

# Genome-wide meta-analysis and omics integration identifies novel genes associated with diabetic kidney disease

Niina Sandholm\*, Joanne B Cole\*, Viji Nair, Xin Sheng, Hongbo Liu, Emma Ahlqvist, Natalie van Zuydam, Emma Dahlström, Damian Fermin, Laura J Smyth, Rany M Salem, Carol Forsblom, Erkka Valo, Valma Harjutsalo, Eoin P Brennan, Gareth McKay, Darrell Andrews, Ross Doyle, Helen C Looker, Robert G Nelson, Colin Palmer, Amy Jayne McKnight, Catherine Godson, Alexander P Maxwell, Leif Groop, Mark I McCarthy, Matthias Kretzler, Katalin Susztak, Joel N Hirschhorn, Jose C Florez, Per-Henrik Groop, for the GENIE Consortium.

## Supplemental Material

### Table of Contents

Supplemental Table 1: A total of ten case – control definitions for diabetic kidney disease (DKD; used as a general term to describe renal complications in diabetes) .....	3
Supplemental Table 2: Number of individuals in each cohort for each phenotypic comparison .....	4
Supplemental Table 3: Key characteristics of the genotyping and statistical analyses in DNCRI and SUMMIT cohorts.....	5
Supplemental Table 4: Association details for the novel genome-wide significant locus rs72831309 in <i>TENM2</i> .....	6
Supplemental Table 5: Kidney eQTL associations with $p < 1 \times 10^{-4}$ in tubular or glomerular eQTL data, or in the kidney eQTL meta-analysis for the lead SNPs. ....	7
Supplemental Table 6: TWAS results with $p < 1 \times 10^{-4}$ .....	8
Supplemental Table 7: Highly correlated reconstituted gene-sets that make-up the “basement membrane” meta-gene set derived in Marouli et al.....	9
Supplemental Table 8: Significant kidney mQTL associations ( $p < 1.46 \times 10^{-11}$ Bonferroni-adjusted genome-wide significance; $1 \times 10^{-7}$ suggestive significance) for lead loci. ....	10
Supplemental Table 9: Correlation between glomerular and tubular gene expression and glomerulosclerosis, fibrosis, and eGFR in nephrectomy samples. ....	11
Supplemental Table 10: Gene centric summary of the lead genes .....	12
Supplemental Table 11: Mendelian Randomization (MR) results for DKD (All vs Ctrl).....	14
Supplemental Table 12. Egger intercepts for Mendelian Randomization analyses on DKD (all vs. Ctrl phenotype).....	15
Supplemental Figure 1: Manhattan and QQ-plots of the ten DKD GWAS meta-analysis results. ....	16
Supplemental Figure 2: Regional association plots for the GWAS meta-analysis lead loci (A-K). ....	19
Supplemental Figure 3: Regional association plot for the <i>COL4A3</i> gene region associated with Severe DKD, indicating a secondary association peak at chr2:228259302 (rs6436688, effect allele (A) frequency 56%, OR = 1.13 (95% confidence interval 1.08 – 1.19), p-value $1.79 \times 10^{-7}$ ).....	22
Supplemental Figure 4: Regional association plots for the gene-level analysis results from MAGMA and PASCAL analysis. ....	23
Supplemental Figure 5: Gene prioritization for the <i>COL4A3</i> gene at lead SNP rs5570367 associated with Severe DKD using multiple intersecting gene prioritization approaches (PoPS, nearest gene, and MAGMA). ....	25
Supplemental Figure 6: Tubular and glomerular gene expression of the lead genes correlate with multiple morphological and pathological renal parameters. ....	26

Supplemental Figure 7: Genetic correlation between DKD and related traits based on LD score regression. ....27

Supplemental Figure 8: Mendelian Randomization scatter plots for SNP effects for the metabolic traits vs. DKD (All vs. Ctrl). ....28

Supplemental Figure 9: rs1260634 intronic in the *ALLC* gene affects the predicted binding motifs for KLF12, KLF4, and SP8 (top to bottom). ....29

**Supplemental Table 1: A total of ten case – control definitions for diabetic kidney disease (DKD; used as a general term to describe renal complications in diabetes)**

Phenotype	Cases	Controls	Note
<b>All vs. Ctrl</b>	microalbuminuria or macroalbuminuria or ESRD	Normal AER	Phenotype abbreviated as "DN" (diabetic nephropathy) in SUMMIT
<b>Severe DKD</b>	macroalbuminuria or ESRD	Normal AER	Phenotype abbreviated as "MACRO" in SUMMIT; as "DN" in DNCRI
<b>Micro</b>	microalbuminuria	Normal AER	
<b>Macro</b>	macroalbuminuria	Normal AER	Not analyzed in SUMMIT
<b>ESRD</b>	ESRD	Normal AER	
<b>ESRD vs. All</b>	ESRD	no ESRD	
<b>ESRD vs. macro</b>	ESRD	macroalbuminuria	Not analyzed in SUMMIT
<b>CKD</b>	eGFR < 60 ml/min/1.73m <sup>2</sup>	eGFR ≥ 60 ml/min/1.73m <sup>2</sup>	
<b>CKD extremes</b>	ESRD or eGFR < 15 ml/min/1.73m <sup>2</sup>	eGFR ≥ 60 ml/min/1.73m <sup>2</sup>	Not analyzed in SUMMIT
<b>CKD-DKD</b>	ESRD, or eGFR < 60 ml/min/1.73m <sup>2</sup> AND microalbuminuria or macroalbuminuria	normal AER and eGFR ≥ 60 ml/min/1.73m <sup>2</sup>	

Normal AER: AER <30mg/24h, or equivalent

Microalbuminuria: 30 mg/24h ≤ AER < 300 mg/24h, or equivalent

Macroalbuminuria: AER ≥ 300 mg/24h, or equivalent

ESRD: End-stage renal disease, dialysis or renal transplant (or eGFR < 15 ml/min/1.73m<sup>2</sup> in SUMMIT)

