetes Control and Complications Trial and Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC)
To understand the potential impact of rs1801239 on microalbuminuria in a high-risk population, we tested the association of rs1801239 with time to persistent microalbuminuria over 20 years of follow-up among 1304 participants of European ancestry with type 1 diabetes from the DCCT/EDIC Study (mean baseline age 26.7 years). With use of survival analysis, the minor C allele was associated with an increased risk of incident persistent microalbuminuria (hazard ratio 1.42 per copy of the C allele, 95% CI: 1.11 to 1.45). This association was essentially unchanged with further multivariable adjustment (Table 2).

DISCUSSION
We have identified and validated a missense SNP in the CUBN gene that is associated with albuminuria. This association is
robust across subgroups defined by diabetes and hypertension, two major risk factors for albuminuria, and among populations of both European and African ancestry. Finally, we have validated this finding in association with persistent microalbuminuria among patients with type 1 diabetes from the DCCT/EDIC Study.

Cubilin was first identified as the intrinsic factor/vitamin B12 complex receptor in the ileal mucosa. In the kidney, cubilin is expressed predominantly in the apical brush border of proximal tubular cells. We queried publicly available expression databases but did not find evidence for altered gene expression associated with the conservative amino acid substitution in cubilin encoded by rs1801239 (I2984V). On the basis of UniProt, the amino acid position 2984 is part of the 22nd, out of a total of 27, CUB domains. In vitro, CUB domains 22 through 27 demonstrated Ca2+-dependent binding to megalin. Together with megalin (LRP2) and amnionless (AMN), cubilin plays a key role in the receptor-mediated endocytotic reabsorption of albumin and other low-molecular-weight proteins. Interrogation of these respective genomic regions in our data did not reveal any significant findings. The process of endocytotic reabsorption of albumin is of importance because an estimated 3 g of albumin per day are not retained by the glomerular filter and enter the primary urine. Yet urine in its final composition is nearly devoid of proteins in healthy humans, highlighting the effectiveness of the tubular reabsorptive process to prevent significant protein loss.

### Table 2. Results for CUBN SNP rs1801239 on chromosome 10 in the CKDGen and CARe Consortia and DCCT/EDIC

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>UACR P Value</th>
<th>Microalbuminuria P value</th>
<th>Odds Ratio (95% CI) for Clinical Outcomes per Copy of Minor C Allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>CKDGen Stage 1 discovery</td>
<td>$3.0 \times 10^{-7}$</td>
<td>$8.7 \times 10^{-7}$</td>
<td>1.25 (1.15 to 1.37)</td>
</tr>
<tr>
<td>CARe IBC discovery in European Americans</td>
<td>$2.9 \times 10^{-10}$</td>
<td>$2.4 \times 10^{-7}$</td>
<td>1.31 (1.18 to 1.45)</td>
</tr>
<tr>
<td>CKDGen Stage 2 follow-up</td>
<td>0.02</td>
<td>0.43</td>
<td>1.01 (0.98 to 1.05)</td>
</tr>
<tr>
<td>Combined populations of European ancestry</td>
<td>$1.1 \times 10^{-11}$</td>
<td>0.001</td>
<td>1.06 (1.02 to 1.09)</td>
</tr>
<tr>
<td>CARe African Americans</td>
<td>0.005</td>
<td>0.008</td>
<td>1.42 (1.10 to 1.84)</td>
</tr>
<tr>
<td>Diabetes stratified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no diabetes</td>
<td>$3.2 \times 10^{-8}$</td>
<td>0.06</td>
<td>1.03 (1.00 to 1.07)</td>
</tr>
<tr>
<td>diabetes</td>
<td>0.006</td>
<td>$4.7 \times 10^{-4}$</td>
<td>1.27 (1.11 to 1.45)</td>
</tr>
<tr>
<td>Hypertension stratified</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no hypertension</td>
<td>$1.3 \times 10^{-6}$</td>
<td>$6.0 \times 10^{-4}$</td>
<td>1.23 (1.09 to 1.38)</td>
</tr>
<tr>
<td>hypertension</td>
<td>$7.5 \times 10^{-8}$</td>
<td>$1.4 \times 10^{-7}$</td>
<td>1.34 (1.20 to 1.49)</td>
</tr>
<tr>
<td>DCCT/EDIC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>simple model</td>
<td>0.02</td>
<td></td>
<td>1.42 (1.08 to 1.88)</td>
</tr>
<tr>
<td>extended model</td>
<td>0.02</td>
<td></td>
<td>1.41 (1.06 to 1.87)</td>
</tr>
<tr>
<td>Persistent MA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>simple model</td>
<td>0.53</td>
<td></td>
<td>1.14 (0.76 to 1.69)</td>
</tr>
<tr>
<td>extended model</td>
<td>0.67</td>
<td></td>
<td>1.10 (0.72 to 1.67)</td>
</tr>
</tbody>
</table>

**Notes:**
- Sample sizes are as follows (microalbuminuria case numbers in parentheses): CKDGen Stage 1: 31,580 (3,698 microalbuminuria cases); CARe European Americans: 19,499 (2,100 microalbuminuria cases).
- Combined populations of European ancestry, includes CKDGen Stage 1 and 2 (after removal of ARIC, CHS, and FHS) and all five cohorts of CARe European Americans: 63,153 (7,383 microalbuminuria cases).
- Combined populations of European ancestry, includes CKDGen Stage 1 (after removal of ARIC, CHS, and FHS) and all five cohorts of CARe European Americans: 36,166 (4,128 microalbuminuria cases).
- Modeled as hazard ratio; $n = 1,304$ including 318 cases of persistent microalbuminuria and 116 cases of severe nephropathy.
observed association of rs1801239 with incident persistent microalbuminuria in the DCCT/EDIC Study.

Our results suggest that levels of albuminuria in the general population are determined in part by tubular reabsorption, and not only by glomerular filtration. Although the prognostic implications of tubular as compared with glomerular albuminuria remain to be determined, a pathogenic role for tubular albuminuria, in addition to glomerular, has been demonstrated in experimental data. The identification of a SNP that is associated with albuminuria among individuals irrespective of diabetes or hypertension status suggests some common pathophysiology that is independent of these known albuminuria risk factors. The CUBN SNP we identified is specifically associated with albuminuria, and not with eGFR. Finally, results from the DCCT/EDIC Study underscore the relative strength of the CUBN SNP in association with persistent microalbuminuria that is comparable in magnitude to other albuminuria risk factors in patients with diabetes, including diabetes duration, BP, hemoglobin A1c, and obesity.

Important strengths of this study include consistency of association across populations of European and African descent and across groups defined by diabetes and hypertension, as well as the known role of cubilin in tubular albumin reabsorption. The exploration of genetic determinants for albuminuria in predominantly population-based cohorts reduces confounding by disease progression, which may be related to important nongenetic factors. Nonetheless, some limitations warrant mention. First, the causal nature of the missense SNP in CUBN is unclear. Second, urine albumin and creatinine were assessed with different assays and at one point in time in most studies, which may lead to misclassification of the outcome and bias our results toward the null hypothesis. However, this is unlikely to yield a true positive finding.

