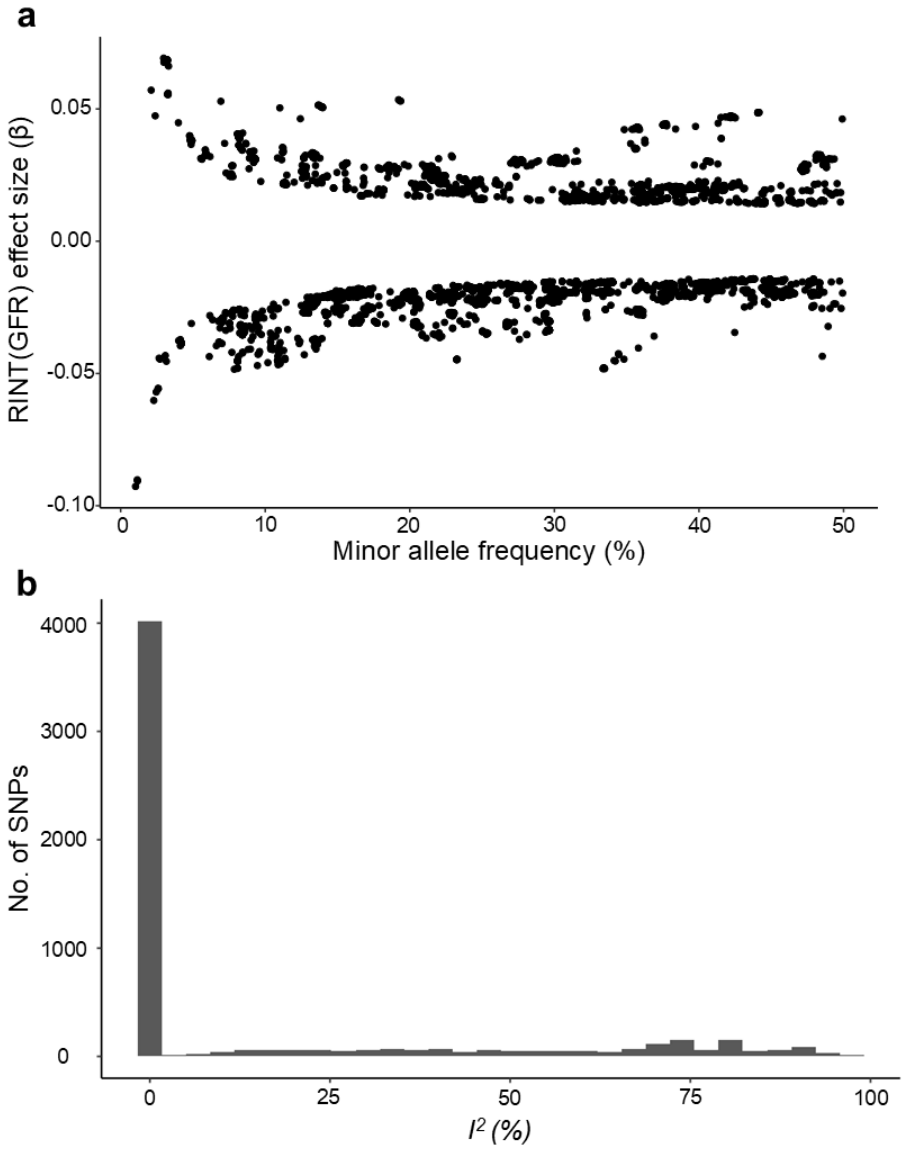
Discovery and prioritization of genetic determinants of kidney function in 297,355 individuals from Taiwan and Japan