**Supplemental Table 2: Number of individuals in each cohort for each phenotypic comparison**

Study	Severe DKD			Macro			ESRD			ESRD vs. All			ESRD vs. Macro			All vs. Ctrl			Micro			CKD			CKD extremes			CKD-DKD		
DNCRI	Cases	Ctrls	Total	Cases	Ctrls	Total	Cases	Ctrls	Total	Cases	Ctrls	Total	Cases	Ctrls	Total	Cases	Ctrls	Total	Cases	Ctrls	Total	Cases	Ctrls	Total	Cases	Ctrls	Total	Cases	Ctrls	Total
Austria	6	71	77	4	71	75	2	71	73	2	88	90	2	4	6	19	71	90	13	71	84	11	80	91	2	80	82	6	66	72
CACTI	35	422	457	29	422	451	6	422	428	6	503	509	6	29	35	87	422	509	52	422	474	45	477	522	6	477	483	24	407	431
EDIC	134	193	327	75	193	268	59	193	252	59	359	418	59	75	134	225	193	418	91	193	284	130	288	418	59	288	347	89	180	269
EDC	84	1016	1100	61	1016	1077	23	1016	1039	23	1282	1305	23	61	84	289	1016	1305	205	1016	1221	80	1218	1298	23	1218	1241	47	999	1046
FinnDiane	1371	2240	3611	535	2268	2803	854	2265	3119	854	3418	4272	854	517	1371	2069	2202	4271	719	2257	2976	1226	3038	4264	838	3069	3907	993	2066	3059
FRANCE	332	627	959	181	625	806	151	627	778	151	920	1071	151	181	332	448	627	1075	124	625	749	281	740	1021	159	740	899	225	578	803
GWU GOKIND	290	311	601	29	311	340	261	311	572	261	340	601	261	29	290	290	311	601	0	311	311	273	325	598	261	325	586	269	309	578
ITALY	180	161	341	38	161	199	142	161	303	142	201	343	142	38	180	180	163	343	0	161	161	157	172	329	148	175	323	155	154	309
JOSLIN	719	1082	1801	475	1082	1557	244	1082	1326	244	2027	2271	244	475	719	1189	1082	2271	470	1082	1552	533	1574	2107	262	1695	1957	402	1013	1415
LatDiane	25	80	105	18	80	98	7	80	87	7	131	138	7	18	25	58	80	138	33	80	113	16	109	125	7	109	116	9	78	87
LitDiane	19	39	58	9	39	48	10	39	49	10	69	79	10	9	19	40	39	79	21	39	60	21	50	71	10	50	60	16	36	52
RomDiane	98	89	187	70	89	159	28	89	117	28	207	235	28	70	98	146	89	235	48	89	137	53	167	220	28	167	195	39	87	126
Scotland	195	3962	4157	144	3984	4128	57	3962	4019	57	4632	4689	57	138	195	727	3962	4689	540	3984	4524	404	4712	5116	80	4712	4792	90	4450	4540
STENO	488	414	902	469	414	883	19	414	433	19	897	916	19	470	489	489	427	916	0	414	414	200	690	890	28	690	718	106	398	504
SWEDEN	51	346	397	35	346	381	20	346	366	20	497	517	20	32	52	51	346	397	85	346	431	42	287	329	20	287	307	21	252	273
UK ROI	704	730	1434	466	730	1196	200	730	930	200	1196	1396	200	466	666	704	730	1434	0	730	730	587	513	1100	200	513	713	266	433	699
WESDR	217	293	510	113	293	406	104	293	397	104	452	556	104	113	217	263	293	556	46	293	339	207	398	605	104	398	502	140	260	400
<b>Total DNCRI</b>	<b>4,948</b>	<b>12,076</b>	<b>17,024</b>	<b>2,751</b>	<b>12,124</b>	<b>14,875</b>	<b>2,187</b>	<b>12,101</b>	<b>14,288</b>	<b>2,187</b>	<b>17,219</b>	<b>19,406</b>	<b>2,187</b>	<b>2,725</b>	<b>4,912</b>	<b>7,274</b>	<b>12,053</b>	<b>19,327</b>	<b>2,447</b>	<b>12,113</b>	<b>14,560</b>	<b>4,266</b>	<b>14,838</b>	<b>19,104</b>	<b>2,235</b>	<b>14,993</b>	<b>17,228</b>	<b>2,897</b>	<b>11,766</b>	<b>14,663</b>
<b>SUMMIT T1D</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>
Eurodiab	203	491	694				84	491	575	84	705	789				298	491	789	95	491	586	113	467	580				210	357	567
NFS-ORPS	47	199	246													197	199	396	150	199	349									
SDR	168	292	460	85	277	362	75	294	369	75	529	604	57	85	142	266	290	556	98	290	388	163	365	528	57	349	406	118	239	357
<b>Total SUMMIT T1D</b>	<b>418</b>	<b>982</b>	<b>1,400</b>	<b>85</b>	<b>277</b>	<b>362</b>	<b>159</b>	<b>785</b>	<b>944</b>	<b>159</b>	<b>1,234</b>	<b>1,393</b>	<b>57</b>	<b>85</b>	<b>142</b>	<b>761</b>	<b>980</b>	<b>1,741</b>	<b>343</b>	<b>980</b>	<b>1,323</b>	<b>276</b>	<b>832</b>	<b>1,108</b>	<b>57</b>	<b>349</b>	<b>406</b>	<b>328</b>	<b>596</b>	<b>924</b>
<b>SUMMIT T2D</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>	<b>Cases</b>	<b>Ctrls</b>	<b>Total</b>
GoDARTS Affy	218	816	1,034	138	816	954	80	816	896	48	1,491	1,539	80	138	218	885	816	1,701	667	816	1,483	1,025	1,553	2,578	80	1,101	1,181	168	716	884
GoDARTS Illumina	179	680	859	130	675	805	48	680	728	80	1,621	1,701	48	130	178	859	680	1,539	680	680	1,360	972	513	1,485	48	1,092	1,140	120	587	707
MNI	66	165	231													188	165	353	122	162	284									
SDR	713	580	1,292	424	556	980	243	580	823	243	1,359	1,602	268	424	692	1,250	580	1,830	520	580	1,100	997	666	1,663	240	628	868	609	307	916
Steno	163	131	294													163	131	294				100	174	274						
<b>Total SUMMIT T2D</b>	<b>1,339</b>	<b>2,372</b>	<b>3,710</b>	<b>692</b>	<b>2,047</b>	<b>2,739</b>	<b>371</b>	<b>2,076</b>	<b>2,447</b>	<b>371</b>	<b>4,471</b>	<b>4,842</b>	<b>396</b>	<b>692</b>	<b>1,088</b>	<b>3,345</b>	<b>2,372</b>	<b>5,717</b>	<b>1,989</b>	<b>2,238</b>	<b>4,227</b>	<b>3,094</b>	<b>2,906</b>	<b>6,000</b>	<b>368</b>	<b>2,821</b>	<b>3,189</b>	<b>897</b>	<b>1,610</b>	<b>2,507</b>
<b>Total ALL</b>	<b>6,705</b>	<b>15,430</b>	<b>22,134</b>	<b>3,528</b>	<b>14,448</b>	<b>17,976</b>	<b>2,717</b>	<b>14,962</b>	<b>17,679</b>	<b>2,717</b>	<b>22,924</b>	<b>25,641</b>	<b>2,640</b>	<b>3,502</b>	<b>6,142</b>	<b>11,380</b>	<b>15,405</b>	<b>26,785</b>	<b>4,779</b>	<b>15,331</b>	<b>20,110</b>	<b>7,636</b>	<b>18,576</b>	<b>26,212</b>	<b>2,660</b>	<b>18,163</b>	<b>20,823</b>	<b>4,122</b>	<b>13,972</b>	<b>18,094</b>

Number of samples in the SUMMIT T1D and T2D cohorts represents those individuals included in the original analysis, containing related individuals; The current meta-analysis was based on effect size estimates from SUMMIT which were derived after excluding related individuals (see Supplemental Table 3 analysis method).

**Supplemental Table 3: Key characteristics of the genotyping and statistical analyses in DNCRI and SUMMIT cohorts.**

	DNCRI	SUMMIT T1D and T2D
N studies	17	3 (T1D), 5 (T2D)
N samples (max)	19,406	1,741 (T1D), 6,000 (T2D)
Genotyping chips	HumanCoreExome Bead arrays 12-1.0, 12-1.1, and 24-1.0	Illumina Omni express array, Affymetrix SNP 6.0 array, Illumina 610Quad assay
Imputation reference panel	1000Genomes Phase 3	1000Genomes Phase 1
Imputation software	Minimac3/Minimac3-omp (version 1.0.14)	Prephasing with SHAPE-IT v2; Imputation with IMPUTEv2
Covariates	Age, sex, diabetes duration, genetic principal components, study specific covariates (e.g. site or genotyping batch)	Age, gender, duration of diabetes, genetic principal components
Main post-analysis SNP QC filters	<ul style="list-style-type: none"> <li>•Imputation quality score: <math>INFO \geq 0.3</math></li> <li>•Minor allele count <math>\geq 10</math> in cases and in controls</li> <li>•Marker must be present in at least 2 studies</li> </ul>	<ul style="list-style-type: none"> <li>•Imputation quality score: <math>INFO \geq 0.4</math></li> <li>•Minor allele count <math>\geq 10</math> in cases and in controls</li> <li>•MAF <math>\geq 0.01</math></li> <li>•T2D: Marker must be present in at least 2 studies (not applied for the three additional phenotypes available only in SDR, GoDARTS 1 and 2)</li> </ul>
Analysis method and software	SNPtest, additive score test	SNPtest, additive score test. Note: in original studies, <i>P</i> values of association were estimated using EMMAX mixed model including related individuals, while only effect size estimates were obtained from SNPtest (excluding related individuals).
Meta-analysis	Inverse-variance fixed effects meta-analysis (METAL software)	Inverse-variance fixed effects meta-analysis (GWAMA or METAL)

**Supplemental Table 4: Association details for the novel genome-wide significant locus rs72831309 in TENM2.** In the meta-analysis, rs72831309 (chr5:166978230) minor A allele was considered the effect allele, major G as the non-effect allele.