In summary, through a series of genetic association analyses, we have identified a missense SNP in the CUBN gene that is associated with higher levels of albuminuria among individuals of European and African descent with and without diabetes or...
hypertension. These findings highlight a novel genetic susceptibility for albuminuria that is consistent across multiple study populations and shared across diverse clinical settings.

CONCISE METHODS

Overall Study Design
Genetic association testing for urinary albumin-to-creatinine ratio (UACR) and MA was performed in the CKDGen and CARe cohorts of European ancestry, with further follow-up genetic analysis of significant SNPs in CARe cohorts of African-American ancestry and in the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Study. A graphical overview is given in Supplemental Figure 4.

CKDGen Stage 1 Discovery Meta-analysis
In the stage 1 (discovery) meta-analysis we searched for SNPs associated with UACR or MA in 12 CKDGen population-based cohorts totaling n = 31,580 patients of European descent. Individual GWA analyses encompassing approximately 2.5 million imputed SNPs were performed within each of the 12 stage 1 (discovery) CKDGen population-based cohorts. In each cohort, these analyses were performed in the overall group and separately in patients without diabetes. Next, we conducted four meta-analyses combining the study-specific UACR or MA GWA analysis results (i) including all patients and (ii) separately in patients without diabetes. From these four sets of meta-analysis results, we selected a list of independent SNPs (pairwise r² < 0.2) with a P value <1 × 10⁻⁶ and minor allele frequencies (MAF) >5%. The 16 highest ranking SNPs from this list were then selected for CKDGen stage 2 follow-up to replicate our findings.

CKDGen Stage 2 Follow-up Meta-analysis
Association testing for each of the 16 SNPs for UACR and MA was performed in each of the 15 independent CKDGen stage 2 cohorts totaling n = 27,746 individuals of European descent, again, including all patients and separately in patients without diabetes. Study-specific association results of the 16 SNPs were then meta-analyzed across stage 2 studies. SNPs showing evidence of replication in CKDGen stage 2 were further evaluated for their effects both in the presence and absence of diabetes and hypertension as major risk factors of albuminuria.

CARe Discovery Association Analysis
The CARe Consortium consists of nine studies. For the present analysis, we included five studies in European Americans (19,499 patients in total) and five studies in African Americans (6981 patients in total) with the IBC SNP chip. Study-specific genetic association analysis of UACR and MA were performed in the same manner as in CKDGen. CARe study-specific results were then meta-analyzed within each ethnic group for both UACR and MA.

Joint CKDGen and CARe Meta-analysis
For SNPs that reached genome-wide significance in the combined CKDGen stage 1 and stage 2 meta-analysis and were also significant in the CARe meta-analysis, we conducted meta-analyses for UACR and MA across a total of 28 nonoverlapping cohorts: in 9 CKDGen stage 1 studies with 16,667 patients (excluding ARIC, CHS, and FHS since they were also members of the CARe Consortium), 14 CKDGen stage 2 follow-up studies with 26,987 patients (excluding ARIC in silico results), and 5 CARe European American cohorts with 19,499 patients, involving a total of 63,153 patients of European descent.

Follow-up Analysis in DCCT/EDIC
Significant SNPs from the joint CKDGen and CARe meta-analysis were replicated in the DCCT/EDIC Study, which currently consists of 1304 Caucasian participants who underwent genotyping on the Illumina 1M SNP chip. This is a longitudinal study using Cox proportional hazards models to analyze time to renal events (see definition of the outcome below).

Study-Specific Information and Statistical Analysis
In all studies, all participants gave informed consent. All studies were approved by their appropriate Research Ethics Committees. A list of all contributing studies is given in Table 1 and more study-specific information including genotyping and imputation methods are given in Supplemental Table 1, A and B, as well as in the Study-Specific Methods Section of the Supplemental material. Details on statistical analyses on the study-specific level as well as for meta-analyses are given in the Supplemental material.

Outcomes and Covariates
In each of the CKDGen and CARe studies, the continuous outcome urinary albumin-to-creatinine ratio (mg/g, measured as described in the Study-Specific Methods section in the Supplemental material) as well as the dichotomous outcome microalbuminuria (MA, defined as urinary albumin-to-creatinine ratio >17 mg/g for men and >25 mg/g for women⁴⁹,⁵¹) were analyzed. For creating the continuous trait UACR used in the analysis, urinary albumin-to-creatinine ratio was log-transformed, and sex-specific residuals were computed by regression on age and, in multicenter studies, on study center.

Hypertension was defined as systolic BP ≥140 mmHg, diastolic BP ≥90 mmHg, or treatment; and diabetes was defined as fasting plasma glucose of at least 126 mg/dl or treatment if not stated otherwise in the study-specific methods. We estimated GFR (eGFR) using the four-variable MDRD formula as described previously⁵² to determine the prevalence of chronic kidney disease (defined as eGFR <60 ml/min per 1.73 m²) in each cohort.

In DCCT/EDIC, renal outcomes were (1) time from DCCT baseline until persistent microalbuminuria, defined as time to two consecutive albumin excretion rates (AERs) >30 mg/24 h (≥20.8 µg/min), and (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).

Genotypes
All the studies in CKDGen stage 1 had genotype data from genome-wide SNP arrays available, whereas the studies in the CARe Discovery studies genotyped the IBC SNP array, a gene-centric array containing 50,000 SNPs tagging genes across a range of cardiovascular, met-
abolic, and inflammatory syndromes. In CKDGen, the genotyped SNPs were imputed to approximately 2.5 million HapMap SNPs based on HAPMAP CEU samples. Imputation provides a common SNP panel across all studies to facilitate a meta-analysis across all contributing SNPs. Information on study-specific genotyping platforms and imputation procedures are presented in Supplemental Table 1, A and B. Information on genotyping in CKDGen stage 2 cohorts are given in the Study-Specific Methods section of the Supplemental material.

ACKNOWLEDGMENTS

Author contributions:


AMISH: The Amish studies are supported by grants and contracts from the NIH including R01 AG18728 (Amish Longevity Study), R01 HL088119 (AmishCalcification Study), U01 GM074518-04 (PAPI Study), U01 HL072515-06 (HAPI Study), U01 HL084756 and NIH K12RR023250 (University of Maryland MCRDP), NIH P30 DK072488 (Clinical Nutrition Research Unit of Maryland), and NIH P60 DK079637 (Baltimore Diabetes Research and Training Center). The University of Maryland General Clinical Research Center, grant M01 RR 16500, the Baltimore Veterans Administration Medical Center Geriatrics Research and Education Clinical Center, and the Paul Beeson Physician Faculty Scholars in Aging Program. We thank our Amish research volunteers for their long-standing partnership in research, and the research staff at the Amish Research Clinic for their hard work and dedication.

ARIC: The Atherosclerosis Risk in Communities study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL086737, and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by grant numbers UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

BLSA: This study was supported in part by the Intramural Research Program of the NIH, National Institute on Aging. A portion of that support was through a R&D contract with MedStar Research Institute.

CHARGE Consortium: We acknowledge the individual participating studies and investigators of the CHARGE (CoHORTs for Heart and Aging Research in Genome Epidemiology) consortium (AGES, ARIC, CHS, Framingham Heart Study, Rotterdam Study).