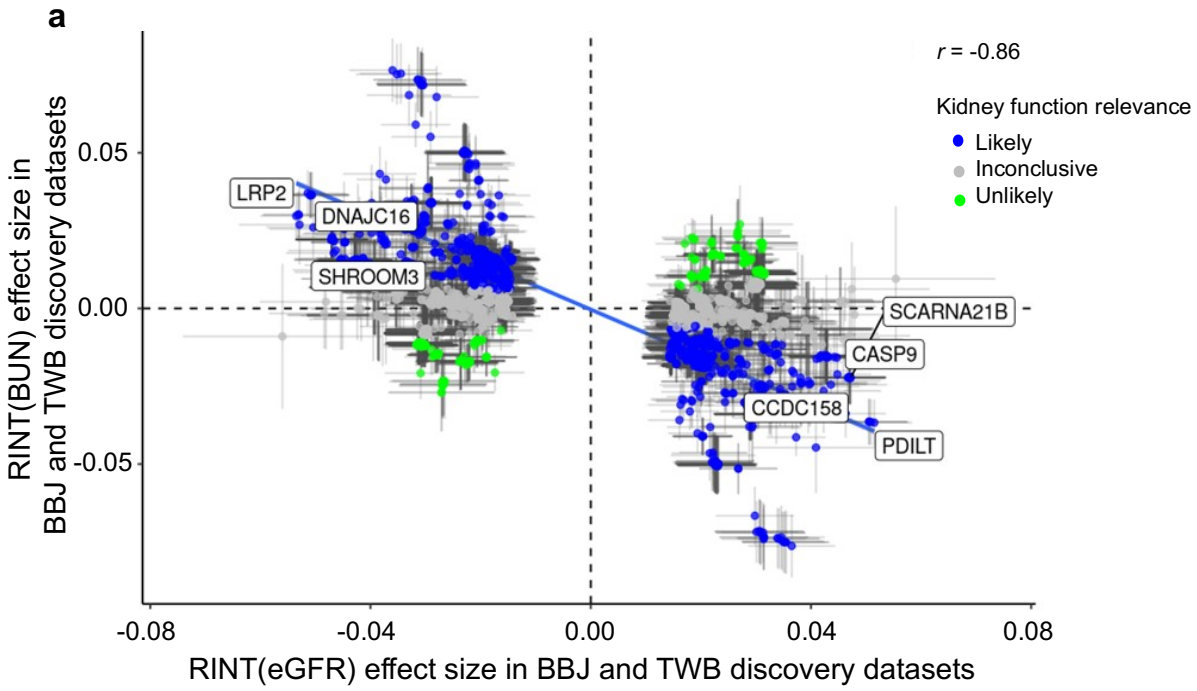
Supplement Information

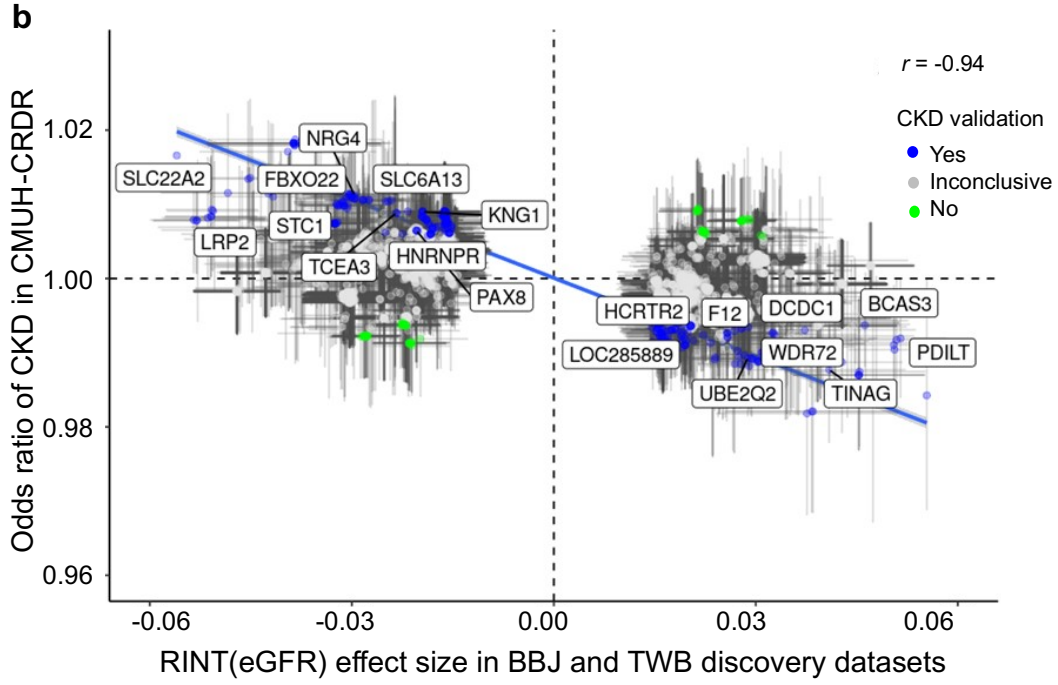
Supplementary Figure 1. Effect size and heterogeneity of 5,790 eGFR-associated SNPs (black circle) from a meta-analysis of BBJ (n=143,658) and TWB-based (n=101,294) discovery datasets. (a) Effect size and MAF of 5,790 eGFR-associated SNPs. Most SNPs with a higher MAF have lower effect sizes. **(b)** Most of the 5,790 SNPs are homogeneous (I^2 smaller than 75%). Further details are provided in the Supplementary Data 2.