COHORT	EAF	BETA	SE	OR	OR_L95	OR_U95	P	INFO	N
GWU_GOKIND	0.02	-0.11	0.76	0.88	0.20	3.93	0.88	0.38	578
UK_ROI	0.03	0.57	0.49	1.40	0.53	3.68	0.25	0.53	699
JOSLIN	0.03	0.33	0.36	1.15	0.57	2.33	0.36	0.51	1415
FinnDiane	0.05	0.82	0.17	1.60	1.15	2.22	1.0×10 <sup>-6</sup>	0.66	3059
EURODIAB	0.03	1.36	0.51	3.89	1.43	10.54	0.008	0.52	567
T2D meta	0.03	0.69	0.31	1.99	1.08	3.67	0.028	NA	1552
<b>Meta all</b>	<b>0.04</b>	<b>0.73</b>	<b>0.13</b>	<b>2.08</b>	1.62	2.67	<b>9.82×10<sup>-9</sup></b>		<b>8322</b>

EAF: effect allele (minor A allele) frequency. BETA: effect size estimate. SE: standard error for BETA. OR: odds ratio. OR\_L95 and OR\_U95: Lower and upper confidence intervals. P: p-value. INFO: imputation quality info metrics. N: Number of samples in the study.

**Supplemental Table 5: Kidney eQTL associations with  $p < 1 \times 10^{-4}$  in tubular or glomerular eQTL data, or in the kidney eQTL meta-analysis for the lead SNPs.** Three top SNPs were queried for each lead locus from GWAS meta-analysis, gene-level (MAGMA or PASCAL) analysis, or TWAS locus.

Tissue	SNP	CHR:POS	P eQTL	eQTL GENE	Index gene	P GWAS
Glomerular	rs28577966	4:99796005	2.13E-07	<i>ADH4</i>	<i>EIF4E</i>	1.05E-06
Glomerular	rs7664964	4:99796439	2.13E-07	<i>ADH4</i>	<i>EIF4E</i>	8.92E-07
Glomerular	rs11725932	4:99799310	2.13E-07	<i>ADH4</i>	<i>EIF4E</i>	9.97E-07
Tubular	rs59113552	6:88236233	5.19E-05	<i>SMIM8</i>	<i>AKIRIN2</i>	2.85E-05
Kidney meta	rs786959	9:115429626	4.59E-07	<i>SNX30</i>	<i>INIP</i>	9.913E-07
Kidney meta	rs6011746	20:61964452	5.75E-05	<i>CHRNA4</i>	<i>COL20A1</i>	1.90E-06

Tissue: kidney eQTL meta-analysis, or glomerular or tubule compartment-specific expression.

**Supplemental Table 6: TWAS results with  $p < 1 \times 10^{-4}$ .** P-values  $< 4.1 \times 10^{-6}$  are significant after correction for multiple testing.

Tissue	Pheno	gene	gene_name	TWAS association			Prediction performance				
				Effect	p-value	var_g	r2	p-value	q-value	n_snps_used	n_snps_in_model
tub	Severe DKD	ENSG00000135334.8	<i>AKIRIN2</i>	0.308	<b>1.11E-06</b>	0.092	0.05	0.013	0.012	39	42
tub	Macro	ENSG00000135334.8	<i>AKIRIN2</i>	0.383	<b>1.70E-06</b>	0.092	0.05	0.013	0.012	39	42
glom	Micro	ENSG00000268208.1	<i>AC008372.1</i>	-7.489	1.59E-05	0.000	0.03	0.083	0.044	1	1
glom	Macro	ENSG00000228696.4	<i>ARL17B</i>	-0.195	2.06E-05	0.331	0.51	1.28E-19	1.77E-18	54	65
glom	Micro	ENSG00000138028.10	<i>CGREF1</i>	-0.228	2.15E-05	0.131	0.14	2.08E-05	3.07E-05	25	26
glom	CKD	ENSG00000078804.8	<i>TP53INP2</i>	0.499	2.73E-05	0.068	0.03	0.080	0.042	34	38
glom	Macro	ENSG00000227057.3	<i>WDR46</i>	-0.583	4.05E-05	0.024	0.03	0.085	0.044	10	10
tub	Severe DKD	ENSG00000205269.4	<i>TMEM170B</i>	0.192	4.62E-05	0.133	0.06	0.007	0.007	90	95
tub	CKD	ENSG00000188283.7	<i>ZNF383</i>	0.327	4.72E-05	0.046	0.05	0.012	0.011	15	16
tub	CKD	ENSG00000075413.13	<i>MARK3</i>	-0.258	5.05E-05	0.085	0.05	0.013	0.012	102	104
tub	Macro	ENSG00000162836.7	<i>ACP6</i>	-0.146	6.44E-05	0.377	0.47	4.19E-18	7.16E-17	30	31

Tissue: tub(ular) or glom(erular). TWAS association: Effect: association effect size for the gene. P-value: p-value for the TWAS association. var\_g: variance of the gene expression. Prediction performance r2, p-value and q-value: statistics for tissue model's correlation to gene's measured transcriptome; n\_snps\_used: number of SNPs from GWAS that were used in the analysis; n\_snps\_in\_model: number of SNPs in the model (i.e. in the transcriptomics data)



**Supplemental Table 7: Highly correlated reconstituted gene-sets that make-up the “basement membrane” meta-gene set derived in Marouli et al.** The “Genes prioritized” column contains all genes prioritized in MAGMA by one of these 26 gene-sets, including the “FBLN2 PPI subnetwork” gene set that prioritized *COL4A3* for Severe DKD.

GENE-SET ID	GENE-SET DESCRIPTION	GENES PRIORITIZED
MP:0003044	impaired basement membrane formation	
<b>ENSG00000163520</b>	<b>FBLN2 PPI subnetwork</b>	<i>ADAMTS15, CD248, <b>COL4A3</b>, NID2</i>
MP:0004272	abnormal basement membrane morphology	
GO:0043256	laminin complex	
ENSG00000130702	LAMA5 PPI subnetwork	
ENSG00000132561	MATN2 PPI subnetwork	
ENSG00000091136	LAMB1 PPI subnetwork	
ENSG00000116962	NID1 PPI subnetwork	
ENSG00000135862	LAMC1 PPI subnetwork	
ENSG00000168487	BMP1 PPI subnetwork	
GO:0034446	substrate adhesion-dependent cell spreading	
ENSG00000125810	CD93 PPI subnetwork	
ENSG00000134871	COL4A2 PPI subnetwork	<i>ABCC9, CCDC102B, COL18A1, CSPG4, CTHRC1, FN1, IGFBP3, LAMC1, LOXL1, OLFML2B, TGFBI</i>
ENSG00000188153	COL4A5 PPI subnetwork	<i>ABCA9, ADAMTS5, AEBP1, ART3, BICC1, C3, C7, COL15A1, COL1A1, COL1A2, COL3A1, <b>COL4A3</b>, COL5A2, COL6A2, ENSG00000259134, ENSG00000259284, FBLN5, FBN1, FIBIN, FNDC1, FSTL1, GALNTL4, GRB14, IGFBP6, LAMA2, LAMB1, LINC00312, LOX, MMP2, NOV, OLFML1, POSTN, SCN7A, SERPING1, SLIT3, SPARC, THBS2, VGLL3, WDR72</i>
ENSG00000112773	FAM46A PPI subnetwork	
ENSG00000100985	MMP9 PPI subnetwork	
ENSG00000110492	MDK PPI subnetwork	
ENSG00000101680	LAMA1 PPI subnetwork	
GO:0043236	laminin binding	
ENSG00000187498	COL4A1 PPI subnetwork	<i>ACTA2, COL14A1, COL4A2, CTHRC1, ENPEP, LHFP, LOXL2, TGFBI</i>
GO:0005605	basal lamina	
ENSG00000213949	ITGA1 PPI subnetwork	
ENSG00000114270	COL7A1 PPI subnetwork	
GO:0050840	extracellular matrix binding	
ENSG00000081052	COL4A4 PPI subnetwork	<i>ABCC9, ACTA2, ADAMTS1, ADAMTS5, ASPN, C1R, C1S, C7, CCDC80, COL4A1, COL5A1, COL6A1, DCN, EFEMP1, FBLN1, FKBP7, IGFBP3, LGALS1, LOX, LUM, MGP, NID2, PID1, PXDN, SCN7A, SERPINF1, SPARCL1, VCAN</i>
GO:0005604	basement membrane	<i>APLN, COL12A1, CTGF, HTRA1, ITGB4, ITGB6, MCAM, PRSS23</i>