CHS: The Cardiovascular Health Study (CHS) research reported in this article was supported by contracts N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, and N01-HC-45133, grants U01 HL080295 and R01 HL087652, and R01 AG027002 from the National Heart, Lung and Blood Institute, with additional contribution from the National Institute of Neurologic Disorders and Stroke. A full list of principal CHS investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center for Research Resources grant M01 RR 16500, the Baltimore Veterans Administration Medical Center Geriatrics Research and Education Clinical Center, and the Paul Beeson Physician Faculty Scholars in Aging Program. We thank our Amish research volunteers for their long-standing partnership in research, and the research staff at the Amish Research Clinic for their hard work and dedication.


AMISH: The Amish studies are supported by grants and contracts from the NIH including R01 AG18728 (Amish Longevity Study), R01 HL088119 (AmishCalcification Study), U01 GM074518-04 (PAPI Study), U01 HL072515-06 (HAPI Study), U01 HL084756 and NIH K12RR023250 (University of Maryland MCRDP), NIH P30 DK072488 (Clinical Nutrition Research Unit of Maryland), and NIH P60 DK079637 (Baltimore Diabetes Research and Training Center). The University of Maryland General Clinical Research Center, grant M01 RR 16500, the Baltimore Veterans Administration Medical Center Geriatrics Research and Education Clinical Center, and the Paul Beeson Physician Faculty Scholars in Aging Program. We thank our Amish research volunteers for their long-standing partnership in research, and the research staff at the Amish Research Clinic for their hard work and dedication.

ARIC: The Atherosclerosis Risk in Communities study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL086737, and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by grant numbers UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

BLSA: This study was supported in part by the Intramural Research Program of the NIH, National Institute on Aging. A portion of that support was through a R&D contract with MedStar Research Institute.

CHARGE Consortium: We acknowledge the individual participating studies and investigators of the CHARGE (CoHORTs for Heart and Aging Research in Genome Epidemiology) consortium (AGES, ARIC, CHS, Framingham Heart Study, Rotterdam Study).

CHS: The Cardiovascular Health Study (CHS) research reported in this article was supported by contracts N01-HC-85079 through N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, and N01-HC-45133, grants U01 HL080295 and R01 HL087652, and R01 AG027002 from the National Heart, Lung and Blood Institute, with additional contribution from the National Institute of Neurologic Disorders and Stroke. A full list of principal CHS investigators and institutions can be found at http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center for Research Resources grant M01 RR 16500, the Baltimore Veterans Administration Medical Center Geriatrics Research and Education Clinical Center, and the Paul Beeson Physician Faculty Scholars in Aging Program. We thank our Amish research volunteers for their long-standing partnership in research, and the research staff at the Amish Research Clinic for their hard work and dedication.
**CKDGen Consortium:** We thank all members of the CKDGen Consortium who contributed toward the eGFR and CKD data; all names are listed at the end of the acknowledgments.

**CoLaus:** The CoLaus authors thank Yolande Barreau, Mathieu Firnan, Vladimir Mayor, Anne-Lise Bastian, Binas Ramie, Martine Moranville, Martine Baumont, Marcy Sagette, Jeanne Ecoffey, and Sylvie Mermod for data collection. The CoLaus study received financial contributions from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, the Swiss National Science Foundation (33CSCO-122661, 3200BO-111361/2, 3100AO-116323/1, and 310000-112552), the Swiss School of Public Health Plus, the Giorgi-Cavaglieri Foundation, and the European Framework Project 6 (EuroDia, Aneuploidy, and Hypergenes projects).

**EPIC:** The EPIC Norfolk Study is funded by program grants from the Medical Research Council UK and Cancer Research UK, and by additional support from the European Union, Stroke Association, British Heart Foundation, Department of Health, Food Standards Agency, and the Wellcome Trust.

**FENLAND:** We are grateful to all of the volunteers for their time and help, and to the general practitioners and practice staff for help with recruitment. We thank the Fenland Study Co-ordination team, the Field Epidemiology team, and the Fenland Study Investigators. The albuminuria measurement was performed by the Department of Clinical Biochemistry and Immunology. The Fenland Study is funded by the Wellcome Trust and the Medical Research Council, as well as by the Support for Science Funding program and CamStrad.

**Framingham Heart Study:** This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart, Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. The analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARE) project. This work was partially supported by the National Heart, Lung and Blood Institute’s Framingham Heart Study (contract no. N01-HC-25195) and its contract with Affymetrix, Inc. for genotyping services (contract no. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. The albuminuria measurement was supported by an American Diabetes Association Research Grant and by donation of microalbuminuria assay reagents from Roche Diagnostics Inc.

**KORA:** The genetic epidemiologic work was funded by the NIH subcontract from Children’s Hospital, Boston, (H.E.W. and I.M.H., prime grant 1 R01 DK075787-01A1), the German National Genome Research Net NGEN2 and NGENplus (H.E.W. 01GS0823; WK project A3, number 01GS0834), the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ, and by the Else Kröner-Fresenius-Stiftung (P48/08/A11/08; C.A.B. and B.K.K.). The kidney parameter measurements in F3 were funded by the Else Kröner-Fresenius-Stiftung (C.A.B. and B.K.K.) and the Regensburg University Medical Center, Germany; in F4 by the University of Ulm, Germany (W.K.). Genome-wide genotyping costs in F3 and F4 were in part funded by the Else Kröner-Fresenius-Stiftung (C.A.B. and B.K.K.). De novo genotyping in F3 and F4 were funded by the Else Kröner-Fresenius-Stiftung (C.A.B. and B.K.K.). The KORA research platform and the MONICA Augsburg studies were initiated and financed by the Helmholtz Zentrum München, German Research Center for Environmental Health, by the German Federal Ministry of Education and Research and by the State of Bavaria. Genotyping was performed in the Genome Analysis Center (GAC) of the Helmholtz Zentrum München. The LINUX platform for computation was funded by the University of Regensburg for the Department of Epidemiology and Preventive Medicine at the Regensburg University Medical Center.

**MICROS:** We owe a debt of gratitude to all participants. We thank the primary care practitioners Raffaella Stocker, Stefan Waldner, Toni Pizzecco, Josef Plangger, Ugo Marcadent, and the personnel of the Hospital of Silandro (Department of Laboratory Medicine) for their participation and collaboration in the research project. In South Tyrol, the study was supported by the Ministry of Health and Department of Educational Assistance, University and Research of the Autonomous Province of Bolzano, and the South Tyrolean Sparkasse Foundation.

**SHIP:** SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs, and the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the “Center of Knowledge Interchange” program of the Siemens AG.