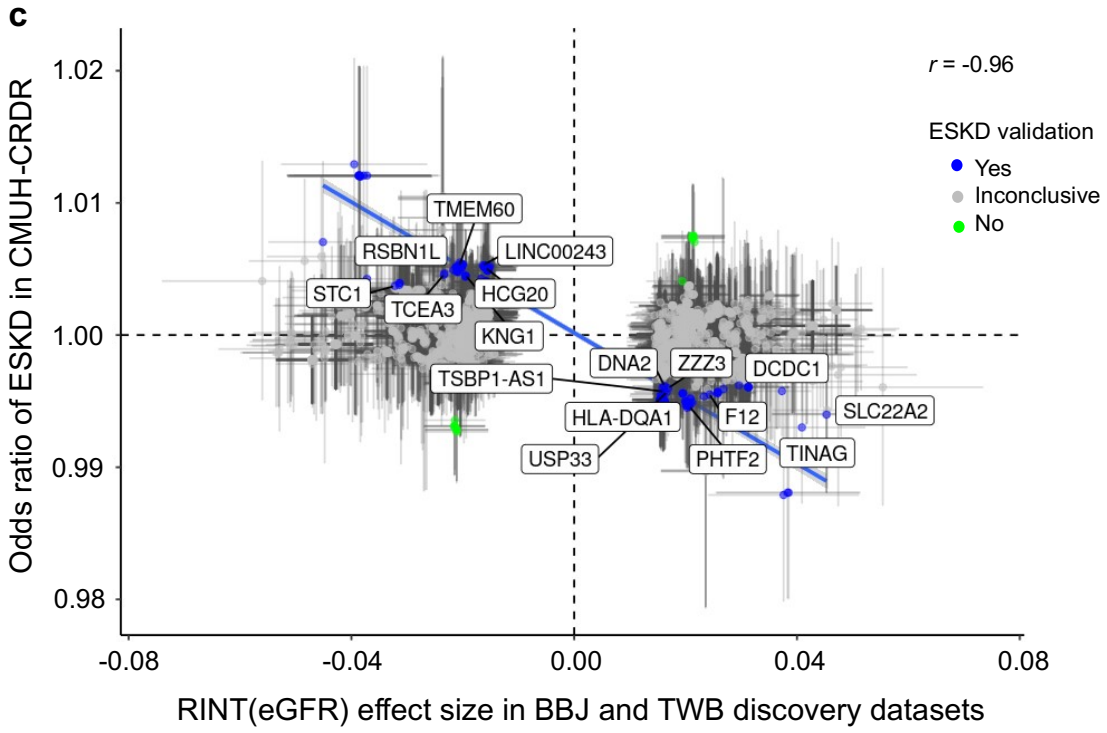
Abbreviations: BBJ, Biobank Japan; eGFR, estimated glomerular filtration rate; MAF, minor allele frequency; RINT, rank-based inverse normal transformation; SNP, single nucleotide polymorphism; TWB, Taiwan Biobank.