**Supplemental Table 8: Significant kidney mQTL associations ( $p < 1.46 \times 10^{-11}$  Bonferroni-adjusted genome-wide significance;  $1 \times 10^{-7}$  suggestive significance) for lead loci.** Three most significant SNPs were queried at each associated loci from single variant, gene-level, and transcriptome-wide association study (TWAS).

Chr	RSID	SNP_Pos	CpG	CpG_Start	P mQTL	Gene	P DKD/eGFR
19	rs668933	34704936	cg14143166	34716204	1.94E-28	<i>LSM14A</i>	0.03 (DKD)
19	rs1260634	34701331	cg14143166	34716204	2.09E-28	<i>LSM14A</i>	0.03 (DKD)
19	rs535440	34694581	cg14143166	34716204	2.12E-28	<i>LSM14A</i>	0.03 (DKD)
13	rs12428319	36542599	cg21746263	36562319	6.81E-22	<i>DCLK1</i>	
13	rs61948262	36533891	cg21746263	36562319	4.19E-21	<i>DCLK1</i>	
20	rs117255010	61953366	cg20706388	61958549	1.79E-12	<i>COL20A1</i>	
20	rs6011746	61964452	cg20706388	61958549	2.14E-12	<i>COL20A1</i>	
20	rs4809528	61943661	cg20706388	61958549	2.55E-12	<i>COL20A1</i>	
20	rs74397198	61950071	cg20706388	61958549	2.88E-12	<i>COL20A1</i>	
20	rs143391037	61971717	cg20706388	61958549	4.76E-12	<i>COL20A1</i>	
2	rs6436131	220151858	cg06895971	220147671	9.20E-12	<i>PTPRN</i>	
4	rs7664964	99796439	cg25974308	99852386	1.10E-11	<i>EIF4E</i>	0.041 (eGFR slope)
4	rs28577966	99796005	cg25974308	99852386	1.10E-11	<i>EIF4E</i>	0.041 (eGFR slope)
4	rs11725932	99799310	cg25974308	99852386	1.10E-11	<i>EIF4E</i>	0.041 (eGFR slope)
6	rs34472900	88405040	cg00551398	88298473	1.12E-11	<i>AKIRIN2</i>	
6	rs151077971	88405605	cg00551398	88298473	1.13E-11	<i>AKIRIN2</i>	
6	rs59113552	88236233	cg00551398	88298473	1.58E-11	<i>AKIRIN2</i>	
9	rs786975	115451231	cg13293976	115516494	2.20E-11	<i>INIP</i>	0.012 (eGFR slope)
2	rs4674377	220201272	cg14891200	220197663	2.72E-10	<i>RESP18</i>	
2	rs2090163	220178435	cg14891200	220197663	1.06E-09	<i>PTPRN</i>	
6	rs59113552	88236233	cg10313604	88493367	3.18E-09	<i>AKIRIN2</i>	
6	rs151077971	88405605	cg10313604	88493367	7.06E-09	<i>AKIRIN2</i>	
6	rs34472900	88405040	cg10313604	88493367	7.11E-09	<i>AKIRIN2</i>	
6	rs151077971	88405605	cg05834092	87792915	7.82E-09	<i>AKIRIN2</i>	
6	rs34472900	88405040	cg05834092	87792915	7.83E-09	<i>AKIRIN2</i>	
19	rs535440	34694581	cg21245903	34711622	9.97E-09	<i>LSM14A</i>	
19	rs1260634	34701331	cg21245903	34711622	1.00E-08	<i>LSM14A</i>	
19	rs668933	34704936	cg21245903	34711622	1.04E-08	<i>LSM14A</i>	
6	rs59113552	88236233	cg05834092	87792915	1.63E-08	<i>AKIRIN2</i>	
6	rs151077971	88405605	cg15059496	88185750	3.76E-08	<i>AKIRIN2</i>	
6	rs34472900	88405040	cg15059496	88185750	3.76E-08	<i>AKIRIN2</i>	
6	rs59113552	88236233	cg20648632	88182160	4.22E-08	<i>AKIRIN2</i>	
2	rs2090163	220178435	cg19020434	220199207	6.40E-08	<i>PTPRN</i>	
2	rs4674377	220201272	cg06895971	220147671	6.77E-08	<i>RESP18</i>	
6	rs59113552	88236233	cg15059496	88185750	7.55E-08	<i>AKIRIN2</i>	
2	rs4674377	220201272	cg19020434	220199207	7.84E-08	<i>RESP18</i>	
19	rs535440	34694581	cg01663383	34676533	8.98E-08	<i>LSM14A</i>	
19	rs1260634	34701331	cg01663383	34676533	9.01E-08	<i>LSM14A</i>	
19	rs668933	34704936	cg01663383	34676533	9.24E-08	<i>LSM14A</i>	

Gene: CpG site annotated gene. P DKD/eGFR: P-value for association between blood methylation at the CpG site and DKD (UK-ROI+FinnDiane EWAS) or with eGFR slope (CRIC EWAS).

**Supplemental Table 9: Correlation between glomerular and tubular gene expression and glomerulosclerosis, fibrosis, and eGFR in nephrectomy samples.** All associations with  $p < 0.05$  are shown, significant associations ( $p < 2.2 \times 10^{-4}$ , corrected for 29 genes  $\times$  2 tissue compartments  $\times$  4 phenotypes) are in bold.

Gene	Tissue	eGFR		Glomerulosclerosis		Fibrosis		Group comparison	
		r	p	r	p	r	p	Direction	p
<i>ALLC</i>	glom								
<i>ALLC</i>	tub	0.26	<b>6.9E-08</b>			-0.50	<b>2.0E-16</b>	lowest in DKD	<b>8.3E-05</b>
<i>COLEC11</i>	glom	0.24	<b>1.0E-05</b>	-0.27	<b>1.8E-06</b>			lowest in DKD	<b>1.8E-05</b>
<i>COLEC11</i>	tub	0.21	<b>1.0E-05</b>			-0.47	<b>2.0E-16</b>	lowest in DKD	1.37E-03
<i>PLEKHA7</i>	glom	0.14	9.5E-03					lowest in DKD	<b>4.3E-06</b>
<i>PLEKHA7</i>	tub	0.30	<b>1.3E-10</b>			-0.49	<b>2.0E-16</b>	lowest in DKD	<b>1.3E-05</b>
<i>SNX30</i>	glom	0.24	<b>1.2E-05</b>	-0.22	<b>8.0E-05</b>			lowest for DKD	<b>5.5E-05</b>
<i>SNX30</i>	tub	0.35	<b>5.8E-14</b>			-0.56	<b>2.0E-16</b>		<b>1.2E-06</b>
<i>DCLK1</i>	glom								
<i>DCLK1</i>	tub	-0.15	1.48E-03			0.39	<b>7.4E-16</b>	Highest in DKD	<b>2.17E-04</b>
<i>TENM2</i>	glom	0.13	0.02	-0.18	1.7E-03				
<i>TENM2</i>	tub	0.27	<b>1.6E-08</b>			-0.29	<b>2.0E-09</b>	lowest in DKD	6.6E-04
<i>COL4A3</i>	glom	0.11	0.05	-0.16	4.8E-03				
<i>COL4A3</i>	tub					0.29	<b>3.2E-09</b>		
<i>ZNF3</i>	glom			0.12	0.04				
<i>ZNF3</i>	tub	-0.13	7.26E-03			0.26	<b>1.4E-07</b>		
<i>TAMM41</i>	glom	-0.11	0.04	0.16	4.1E-03				
<i>TAMM41</i>	tub	-0.20	<b>2.0E-05</b>			0.26	<b>1.5E-07</b>	highest in DKD	5.6E-03
<i>AKIRIN2</i>	glom								
<i>AKIRIN2</i>	tub					0.25	<b>2.8E-07</b>		
<i>EIF4E</i>	glom	-0.12	0.03					highest in DKD	<b>2.4E-06</b>
<i>EIF4E</i>	tub					-0.18	<b>1.9E-04</b>		
<i>LSM14A</i>	glom								
<i>LSM14A</i>	tub	0.22	<b>2.9E-06</b>			-0.13	0.01		
<i>INIP</i>	glom			-0.15	9.3E-03				
<i>INIP</i>	tub	0.22	<b>5.5E-06</b>			-0.21	<b>2.2E-05</b>		
<i>MFF</i>	glom								
<i>MFF</i>	tub					-0.21	<b>2.1E-05</b>		
<i>MBLAC1</i>	glom			-0.11	0.05			lowest in DKD	2.8E-03
<i>MBLAC1</i>	tub					-0.17	5.2E-04		
<i>STAC</i>	glom			-0.17	3.0E-03				
<i>STAC</i>	tub								