**CARe Consortium:** The authors wish to acknowledge the support of the National Heart, Lung and Blood Institute and the contributions of the research institutions, study investigators, field staff, and study participants in creating this resource for biomedical research. The following nine parent studies have contributed parent study data, ancillary study data, and DNA samples through the Massachusetts Institute of Technology-Broad Institute (N01-HC-65226) to create this genotype/phenotype database for wide dissemination to the biomedical research community:

- **Atherosclerotic Risk in Communities (ARIC):** University of North Carolina at Chapel Hill (N01-HC-55015), Baylor Medical College (N01-HC-55016), University of Mississippi Medical Center (N01-HC-55021), University of Minnesota (N01-HC-55019), Johns Hopkins University (N01-HC-55020), University of Texas, Houston (N01-HC-55022), University of North Carolina, Forsyth County (N01-HC-55018).
- **Cardiovascular Health Study (CHS):** University of Washington (N01-HC-85079), Wake Forest University (N01-HC-85080), Johns Hopkins University (N01-HC-85081), University of Pittsburgh (N01-HC-85082), University of California, Davis (N01-HC-85083), University of California, Irvine (N01-HC-85084), New England Medical Center (N01-HC-85085), University of Vermont (N01-HC-85086), Georgetown University (N01-HC-35129), Johns Hopkins University (N01-HC-15103), University of Wisconsin (N01-HC-75150), Geisinger Clinic (N01-HC-45133), University of Washington (N01 HC-55222, U01 HL080295).
Study (CFS): Case Western Reserve University (NIH HL 46380, M01RR00080), Cooperative Study of Sickle Cell Disease (CSSCD): University of Illinois (N01-HB-72982, N01-HB-97062), Howard University (N01-HB-72991, N01-HB-97061), University of Miami (N01-HB-72992, N01-HB-97064), Duke University (N01-HB-72993), George Washington University (N01-HB-72994), University of Tennessee (N01-HB-72995, N01-HB-97070), Yale University (N01-HB-72996, N01-HB-97072), Children’s Hospital-Philadelphia (N01-HB-72997, N01-HB-97056), University of Chicago (N01-HB-72998, N01-HB-97053), Medical College of Georgia (N01-HB-73000, N01-HB-97060), WA University (N01-HB-73001, N01-HB-97071), Jewish Hospital and Medical Center of Brooklyn (N01-HB-73002), Trustees of Health and Hospitals of the City of Boston, Inc. (N01-HB-73003), Children’s Hospital-Oakland (N01-HB-73004, N01-HB-97054), University of Mississippi (N01-HB-73005), St. Luke’s Hospital-New York (N01-HB-73006), Alta Bates-Herrick Hospital (N01-HB-97051), Columbia University (N01-HB-97058), St. Jude’s Children’s Research Hospital (N01-HB-97066), Research Foundation, State University of New York-Albany (N01-HB-97068, N01-HB-97069), New England Research Institute (N01-HB-97073), Interfaith Medical Center-Brooklyn (N01-HB-97085), Coronary Artery Risk in Young Adults (CARDIA); University of Alabama at Birmingham (N01-HC-48047), University of Minnesota (N01-HC-48048), Northwestern University (N01-HC-48049), Kaiser Foundation Research Institute (N01-HC-48050), University of Alabama at Birmingham (N01-HC-95095), Tufts-New England Medical Center (N01-HC-45204), Wake Forest University (N01-HC-45205), Harbor-UCLA Research and Education Institute (N01-HC-05187), University of California, Irvine (N01-HC-45134, N01-HC-95100), Framingham Heart Study (FHS): Boston University (N01-HC-25195, R01 HL092577-01A1, R01 HL076784, R01 AG028321), Jackson Heart Study (JHS): Jackson State University (N01-HC-95170), University of Mississippi (N01-HC-95171), Tougalo campus (N01-HC-95172), Multi-Ethnic Study of Atherosclerosis (MESA); University of Washington (N01-HC-95159), Regents of the University of California (N01-HC-95160), Columbia University (N01-HC-95161), Johns Hopkins University (N01-HC-95162, N01-HC-95168), University of Minnesota (N01-HC-95163), Northwestern University (N01-HC-95164), Wake Forest University (N01-HC-95165), University of Vermont (N01-HC-95166), New England Medical Center (N01-HC-95167), Harbor-UCLA Research and Education Institute (N01-HC-95169), Cedars-Sinai Medical Center (R01-HL-071205, University of Virginia (subcontract to R01-HL-071205), Sleep Heart Health Study (SHHS): Johns Hopkins University (U01 HL064360), Case Western University (U01 HL063463), University of California, Davis (U01 HL053916), University of Arizona (U01 HL053938), University of Minnesota (relocating in 2006 to Univ Arizona) (U01 HL053934), University of Pittsburgh (U01 HL077813), Boston University (U01 HL053941), MedStar Research Institute (U01 HL063429), Johns Hopkins University (U01 HL053937).

Stage 2 Cohorts: AGES: The Age, Gene/Environment Susceptibility—Reykjavik Study has been funded by NIH contract N01-AG-12100, the NIA Intramural Research Program, Hjartavard (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The researchers are indebted to the participants for their willingness to participate in the study.

AMESH: see above details for the Amish study; Replication genotyping was funded by the Else Kröner-Fresenius-Stiftung (P48/08// A11/08; to C.A. Böger, B.K. Krämer).

DCCT/EDIC: The DCCT/EDIC Research Group is sponsored through research contracts from the National Institute of Diabetes, Endocrinology and Metabolic Diseases of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the National Institutes of Health. The authors are grateful to the patients in the DCCT/EDIC cohort for their long-term participation. A.D.P. holds a Canada Research Chair in the Genetics of Complex Diseases. This work has received support from National Institute of Diabetes and Digestive and Kidney Diseases Grant No. 1DK-48050, through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging.

GENOA: The GENOA Study genome-wide association study was funded by NHLBI R01-HL087660.

Health ABC: This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The genome-wide association study was funded by NIA grant IR01AG032098-01A1 to Wake Forest University and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging.

HUNT: The HUNT Study is a collaboration between the HUNT Research Center, Faculty of Medicine, Norwegian University of Science and Technology; The Norwegian Institute of Public Health; Nord-Trøndelag County Council; and Central Norway Regional Health Authority. The authors thank the health service and people of Nord-Trøndelag for their endurance and participation.

KORCULA: This study was supported through the grants from the Medical Research Council UK; and Ministry of Science, Education and Sport of the Republic of Croatia (no. 108-1080315-0302) and the European Union framework program 6 EUROSPIAN project (contract LSHG-CT-2006-018947). We would like to acknowledge the invaluable contributions of the recruitment team in Korcula, the administrative teams in Croatia and Edinburgh, and the people of Korcula. The albuminuria measurements were supported by grants from the Belgian agencies FNRS and FRSM (3.4.592.06F), the “Fondation Alphonse & Jean Forton”, the Programme d’excellence “Marshall” DIANE convention from the Region Walonne, the Interuniversity Attraction Pole (IUAP P6/05), and the Genecure (FP6) and EUNIFRON (FP7, GA#201590) programs of the European Community. We would also like to thank S. Druart and N. Amraoui for excellent technical assistance.
Nurses Health Study: This study is supported by the National Institutes of Health grants DK007791, DK066574, CA087969, and CA050385. Replication genotyping was funded by the NHLBI Intramural research program and the Center for Population Studies.

PREVEND: PREVEND study is supported by the Dutch Kidney Foundation (Grant E033), The Netherlands Heart Foundation (Grant 2006B140, 2006T003), and the EU project grant GENECURE (FP-6 LSHE CT 2006 037697). P. van der Harst is supported by NWO VENI grant 916.76.170 and the ICIN.