Supplementary Figure 2. Association of replicated eGFR-associated SNPs with kidney function and kidney disease. (a) In total, 2,064 replicated eGFR-associated SNPs were negatively correlated with BUN level from a meta-analysis of BBJ (n=143,658) and TWB (n=101,294) discovery datasets (Pearson's $r = -0.86$; $p < 0.0001$; Supplementary Data 5). **(b)** In total, 397 replicated eGFR-associated SNPs were negatively correlated with CKD in CMUH-CRDR (Pearson's $r = -0.94$, $p < 0.0001$; Supplementary Data 6). **(c)** In total, 162 replicated eGFR-associated SNPs were negatively correlated with ESKD in CMUH-CRDR (Pearson's $r = -0.96$, $p < 0.0001$; Supplementary Data 7). Logistic regression was applied for the association of SNPs with CKD and ESKD. For kidney function relevance, "Likely" is indicated by blue dots: the effect direction is opposite, and BUN p is <0.05; "Unlikely" is indicated by green dots: the effect direction is consistent, and BUN p is <0.05; "Inconclusive" is indicated by grey dots: BUN $p \geq 0.05$. For kidney disease validation, "Yes" is indicated by blue dots: the effect direction is opposite, and kidney disease p is <0.05; "No" is indicated by green dots: the effect direction is consistent, and kidney disease p is <0.05; "Inconclusive" is indicated by grey dots: kidney disease $p \geq 0.05$. The blue line represents the best fit of blue dots. Pearson's correlation coefficient (*r)* was applied. All statistical tests employed two-sided p values.
***Abbreviations*:** BBJ, Biobank Japan; BUN, blood urea nitrogen; CKD, chronic kidney disease; CMUH-CRDR, China Medial University Hospital-Clinical Research Data Repository; eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease; RINT, rank-based inverse normal transformation; SNP, single nucleotide polymorphism; TWB, Taiwan Biobank.

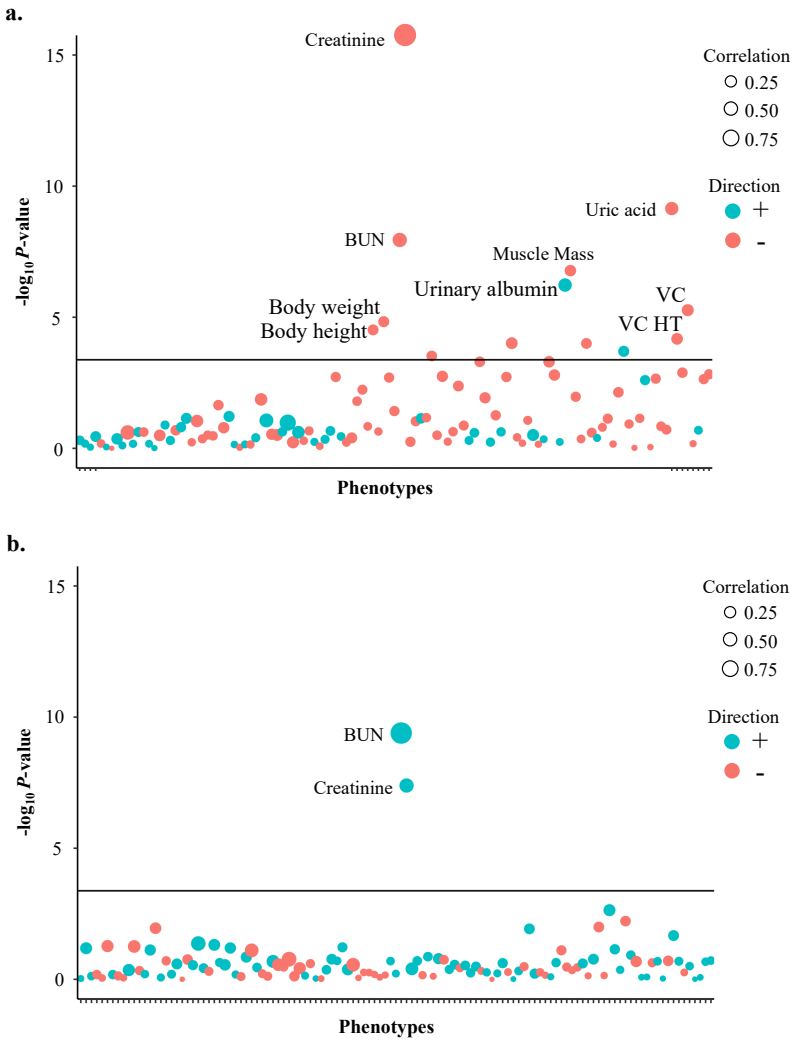

a
r = -0.86
Kidney function relevance
Likely
Inconclusive
Unlikely
LRP2
DNAJC16
SHROOM3
SCARNA21B
CASP9
CCDC158
PDILT
RINT(BUN) effect size in BBJ and TWB discovery datasets
0.05
0.00
-0.05
-0.08
-0.04
0.00
0.04
0.08
RINT(eGFR) effect size in BBJ and TWB discovery datasets