r: Pearson correlation coefficient between the phenotype and  $\log_2$ (Fragments per kilobases of transcript per 1 million mapped reads [FPKM]) of gene expression in glomeruli/tubules. p: p-value. Group comparison: ANOVA test for group comparison (Controls, chronic kidney disease [CKD], diabetic kidney disease [DKD], diabetes mellitus [DM] (without DKD), hypertension [HTN]) vs.  $\log_2$ (FPKM) gene expression.

**Supplemental Table 10: Gene centric summary of the lead genes**

GENE	Pheno	Indication	PoPS	P min	P min	lead SNP PCHIC Max score	Glom/ Tub eQTL	P min	Nephrectomy Gene expression vs. phenotype correlations								NephroSeq DN vs. healthy				Pima BX1 correlations	Pima BX2 correlations		
									eGFR		Glomerulosclerosis		fibrosis		Group comparison		Woroniciecka DN vs healthy		Ju DN vs. healthy					
									r	p	r	p	r	p	direction	p	P	FC	P	FC				
TENM2	CKD+DKD	Underlying lead SNP	Yes			8.61	glom		0.1	0.022	-0.2	0.002												
							tub		0.3	1.6E-8			-0.3	2.0E-9	lowest in DKD	6.6E-4								
DCLK1	ESRD vs. macro	Gene-based test				6.81E-22	glom																	
							tub	0.035	-0.2	0.001			0.4	7.4E-16	Highest in DKD	2.17E-4	0.003	2.09	0.001	1.32	1.2E-4	1.98	ACR: r=0.29 p=0.016; MesVol: r=0.25 p=0.041; FootProcW: r=0.71 p=0.0067; Fibrosis: r=0.52 p=0.0003;	GlomWidth: r=0.32 p=0.024; MesVol: r=0.29 p=0.038; ACR: r=0.28 p=0.04; Fibrosis: r=0.36 p=0.015;
COL4A3	DKD	Missense variant	Yes			8.89	glom		0.1	0.047	-0.2	0.005					8.3E-5	-1.48	0.032	-0.43			SV: r=-0.28 p=0.024;	
							tub						0.3	3.2E-09										
COLEC11	CKD	Nearby				9.67	glom		0.2	1.0E-5	-0.3	1.8E-6			lowest in DKD	1.8E-5			0.019	-1.59				
							tub		0.2	1.0E-5			-0.5	2.0E-16	lowest in DKD	0.001			0.036	-1.47			ACR: r=-0.32 p=0.016; HbA1c: r=-0.29 p=0.034;	
ALLC	CKD	Underlying lead SNP				10.42	glom												0.04	-1.19			GlomWidth: r=0.24 p=0.054; ΔSV: r=-0.46 p=0.01; HbA1c: r=0.3 p=0.041;	MesVol: r=-0.29 p=0.04;
							tub		0.3	6.9E-8			-0.5	2.0E-16	lowest in DKD	8.3E-5	3.1E-4	-1.43						
PLEKHA7	Micro	Underlying lead SNP	Yes			9.49	glom		0.1	9.5E-3					lowest in DKD	4.3E-6								
							tub		0.3	1.3E-10			-0.5	2.0E-16	lowest in DKD	1.3E-5								
SNX30	DKD	Gene-based test; kidney eQTL				4.6E-7	glom	0.012	0.2	1.2E-05	-0.2	8.0E-5			lowest in DKD	5.5E-5								
							tub	0.002	0.4	5.8E-14			-0.6	2.0E-16	lowest in DKD	1.2E-6							eGFR: r=0.33 p=0.013;	
AKIRIN2	Severe DKD	TWAS		0.017		1.12E-11	glom																	
							tub	0.001					0.3	2.8E-7										
EIF4E	ESRD vs. macro	Gene-based test				1.10E-11	glom		-0.1	0.027					highest in DKD	2.4E-6	0.010	-1.88	0.027	-1.27			ACR: r=0.25 p=0.041;	
							tub						-0.2	1.9E-4									eGFR: r=-0.4 p=0.0028;	
MFF	Severe DKD	Gene-based test		0.032		10.19	glom	0.007									5.4E-4	-1.40					ΔMesVol: r=-0.46 p=0.011; ΔeGFR: r=0.43 p=0.017; Progr to ESRD: p=0.007;	
							tub	0.017					-0.2	2.1E-5										
ZNF3	Micro	Underlying tag SNP					glom				0.1	0.038											GlomVol: r=0.25 p=0.05; HbA1c: r=-0.25 p=0.043;	
							tub		-0.1	0.007			0.3	1.4E-7										
TAMM41	Micro	Nearby				10.65	glom		-0.1	0.037	0.2	0.004												
							tub		-0.2	2.0E-5			0.3	1.5E-7	highest in DKD	0.006							ACR: r=0.34 p=0.011;	
LSM14A	Severe DKD	Gene-based test				1.9E-28	glom	0.004									7.4E-4	-1.27	0.001	-1.27			ΔMesVol: r=-0.59 p=0.00065; ΔSV: r=0.5 p=0.0052;	
							tub	0.017	0.2	2.9E-6			-0.1	0.011										
STAC	ESRD vs. All	Underlying lead SNP				10.87	glom				-0.2	0.003					0.030	-1.53	0.013	-1.38				
							tub										0.018	-1.18						
PTPRN	CKD	Gene-based test		0.007		9.2E-12	glom										0.016	1.14	0.002	-1.16			FootProcW: r=0.29 p=0.018;	