SAPHIR: The SAPHIR study was partially supported by a grant from the Kamillo Eisner Stiftung from the Medizinische Forschungsgesellschaft Salzburg, from the Paracelsus Medical University Salzburg (08/07/039) to B. Paulweber, and by grants from the “Genomics of Lipid-associated Disorders – GOLD” of the “Austrian Genome Research Programme GEN-AU” and the Austrian Heart Fund to F. Kronenberg. Replication genotyping was funded by the Else Kröner-Fresenius-Stiftung (P48/08//A11/08; to C.A. Böger and B.K. Krämer).

SORBS: This work was supported by grants from the Interdisciplinary Centre for Clinical Research at the University of Leipzig (B27 to M.S., P.K., and A.T.; N06 to P.K.) and from the German Diabetes Association (to A.T. and P.K.). M.S. is supported by a grant from the DFG (KFO 152). R.M. is funded by European Commission under the Marie Curie Intra-European Fellowship. We would like to thank Knut Krohn from the Microarray Core Facility of the Interdisciplinary Centre for Clinical Research (IZKF), University of Leipzig, Germany, Joachim Thiery, Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University of Leipzig, Germany, Nigel William Rayner from the Wellcome Trust Centre for Human Genetics, University of Oxford, U.K., and John Broxholm from the Bioinformatics Core Unit of the Wellcome Trust Centre for Human Genetics, Oxford, U.K., for their excellent support.

SPLIT: This study was supported through the grants from the Medical Research Council UK and Ministry of Science, Education and Sport of the Republic of Croatia (no. 108-1080315-0302). We would like to acknowledge the invaluable contributions of the recruitment team from the Croatian Centre for Global Health, University of Split, the administrative teams in Croatia and Edinburgh (Rosa Bisset), and the people of Split. The albuminuria measurements were supported by grants from the Belgian agencies FNRS and FRSM (3.4.592.06F), the “Fondation Alphonse & Jean Forton”, the Programme d’excellence “Marshall” DIANE convention from the Region Walonne, the Interuniversity Attraction Pole (IUAP P6/05), and the Geneucore (FP6) and EUNEFRON (FP7, GA#201590) programs of the European Community. We would also like to thank S. Druart and N. Amraoui for excellent technical assistance.

ULSAM: Genotyping was performed by the SNP Technology Platform in Uppsala (www.genotyping.se). We thank Tomas Axelsson, Ann-Christine Wiman, and Caisa Pöntinen for their excellent assistance with genotyping. The SNP Technology Platform is supported by Uppsala University and the Knut and Alice Wallenberg Foundation. E.I. is supported by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Society of Medicine, the Swedish Foundation for Strategic Research, and the Royal Swedish Academy of Science. J.A. is supported by The Swedish Research Council (2006-6555), Swedish Heart-Lung foundation, Loo och Hans Ostermans foundation, Dalarna University, and Uppsala University.


DISCLOSURES
Dawn Waterworth and Vincent Mooser are employees of GlaxoSmithKline.

REFERENCES
minuria is associated with increased risk of coronary heart disease and death independently of renal function, hypertension, and dia-

7. Hemmelgarn BR, Manns BJ, Lloyd A, James MT, Klarenbach S, Quinn RR, Wiebe N, Tonelli M: Relation between kidney function, protein-
uria, and adverse outcomes. JAMA 303: 423–429, 2010


9. Fox CS, Yang Q, Guo CY, Cupples LA, Wilson PW, Levy D, Meigs JB: Genome-wide linkage analysis to urinary microalbuminuria in a com-


11. Tryggvason K, Patrakka J, Wartiovaara J: Hereditary proteinuria syn-


15. Kumar P, Henikoff S, Ng PC: Predicting the effects of coding non-


Edinburgh, Scotland; 67University of Michigan School of Public Health, Department of Epidemiology, University of Michigan, Ann Arbor, Michigan; 68Gen-Info Ltd., Zagreb, Croatia; 69Brigham and Women’s Hospital, Harvard Medical School, Boston, Massachusetts; 70Center for Population Health Sciences, University of Edinburgh Medical School, Edinburgh, Scotland; 71Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden; 72University of Maryland School of Medicine, Baltimore, Maryland; 73Welch Center for Prevention, Epidemiology & Clinical Research, Johns Hopkins University, Baltimore, Maryland; 74General Internal Medicine, University of California, San Francisco, San Francisco, California; 75Samuel Lunenfeld Research Institute of Mount Sinai Hospital, Prosserman Centre for Health Research, Toronto, Ontario, Canada; 76The Hospital for Sick Children, Toronto, Ontario, Canada; 77Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany; 78Klinikum Grosshadern, Munich, Germany; 79Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland; 80Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts; 81Institute of Physiology, University of Greifswald, Greifswald, Germany; 82University of Mississippi Division of Nephrology, University of Mississippi, Jackson, Mississippi; 83University Institute of Social and Preventive Medicine, Centre Hospitalier Universitaire Vaudois and University of Lausanne, IUMSP, Lausanne, Switzerland; and 84Division of Endocrinology, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts
Genetic Association Studies Identify CUBN as a Gene Locus for Albuminuria