b
r = -0.94
CKD validation
Yes
Inconclusive
No
NRG4
SLC22A2
FBXO22
SLC6A13
KNG1
STC1
LRP2
TCEA3
HNRNPR
PAX8
HCRTR2
F12
DCDC1
BCAS3
LOC285889
WDR72
PDILT
UBE2Q2
TINAG
Odds ratio of CKD in CMUH-CRDR
1.02
1.00
0.98
0.96
-0.06
-0.03
0.00
0.03
0.06
RINT(eGFR) effect size in BBJ and TWB discovery datasets


c
Odds ratio of ESKD in CMUH-CRDR
1.02
1.01
1.00
0.99
0.98
-0.08
-0.04
0.00
0.04
0.08
RINT(eGFR) effect size in BBJ and TWB discovery datasets
r = -0.96
ESKD validation
Yes
Inconclusive
No
TMEM60
RSBN1L
LINC00243
STC1
TCEA3
HCG20
KNG1
TSBP1-AS1
DNA2
ZZZ3
DCDC1
HLA-DQA1
F12
SLC22A2
USP33
PHTF2
TINAG

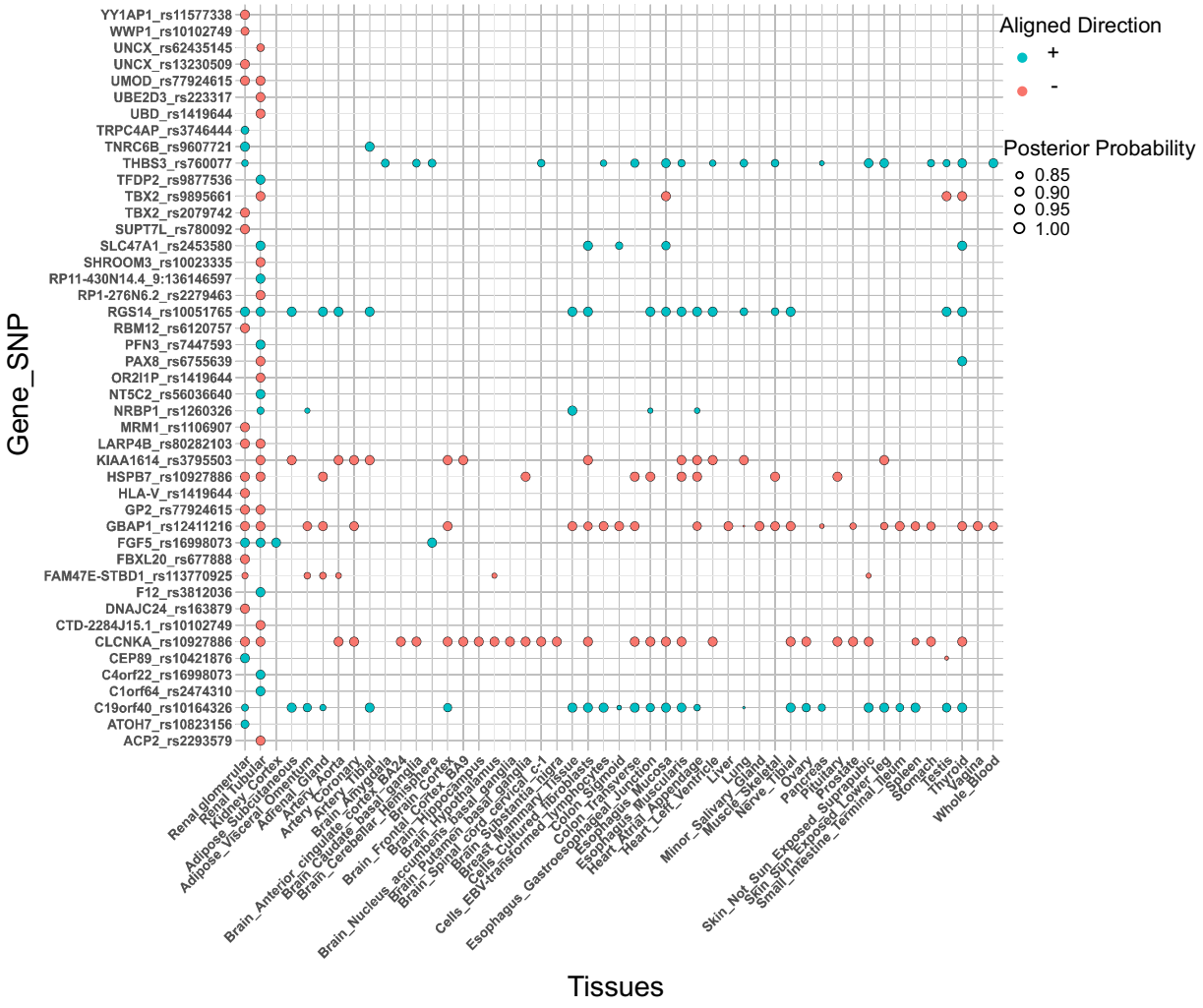
Supplementary Figure 3. Genetic correlation of meta-GWASs from eGFR and BUN with other phenotypes from Taiwan Biobank based on GWAS summary statistics. (a) The strongest genetic correlations (r_g) observed between eGFR and other phenotypes, except for serum creatinine, were those between eGFR and BUN ($r_g = -0.30$; $p = 1.13 \times 10^{-8}$), urinary albumin ($r_g = 0.25$; $p = 5.92 \times 10^{-7}$), uric acid ($r_g = -0.23$; $p = 7.15 \times 10^{-10}$), and muscle mass ($r_g = -0.14$; $p = 1.67 \times 10^{-7}$). **(b)** For BUN, the strongest genetic correlation observed was that with serum creatinine ($r_g = 0.28$; $p = 4.06 \times 10^{-8}$). Red dots represent negative associations, while blue dots represent positive associations. Further details about phenotypes are provided in the Supplementary Data 8. The horizontal line represents a p value of 4.2×10^{-4}. All statistical tests employed two-sided p values.