					<i>tub</i>													
<i>INIP</i>	DKD	Gene-based test		<b>2.2E-11</b>	<i>glom</i>	0.004												0.011 -1.21 9.5E-4 -1.11 HbA1c: r=0.28 p=0.059;
					<i>tub</i>		0.2	<b>5.5E-6</b>										
<i>CNTN6</i>	ESRD	PoPS	<b>Yes</b>	0.046	<i>glom</i>													
					<i>tub</i>													
<i>MBLAC1</i>	Micro	Nearby			<i>glom</i>													
					<i>tub</i>													
<i>COL20A1</i>	CKD extremes	Gene-based test		<b>1.8E-12</b>	<i>glom</i>													
					<i>tub</i>	0.018												
<i>DCLK3</i>	ESRD vs. All	Nearby			<i>glom</i>													
					<i>tub</i>													
<i>MUC7</i>	Severe DKD	Nearby			<i>glom</i>													
					<i>tub</i>													
<i>RESP18</i>	CKD	Gene-based test		<b>2.7E-10</b>	<i>glom</i>													
					<i>tub</i>													
<i>AMTN</i>	Severe DKD	Nearby			<i>glom</i>													
					<i>tub</i>													
<i>GPR158</i>	Severe DKD	Gene-based test			<i>glom</i>													
					<i>tub</i>													
<i>LINC01266</i>	ESRD	Underlying lead SNP			<i>glom</i>													
					<i>tub</i>													
<i>PRNCR1</i>	ESRD vs. macro	Underlying lead SNP			<i>glom</i>													
					<i>tub</i>													
<i>CTB-178M22.2</i>	CKD+DKD	lead SNP kidney eQTL	0.007	<b>NA</b>	<i>glom</i>		<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>
					<i>tub</i>		<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>
<i>ADH4</i>	ESRD vs. macro	glom eQTL for EIF4E	0.008	<b>NA</b>	<i>glom</i>	<b>2.1E-7</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>
					<i>tub</i>	0.002	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>
<i>SMIM8</i>	Severe DKD	tub eQTL for AKIRIN2	1.1E-4	<b>NA</b>	<i>glom</i>		<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>
					<i>tub</i>	<b>5.2E-5</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>
<i>CHRNA4</i>	CKD extremes	kidney eQTL for COL20A1	<b>5.8E-5</b>	<b>NA</b>	<i>glom</i>	0.009	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>
					<i>tub</i>	0.007	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>	<b>NA</b>

Indication: why gene was listed as a lead gene. PoPS: was prioritized by PoPS? Kidney eQTL: minimum P-value for eQTL association between the lead SNPs and the gene in the kidney eQTL meta-analysis. Kidney mQTL: minimum P-value for kidney methylation, between the lead SNPs and the gene, as assigned in the mQTL annotation. Lead SNP PCHiC: highest score for chromatin 3D conformation capture data at the chicp.org for the GWAS meta-analysis lead loci. Glom/Tub eQTL: minimum P-value for eQTL association between the lead SNPs and the gene in glomerular/ tubular eQTL data. Nephrectomy Gene expression vs. phenotype correlations: Pearson correlation (r) and p-value for correlation between glomerular/tubular gene expression and the phenotype. NephroSeq DN vs. healthy: Fold change (FC) and p-value for differential glomerular/tubular gene expression in DN vs. healthy in the Woroniecka and Ju CKD data sets. Pima BX1/BX2 correlations: Pearson correlation coefficient r and p-value for glomerular/tubular gene expression vs. morphological parameters at the first (BX1) or second (BX2) biopsy.

**Supplemental Table 11: Mendelian Randomization (MR) results for DKD (All vs Ctrl)**

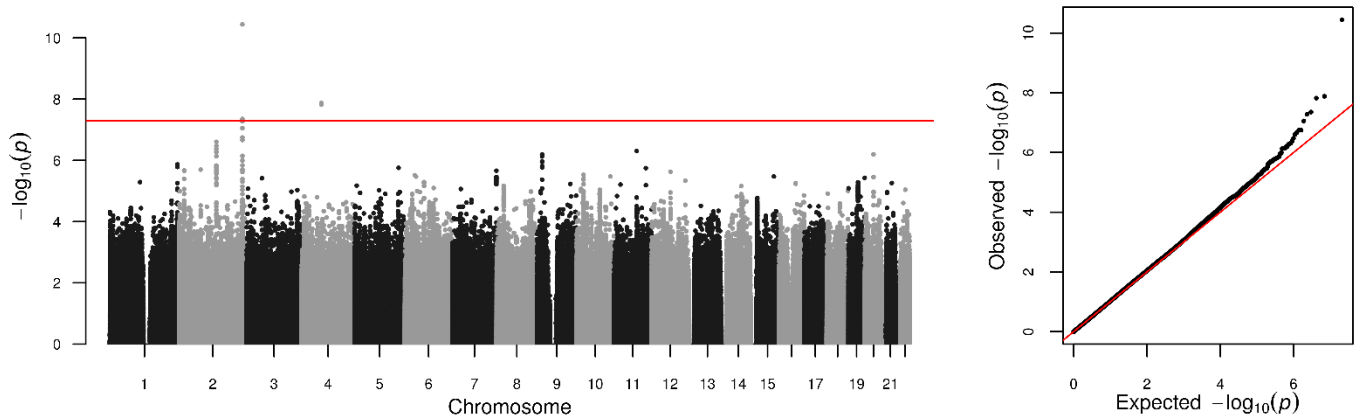
Exposure	method	nsnp	OR (95%CI)	se	p-val	Q-pval	I <sup>2</sup> (%)
<b>Body fat</b>	Inverse variance weighted	9	1.48 (0.73-2.98)	0.36	0.27	0.02	60.2
<b>Body fat</b>	Weighted median	9	1.26 (0.64-2.49)	0.35	0.50	NA	NA
<b>Body fat</b>	MR Egger	9	1.04 (0.01-98.21)	2.32	0.99	0.01	65.1
<b>Body mass index</b>	Inverse variance weighted	78	1.86 (1.55-2.23)	0.09	2.56e-11	0.50	0.00
<b>Body mass index</b>	Weighted median	78	1.76 (1.31-2.36)	0.15	1.85e-04	NA	NA
<b>Body mass index</b>	MR Egger	78	1.98 (1.28-3.08)	0.22	3.18e-03	0.48	0.36
<b>Obesity class 1</b>	Inverse variance weighted	17	1.28 (1.14-1.44)	0.06	1.90e-05	0.10	31.6
<b>Obesity class 1</b>	Weighted median	17	1.24 (1.08-1.42)	0.07	2.26e-03	NA	NA
<b>Obesity class 1</b>	MR Egger	17	1.21 (0.87-1.69)	0.17	0.28	0.08	35.2
<b>Obesity class 2</b>	Inverse variance weighted	11	1.16 (1.05-1.28)	0.05	2.83e-03	0.10	37.2
<b>Obesity class 2</b>	Weighted median	11	1.13 (1.02-1.26)	0.05	0.02	NA	NA
<b>Obesity class 2</b>	MR Egger	11	1.21 (0.89-1.65)	0.16	0.25	0.07	42.9
<b>Overweight</b>	Inverse variance weighted	14	1.47 (1.22-1.77)	0.09	4.51e-05	0.07	38.6
<b>Overweight</b>	Weighted median	14	1.31 (1.05-1.64)	0.11	0.02	NA	NA
<b>Overweight</b>	MR Egger	14	1.11 (0.59-2.08)	0.32	0.75	0.07	39.3
<b>Hip circumference</b>	Inverse variance weighted	49	1.74 (1.34-2.25)	0.13	2.68e-05	0.07	24.5
<b>Hip circumference</b>	Weighted median	49	1.89 (1.33-2.67)	0.18	3.29e-04	NA	NA
<b>Hip circumference</b>	MR Egger	49	3.45 (1.35-8.77)	0.48	0.01	0.09	22.5
<b>Waist circumference</b>	Inverse variance weighted	45	1.90 (1.49-2.42)	0.12	1.71e-07	0.28	10.0
<b>Waist circumference</b>	Weighted median	45	1.86 (1.29-2.67)	0.19	8.35e-04	NA	NA
<b>Waist circumference</b>	MR Egger	45	2.39 (1.27-4.52)	0.32	0.01	0.27	10.8
<b>Waist-to-hip ratio</b>	Inverse variance weighted	30	1.34 (0.97-1.84)	0.16	0.08	0.43	2.5
<b>Waist-to-hip ratio</b>	Weighted median	30	1.15 (0.72-1.84)	0.24	0.55	NA	NA
<b>Waist-to-hip ratio</b>	MR Egger	30	1.61 (0.37-6.95)	0.75	0.53	0.38	5.7
<b>Coronary artery disease</b>	Inverse variance weighted	61	0.99 (0.91-1.08)	0.04	0.79	0.65	0.00
<b>Coronary artery disease</b>	Weighted median	61	0.95 (0.83-1.09)	0.07	0.47	NA	NA
<b>Coronary artery disease</b>	MR Egger	61	0.92 (0.77-1.09)	0.09	0.34	0.65	0.00
<b>Type 2 diabetes</b>	Inverse variance weighted	25	1.14 (1.02-1.27)	0.05	0.02	0.06	32.5
<b>Type 2 diabetes</b>	Weighted median	25	1.11 (0.97-1.27)	0.07	0.14	NA	NA
<b>Type 2 diabetes</b>	MR Egger	25	1.20 (0.79-1.83)	0.22	0.40	0.046	35.1
<b>HDL cholesterol</b>	Inverse variance weighted	84	0.91 (0.82-1.02)	0.06	0.11	0.73	0.00
<b>HDL cholesterol</b>	Weighted median	84	1.03 (0.87-1.21)	0.09	0.77	NA	NA
<b>HDL cholesterol</b>	MR Egger	84	1.16 (0.94-1.43)	0.11	0.17	0.87	0.00
<b>Urate</b>	Inverse variance weighted	25	1.07 (0.94-1.23)	0.07	0.31	0.11	26.8
<b>Urate</b>	Weighted median	25	1.08 (0.90-1.31)	0.10	0.40	NA	NA
<b>Urate</b>	MR Egger	25	0.90 (0.71-1.15)	0.12	0.41	0.17	21.3
<b>Smoking status: Never</b>	Inverse variance weighted	77	0.54 (0.27-1.06)	0.35	0.07	0.10	17.7
<b>Smoking status: Never</b>	Weighted median	77	0.84 (0.31-2.24)	0.50	0.72	NA	NA
<b>Smoking status: Never</b>	MR Egger	77	1.69 (0.08- 35.31)	1.55	0.74	0.09	18.2
<b>Smoking status: Current</b>	Inverse variance weighted	15	0.94 (0.10-8.84)	1.15	0.96	0.95	0
<b>Smoking status: Current</b>	Weighted median	15	0.24 (0.01-4.54)	1.50	0.34	NA	NA
<b>Smoking status: Current</b>	MR Egger	15	4.11 (0.00-1.4e5)	5.34	0.80	0.93	0