SUPPLEMENTARY MATERIAL

Overview

1. Supplementary Tables Page 2
2. Statistical Methods Page 8
3. Study-Specific Methods Page 10
   3.a. CKDGen Stage 1 cohorts Page 10
   3.b. CARe Consortium Page 16
   3.c. CKDGen Stage 2 in silico cohorts Page 18
   3.d. CKDGen Stage 2 denovo genotyped cohorts Page 22
4. References Page 27
5. Supplementary Figure Legends Page 30
   5.a. Supplementary Figure 1 Page 31
   5.b. Supplementary Figure 2 Page 32
   5.c. Supplementary Figure 3 Page 33
   5.d. Supplementary Figure 4 Page 34
<table>
<thead>
<tr>
<th>Study Name</th>
<th>Array type</th>
<th>Genotype calling</th>
<th>QC filters for genotyped SNPs used for imputation</th>
<th>No of SNPs used for imputation</th>
<th>Imputation</th>
<th>Imputation Backbone (NCBI build)</th>
<th>Filtering of imputed genotypes</th>
<th>Data management and statistical analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGES</td>
<td>Illumina 370 HuCNV</td>
<td>Illumina Beadstudio</td>
<td>call rate &lt;95%, pHWE &lt;10E-6</td>
<td>325094</td>
<td>MACH v1.0.16</td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
<td>R</td>
</tr>
<tr>
<td>Amish Studies</td>
<td>Affymetrix 500K</td>
<td>BRLMM</td>
<td>MAF &lt; 1%, non-HapMap, call rate &lt;95%, pHWE &lt;10E-6</td>
<td>338,598</td>
<td>MACH version 1.0.15</td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
<td>Measured genotype accounting for polygenic component</td>
</tr>
<tr>
<td>ARIC (Stage 1)</td>
<td>Affymetrix 6.0</td>
<td>Birdseed</td>
<td>call rate &lt;95%, MAF &lt; 1%, pHWE &lt;10E-5</td>
<td>602,642</td>
<td>MACH v1.0.16</td>
<td>HapMap release 21 (build 35)</td>
<td>none</td>
<td>R, probABEL, PLINK</td>
</tr>
<tr>
<td>ARIC (Stage 2)</td>
<td>Affymetrix 6.0</td>
<td>Birdseed</td>
<td>MAF &lt;1%, call rate &lt;95%, pHWE &lt;10E-5</td>
<td>669450</td>
<td>MACH Imputation (v1.0.16)</td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
<td>R, probABEL, PLINK</td>
</tr>
<tr>
<td>Baltimore Longitudinal Study on Aging</td>
<td>Illumina 550K version 3</td>
<td>Illumina Beadstudio</td>
<td>call rate &lt;99%, MAF &lt;1%, pHWE &lt;10E-4</td>
<td>501,764</td>
<td>MACH version 1.0.15</td>
<td>HapMap release 21 (build 35)</td>
<td>r2hat&lt;0.3</td>
<td>MERLIN</td>
</tr>
<tr>
<td>Cardiovascular Health Study</td>
<td>Illumina 370CNV</td>
<td>Illumina BeadStudio</td>
<td>call rate &lt;97%, heterozygotes=0, pHWE&lt;10E-5, SNP not in HapMap</td>
<td>306,655</td>
<td>BimBam</td>
<td>HapMap release 21A (build 36)</td>
<td>dosage variance &lt;0.01</td>
<td>R, robust variance option</td>
</tr>
<tr>
<td>CoLaus Study</td>
<td>Affymetrix 500K</td>
<td>BRLMM</td>
<td>call rate &lt;70%, pHWE &lt;10E-7</td>
<td>390631</td>
<td>IMPUTE version 0.2.0</td>
<td>HapMap release 21 (build 35)</td>
<td>none</td>
<td>Matlab</td>
</tr>
<tr>
<td>EPIC/cohort</td>
<td>Affymetrix 500K</td>
<td>BRLMM</td>
<td>call rate &lt;90%, pHWE &lt;10E-6</td>
<td>382,037</td>
<td>IMPUTE version 0.3.1</td>
<td>HapMap release 21 (Build 35)</td>
<td>none</td>
<td>SAS, Stata, Linux scripts</td>
</tr>
<tr>
<td>Fenland Study</td>
<td>Affymetrix 500K</td>
<td>BRLMM</td>
<td>call rate &lt;90%, pHWE &lt;10E-6, MAF &lt;1%</td>
<td>362,055</td>
<td>IMPUTE v0.4.2</td>
<td>HapMap release 22 (build 36)</td>
<td>proper_info&lt;0.4</td>
<td>LINUX Stata 10.1 SNPTEST v1.1.5</td>
</tr>
<tr>
<td>Framingham Heart Study</td>
<td>Affymetrix 500K</td>
<td>BRLMM</td>
<td>call rate &lt;95%, pHWE &lt;10E-6</td>
<td>503526</td>
<td>MACH version 1.0.15</td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
<td>R, GWAF</td>
</tr>
<tr>
<td>GENOA</td>
<td>Affymetrix 6.0</td>
<td>Birdseed-v2</td>
<td>call rate &lt;95%, pHWE &lt;10E-6</td>
<td>+/-50kb of target snps</td>
<td>MACH version</td>
<td>HapMap release 22</td>
<td>none</td>
<td>PLINK, R</td>
</tr>
</tbody>
</table>
## Supplementary Table 1a-Genotyping and Imputation Information in CKDGen Consortium

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Genotyping Platform</th>
<th>Imputation Software</th>
<th>Data Version</th>
<th>Quality Control Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>HABC</td>
<td>Illumina Human1M-Duo</td>
<td>Beadstudio</td>
<td>1.0.16</td>
<td>MACH version 1.0.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hapmap CEU release 22</td>
<td>none</td>
</tr>
<tr>
<td>KORA F3</td>
<td>Affymetrix 500K</td>
<td>BRLMM</td>
<td>380407</td>
<td>MACH version 1.0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HapMap release 22 (build 35)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>none</td>
<td>ProbABEL, R, SAS, Visual Basic</td>
</tr>
<tr>
<td>KORA F4</td>
<td>Affymetrix 6.0</td>
<td>BRLMM</td>
<td>629,893</td>
<td>MACH version 1.0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>none</td>
<td>ProbABEL, R, SAS, Visual Basic</td>
</tr>
<tr>
<td>KORCULA</td>
<td>HAP370CNV</td>
<td>Beadstudio</td>
<td>316730</td>
<td>MACH version 1.0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>none</td>
<td>R, GenABEL, ProbABEL</td>
</tr>
<tr>
<td>MICROS (South Tyrol)</td>
<td>Illumina, 300K</td>
<td>Beadstudio</td>
<td>229,917</td>
<td>MACH v1.0.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>none</td>
<td>R, GenABEL, ProbABEL; details in Study-Specific Methods</td>
</tr>
<tr>
<td>Sorbs (Slavonic population from Eastern Germany)</td>
<td>Affymetrix 500K; Affymetrix 6.0</td>
<td>500K:BRLMM algorithm Affymetrix 6.0:Birdseed</td>
<td>378513</td>
<td>IMPUTE version 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HapMap release 21 (build 35)</td>
<td>MAF&gt;1%, HWE&lt;10-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SNPTEST</td>
<td></td>
</tr>
<tr>
<td>SPLIT</td>
<td>HAP370CNV</td>
<td>Beadstudio</td>
<td>330997</td>
<td>MACH version 1.0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>none</td>
<td>R(GenABEL, ProbABEL)</td>
</tr>
<tr>
<td>Study of Health in Pomerania</td>
<td>Affymetrix 6.0</td>
<td>Birdsuite-V2</td>
<td>869224</td>
<td>Impute version 0.5.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HapMap release 22 (build 36)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>none</td>
<td>SNPTEST v1.1.5, Quicktest 0.94, R, InforSense, InterSystems Caché</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Overall number Genotyped</th>
<th>Total Number of Samples Genotyped that Pass QC</th>
<th>Data management and statistical analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIC</td>
<td>15103</td>
<td>3029 African American; 9588 Caucasians</td>
<td>PLINK</td>
</tr>
<tr>
<td>CARDIA</td>
<td>3490</td>
<td>1295 African American; 1443 Caucasians</td>
<td>PLINK</td>
</tr>
<tr>
<td>CHS</td>
<td>5362</td>
<td>752 African American; 3952 Caucasians</td>
<td>PLINK</td>
</tr>
<tr>
<td>FHS</td>
<td>8016</td>
<td>7556 Caucasians</td>
<td>FHS R-LME/GEE</td>
</tr>
<tr>
<td>JHS</td>
<td>2157</td>
<td>2036 African American</td>
<td>PLINK</td>
</tr>
<tr>
<td>MESA</td>
<td>6482</td>
<td>1612 African American; 2298 Caucasian; 637 Asians; 1302 Hispanics</td>
<td>PLINK</td>
</tr>
<tr>
<td>Trait</td>
<td>SNP ID</td>
<td>Chr position (b36)</td>
<td>Genes Within 60 kb or closest gene*</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>-------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>UACR</td>
<td>Overall rs7576149</td>
<td>2 20789943</td>
<td>C2orf43, GDF7</td>
</tr>
<tr>
<td></td>
<td>Overall rs3734316</td>
<td>6 37705107</td>
<td>MDGA1</td>
</tr>
<tr>
<td></td>
<td>Overall rs1801239</td>
<td>10 16959058</td>
<td>CUBN, RSU1</td>
</tr>
<tr>
<td></td>
<td>No DM rs9372871</td>
<td>6 127849645</td>
<td>C6orf174</td>
</tr>
<tr>
<td></td>
<td>No DM rs2838520</td>
<td>21 44440324</td>
<td>ICOSLG; C21orf33; DNMT3L</td>
</tr>
<tr>
<td>MA</td>
<td>Overall rs6829211</td>
<td>4 4601533</td>
<td>STX18</td>
</tr>
<tr>
<td></td>
<td>Overall rs17319721</td>
<td>4 77587871</td>
<td>SHROOM3</td>
</tr>
<tr>
<td></td>
<td>Overall rs2635165</td>
<td>8 98365385</td>
<td>TSPYL5</td>
</tr>
<tr>
<td></td>
<td>No DM rs850898</td>
<td>2 185921053</td>
<td>ZNF804A, F3IP2</td>
</tr>
<tr>
<td></td>
<td>No DM rs7649382</td>
<td>3 34071517</td>
<td>PDCD6IP</td>
</tr>
<tr>
<td></td>
<td>No DM rs13104825</td>
<td>4 186237102</td>
<td>HELT</td>
</tr>
<tr>
<td></td>
<td>No DM rs13160548</td>
<td>5 38814607</td>
<td>OSMR</td>
</tr>
<tr>
<td></td>
<td>No DM rs13215455</td>
<td>6 52847613</td>
<td>GSTA5; GSTA3</td>
</tr>
<tr>
<td></td>
<td>No DM rs4935985</td>
<td>11 126866026</td>
<td>KIRREL3</td>
</tr>
<tr>
<td></td>
<td>No DM rs13337289</td>
<td>16 75895552</td>
<td>ADAMTS18</td>
</tr>
<tr>
<td></td>
<td>No DM rs159782</td>
<td>20 4295347</td>
<td>ADRA1D</td>
</tr>
</tbody>
</table>