***Abbreviations*:** BUN, blood urea nitrogen; VC, Vital capacity; VC_HT, vital capacity/height.



Supplementary Figure 4. Colocalization of eGFR association signals with gene expression in kidney tissues. *cis*-eQTLs of tissue-specific genes were colocalized with regions surrounding eGFR-associated lead SNPs (±100 kb). Genes with SNPs in kidney tissues that have a PP of >0.80 were highlighted. The size of the circles was relative to the PP of SNPs in the colocalization analysis. Blue dots indicate that increased gene expression is correlated with a higher eGFR, and red dots indicates that increased gene expression is correlated with a lower eGFR. Further details are provided in Supplementary Data 14.

***Abbreviations*:** eGFR, estimated glomerular filtration rate; eQTL, expression quantitative trait loci; PP, posterior probability; SNP, single nucleotide polymorphism.



Supplementary Figure 5. The sex-adjusted hazard ratio of CKD was evaluated based on PRS stratification in two external datasets: CMUH-CDRD (Taiwanese, n=25,345) and UKB (white British, n=260,245). A dose-response pattern was observed in the sex-adjusted hazard ratio (HR) for CKD across patients with a PRS_{CKD} of 2 SDs above the mean and a PRS_{CKD} of within 2 SDs of the mean, compared with those with a PRS_{CKD} of 2 SDs below the mean, with corresponding HRs of 2.3 (95% CI = 1.7-3.0) and 1.6 (95% CI = 1.3-2.1) in CMHU-CRDR, and corresponding HRs of 3.2 (95% CI = 2.6-4.0) and 1.9 (95% CI = 1.6-2.3) in the White British dataset, UKB.

***Abbreviations*:** CI, confidence interval; CKD, chronic kidney disease; CMUH-CRDR, China Medial University Hospital-Clinical Research Data Repository; HR, hazard ratio; PRS, polygenic risk score; SD, standard deviation; UKB, UK Biobank.