**Supplemental Table 12. Egger intercepts for Mendelian Randomization analyses on DKD (all vs. Ctrl phenotype).**

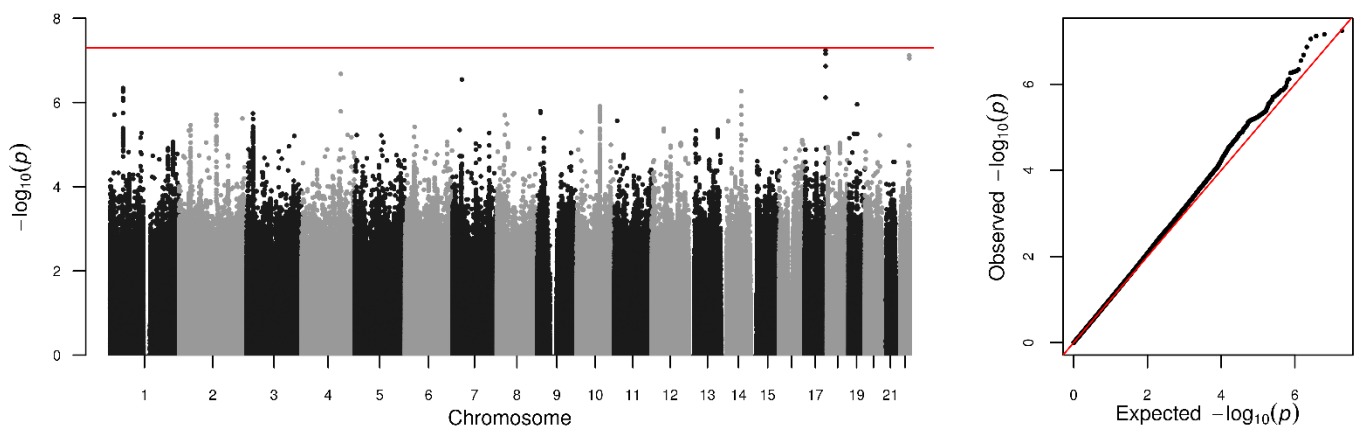
<b>Exposure</b>	<b>Egger intercept</b>	<b>Intercept SE</b>	<b>Intercept p-value</b>
<b>Body fat</b>	0.00094	0.051	0.99
<b>BMI</b>	-0.0020	0.0063	0.76
<b>Obesity class I</b>	0.0072	0.019	0.71
<b>Obesity class II</b>	-0.0083	0.027	0.77
<b>Overweight</b>	0.024	0.026	0.37
<b>Waist circumference</b>	-0.0095	0.011	0.39
<b>Hip circumference</b>	-0.0071	0.0093	0.45
<b>WHR</b>	0.0057	0.017	0.74
<b>CAD</b>	0.0086	0.0064	0.18
<b>T2D</b>	0.012	0.012	0.31
<b>HDL cholesterol</b>	-0.013	0.0047	0.0085
<b>Urate</b>	0.014	0.0090	0.12
<b>Ever smoking</b>	-0.0084	0.013	0.51
<b>Current smoking</b>	-0.0078	0.027	0.78

**Supplemental Figure 1: Manhattan and QQ-plots of the ten DKD GWAS meta-analysis results.**

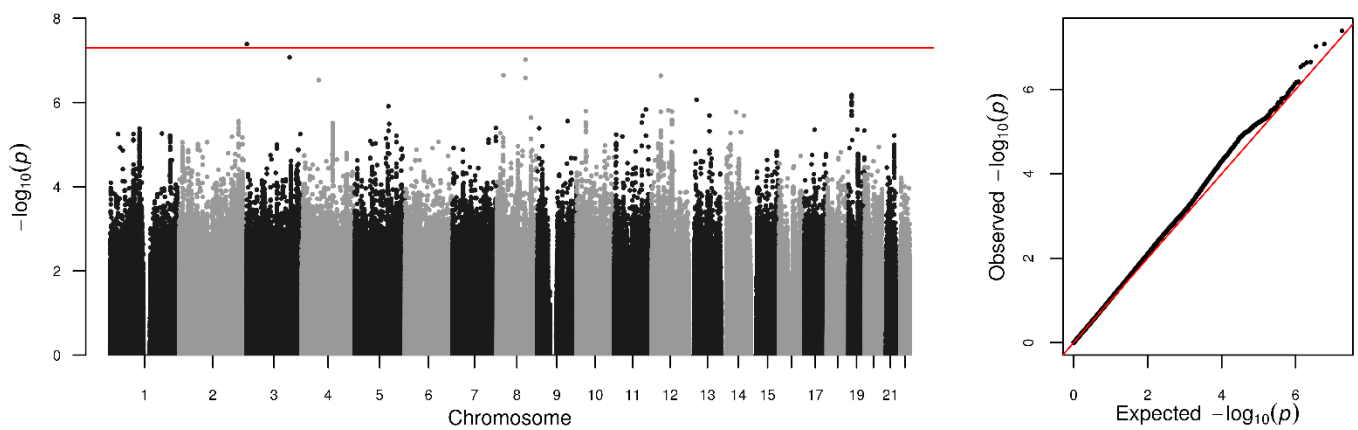
**A: Severe DKD** (Macroalbuminuria or ESRD vs. normal AER).  $\lambda_{GC} = 1.029$ , LDSR regression (LDSR) intercept = 1.019.