*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
<table>
<thead>
<tr>
<th>SNP</th>
<th>AMISH</th>
<th>ARIC</th>
<th>BLSA</th>
<th>CHS</th>
<th>COLAUS</th>
<th>EPIC</th>
<th>FENLAND</th>
<th>FHS</th>
<th>KORAF3</th>
<th>KORAF4</th>
<th>MICROS</th>
<th>SHIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs7576149</td>
<td>0.58</td>
<td>0.92</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td>0.62</td>
<td>0.64</td>
<td>0.92</td>
<td>0.69</td>
<td>0.89</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>rs850898</td>
<td>0.96</td>
<td>1.00</td>
<td>1.00</td>
<td>0.97</td>
<td>0.97</td>
<td>0.96</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>0.96</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>rs7649382</td>
<td>0.89</td>
<td>0.85</td>
<td>0.99</td>
<td>0.72</td>
<td>0.66</td>
<td>0.65</td>
<td>0.73</td>
<td>0.79</td>
<td>0.63</td>
<td>0.83</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td>rs13104825</td>
<td>0.44</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.01</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>rs13160548</td>
<td>0.99</td>
<td>1.00</td>
<td>0.99</td>
<td>0.08</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.89</td>
<td>0.90</td>
<td>0.89</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>rs3734316</td>
<td>0.91</td>
<td>0.88</td>
<td>0.88</td>
<td>0.91</td>
<td>0.82</td>
<td>0.88</td>
<td>0.88</td>
<td>0.89</td>
<td>0.90</td>
<td>0.88</td>
<td>0.87</td>
<td>0.92</td>
</tr>
<tr>
<td>rs13215455</td>
<td>0.62</td>
<td>0.91</td>
<td>0.91</td>
<td>0.90</td>
<td>0.92</td>
<td>0.91</td>
<td>0.96</td>
<td>0.88</td>
<td>0.96</td>
<td>0.96</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>rs9372871</td>
<td>0.62</td>
<td>0.89</td>
<td>0.91</td>
<td>0.88</td>
<td>0.92</td>
<td>0.88</td>
<td>0.96</td>
<td>0.88</td>
<td>0.92</td>
<td>0.96</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td>rs2635165</td>
<td>0.67</td>
<td>0.97</td>
<td>0.97</td>
<td>0.88</td>
<td>0.94</td>
<td>0.97</td>
<td>0.95</td>
<td>0.88</td>
<td>0.93</td>
<td>0.92</td>
<td>0.92</td>
<td>0.96</td>
</tr>
<tr>
<td>rs1801239</td>
<td>0.73</td>
<td>0.97</td>
<td>0.97</td>
<td>0.75</td>
<td>0.90</td>
<td>0.86</td>
<td>0.75</td>
<td>0.75</td>
<td>0.86</td>
<td>0.85</td>
<td>0.91</td>
<td>0.85</td>
</tr>
<tr>
<td>rs4935985</td>
<td>0.36</td>
<td>0.97</td>
<td>0.97</td>
<td>0.85</td>
<td>0.87</td>
<td>0.85</td>
<td>0.88</td>
<td>0.90</td>
<td>0.87</td>
<td>0.85</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>rs13337289</td>
<td>0.59</td>
<td>0.97</td>
<td>0.97</td>
<td>0.75</td>
<td>0.88</td>
<td>0.88</td>
<td>0.93</td>
<td>0.75</td>
<td>0.93</td>
<td>0.93</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td>rs159782</td>
<td>0.46</td>
<td>1.00</td>
<td>1.00</td>
<td>0.84</td>
<td>0.88</td>
<td>0.95</td>
<td>0.88</td>
<td>0.95</td>
<td>0.88</td>
<td>0.92</td>
<td>0.89</td>
<td>0.89</td>
</tr>
<tr>
<td>rs2838520</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>SNPID</td>
<td>Chr</td>
<td>Position</td>
<td>Coded allele</td>
<td>Genes Within 60 kb or closest gene*</td>
<td>UACR Overall</td>
<td>UACR No Diabetes</td>
<td>Microalbuminuria Overall</td>
<td>Microalbuminuria No Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----</td>
<td>------------</td>
<td>--------------</td>
<td>-------------------------------------</td>
<td>--------------</td>
<td>------------------</td>
<td>-------------------------</td>
<td>---------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7576149</td>
<td>2</td>
<td>20789943 t</td>
<td></td>
<td>C2orf43, GDF7</td>
<td>2.2E-06</td>
<td>4.0E-06</td>
<td>6.5E-02</td>
<td>1.2E-01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs850898</td>
<td>2</td>
<td>185921053 a</td>
<td>ZNF804A, FSIP2</td>
<td>4.0E-02, 9.0E-01, 1.3E-01</td>
<td>1.6E-02</td>
<td>1.7E-04</td>
<td>2.2E-06</td>
<td>2.2E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7649382</td>
<td>3</td>
<td>34071517 a</td>
<td>PDCD6IP</td>
<td>9.5E-02, 6.2E-01, 4.0E-01</td>
<td>4.6E-03</td>
<td>3.6E-05</td>
<td>6.8E-07</td>
<td>6.8E-07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs6289211</td>
<td>4</td>
<td>4601533 a</td>
<td>STX8</td>
<td>1.4E-02, 6.1E-01, 2.6E-02</td>
<td>1.4E-01</td>
<td>1.4E-04</td>
<td>1.9E-06</td>
<td>1.2E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17319721</td>
<td>4</td>
<td>77587871 a</td>
<td>SHROOM3</td>
<td>1.7E-04, 3.9E-03, 7.0E-07</td>
<td>2.4E-04</td>
<td>6.3E-02</td>
<td>2.1E-01</td>
<td>2.1E-01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs13104825</td>
<td>4</td>
<td>186237102 c</td>
<td>HELT</td>
<td>1.2E-03, 1.6E-01, 1.4E-01</td>
<td>2.3E-03</td>
<td>2.6E-02</td>
<td>3.2E-07, 1.3E-02</td>
<td>3.2E-07, 1.3E-02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs13160548</td>
<td>5</td>
<td>38814607 t</td>
<td>OSMR</td>
<td>3.7E-03, 8.3E-01, 2.1E-02</td>
<td>1.2E-03</td>
<td>2.3E-05</td>
<td>3.4E-07</td>
<td>3.4E-07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs3734316</td>
<td>6</td>
<td>37705107 t</td>
<td>MDGA1</td>
<td>2.3E-06, 1.4E-01, 5.2E-06</td>
<td>1.9E-05</td>
<td>2.3E-05</td>
<td>1.8E-07</td>
<td>1.8E-07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs13215455</td>
<td>6</td>
<td>52847613 t</td>
<td>GSTA5,GSTA3</td>
<td>2.2E-03, 2.8E-01, 1.6E-01</td>
<td>1.0E-04</td>
<td>2.5E-04</td>
<td>1.2E-06</td>
<td>1.2E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9372871</td>
<td>6</td>
<td>127849645 a</td>
<td>C6orf174</td>
<td>9.2E-03, 5.1E-01, 2.0E-02</td>
<td>5.7E-06</td>
<td>4.9E-05</td>
<td>3.8E-06</td>
<td>3.8E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2635166</td>
<td>8</td>
<td>98363585 a</td>
<td>TSPYL5</td>
<td>2.6E-03, 6.7E-01, 3.7E-03</td>
<td>3.9E-03</td>
<td>2.1E-03</td>
<td>3.1E-06</td>
<td>3.1E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1801239</td>
<td>10</td>
<td>16959058 t</td>
<td>CUBN, RSU1</td>
<td>3.0E-07, 2.0E-02, 4.0E-08</td>
<td>3.6E-04</td>
<td>8.7E-07</td>
<td>2.0E-05</td>
<td>2.0E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4935985</td>
<td>11</td>
<td>126086026 a</td>
<td>KIRREL3</td>
<td>4.6E-02, 9.4E-01, 1.1E-01</td>
<td>6.8E-03</td>
<td>4.1E-04</td>
<td>1.6E-06</td>
<td>1.6E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs13337289</td>
<td>16</td>
<td>75895552 t</td>
<td>ADAMTS18</td>
<td>1.2E-04, 2.1E-02, 2.1E-01</td>
<td>5.9E-05</td>
<td>1.5E-05</td>
<td>3.2E-06</td>
<td>3.2E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs159782</td>
<td>20</td>
<td>4295347 a</td>
<td>ADRA1D, ICOSLG,C21orf33, DNMT3L</td>
<td>7.0E-02, 6.7E-02, 1.1E-02</td>
<td>1.5E-02</td>
<td>8.2E-06</td>
<td>2.4E-06</td>
<td>2.4E-06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2838520</td>
<td>21</td>
<td>44440324 t</td>
<td></td>
<td>3.8E-06, 1.6E-01, 2.2E-02</td>
<td>1.2E-06</td>
<td>2.6E-02</td>
<td>2.9E-02</td>
<td>3.3E-01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Bold indicates the trait and stratrum 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](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] 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.\textsuperscript{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°C, 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.\textsuperscript{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;\textsuperscript{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 Jaffé method. 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 \(\geq 126\) mg/dl (7.0 mmol/L), non-fasting glucose of \(\geq 200\) 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. 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. 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. 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). 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. 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. 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. 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. In 1971, enrollment of the Offspring Cohort was started
(5,124 participants); the design and methodology has been described.\textsuperscript{59,60} In 2002, enrollment began for the Third Generation cohort (n=4095).\textsuperscript{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;\textsuperscript{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\textsuperscript{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, Valelunga and Martello. A detailed description of the MICROS study is available elsewhere. 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. 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. Briefly, CARDIA recruited a cohort of young black and white adults, age 18-30 at the time of enrollment. From 1985-1986 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. 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 Ortho-Clinical Diagnostics 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. 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. 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\textsuperscript{71} with coefficients of variation and coefficients of reliability of 14% and 95% respectively.\textsuperscript{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](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. 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 automated 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. 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. 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 (11 mmol/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. 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 (>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%). 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. 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. 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). 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

27


**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: $\lambda_{\text{meta}}$ represents the genomic control parameters after discovery meta-analysis, and the $\lambda$ 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 $\lambda_{\text{meta}}>1$). 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.

$\lambda_{meta}$ represents the genomic control parameters after discovery meta-analysis, and the $\lambda$ 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 $\lambda<1$. 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.

A. CKDGen discovery
   - Amish
   - ARIC
   - BLSA
   - CHS
   - Colaus
   - EPIC
   - Fenland
   - FHS
   - KORA F3
   - KORA F4
   - SHIP

   CKDGen/CARe follow up
   - AGES
   - Amish
   - GENOA
   - Health ABC
   - HUNT
   - KORA F3
   - KORA F4
   - KORCULA
   - Nurses Health Study
   - PREVEND
   - SAPHIR
   - SORBS
   - SPLIT
   - ULSAM
   - Cardia
   - MESA

B. CARe African American
   - ARIC
   - CARDIA
   - CHS
   - Jackson Heart Study
   - MESA

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.

GWA analysis in each of 12 cohorts
2.5 million SNPs
European Ancestry

GWA data from 12 cohorts

CKDGen stage 1 discovery
2.5 million SNPs
Meta-analysis
12 cohorts’ GWA results
European Ancestry
n=31580 samples

16 SNPs

CKDGen stage 2 follow up
16 SNPs
Meta-analysis
15 cohorts
European Ancestry
n=27746 samples

16 SNPs

CKDGen Stage 1 and Stage 2
16 SNPs
Joint Meta-analysis
27 cohorts
European Ancestry
n= 59326 samples

1 SNP p < 5e-8:
rs1801239
23 non-overlapping cohorts with
n=43654 samples

CKDGen + CARe joint analysis
1 SNP
Meta-analysis
28 non-overlapping cohorts
European Ancestry
n = 63,153 samples

1 SNP p<5e-8:
rs1801239

CARe discovery
50,000 SNPs (IBC chip)
Meta-analysis
5 cohorts
European Ancestry
n=19499 samples

2 SNPs p < 2.2e-6:
rs1801239 rs13177732
5 non-overlapping cohorts with
n=19499 samples

CARe
Meta-analysis
5 cohorts
African American
n=6981 samples

DCCT/EDIC
Association analysis
1 cohort
Diabetes mellitus type 1
n = 1304 samples