**B: Macro** (Macroalbuminuria vs. normal AER).  $\lambda_{GC} = 1.002$ , LDSR intercept = 1.028.

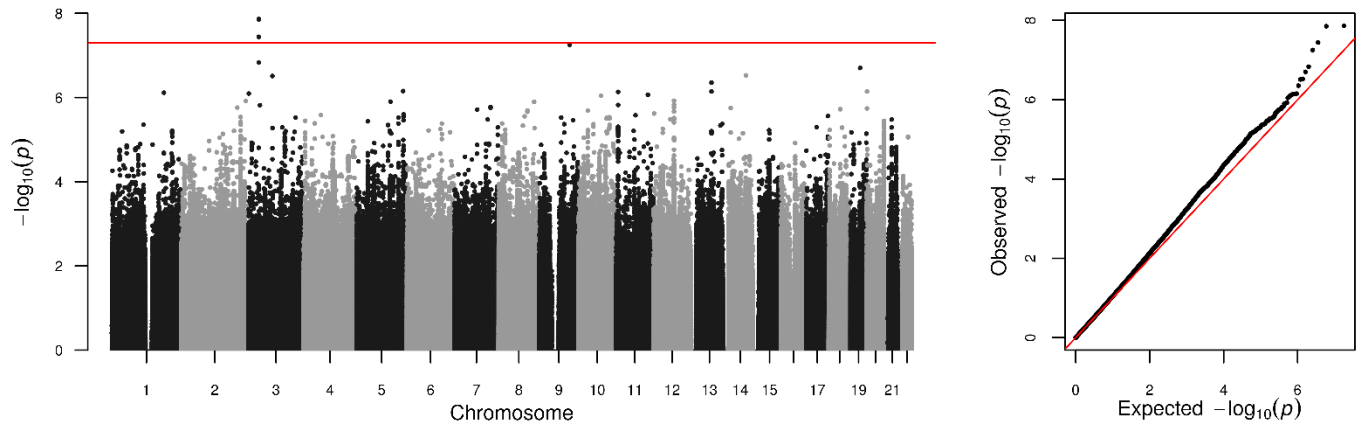


**C: ESRD** (ESRD vs. normal AER).  $\lambda_{GC} = 1.011$ , LDSR intercept = 1.018.

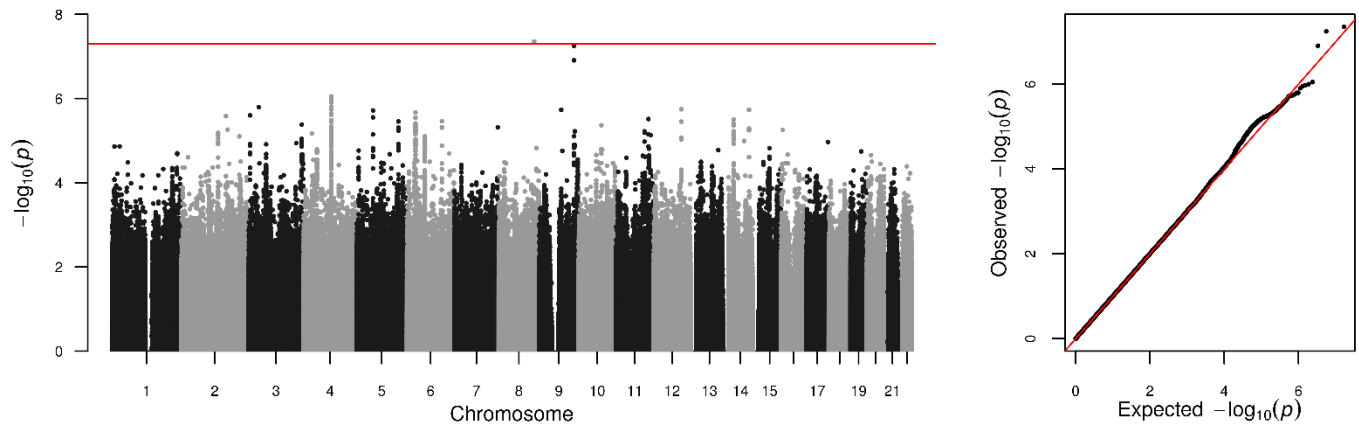




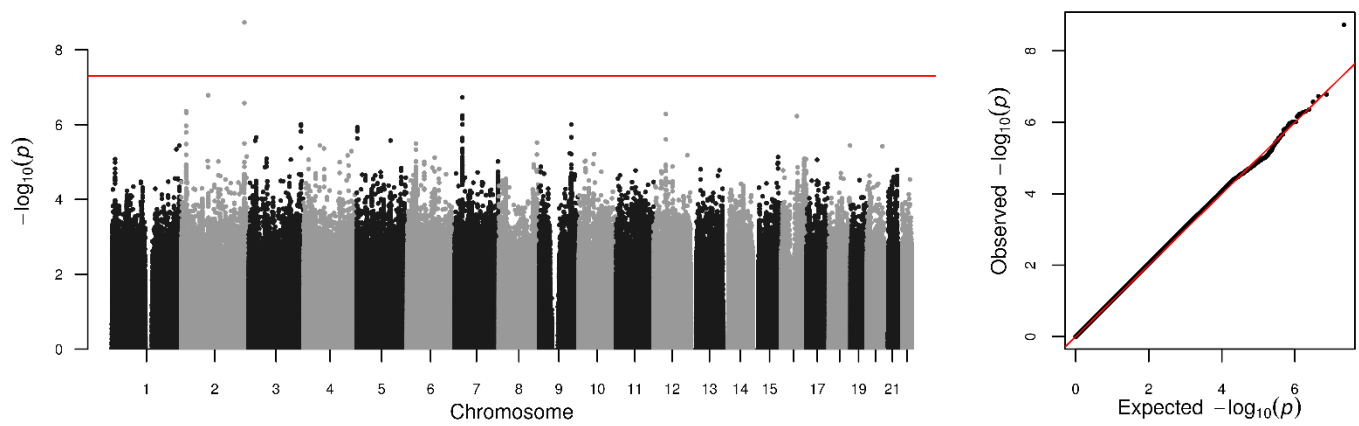
**D: ESRD vs. All** (ESRD vs. macro- or microalbuminuria or normal AER).  $\lambda_{GC} = 1.029$ , LDSR intercept = 1.054.



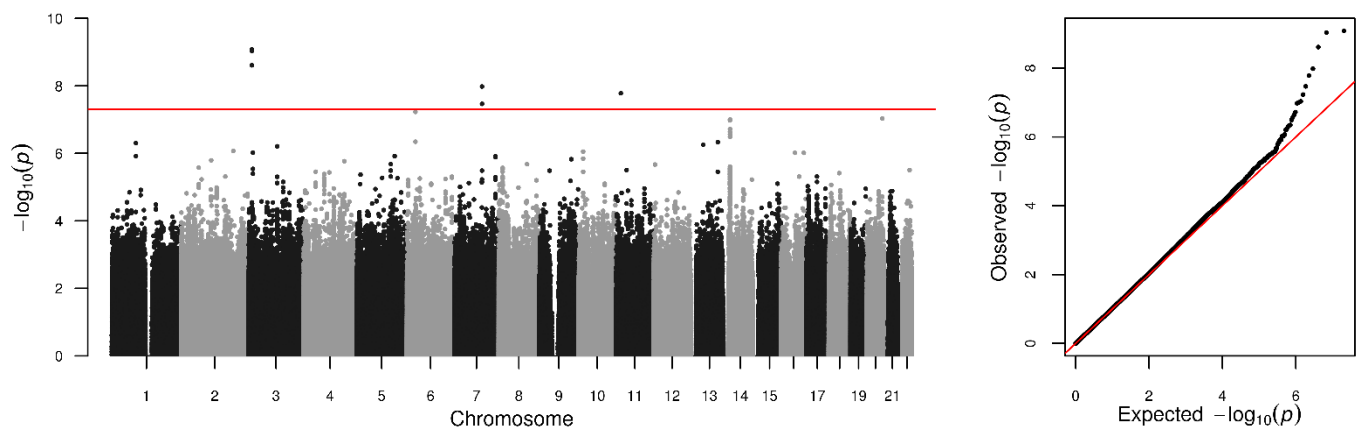
**E: ESRD vs. Macro** (ESRD vs. macroalbuminuria).  $\lambda_{GC} = 1.011$ , LDSR intercept = 1.009.



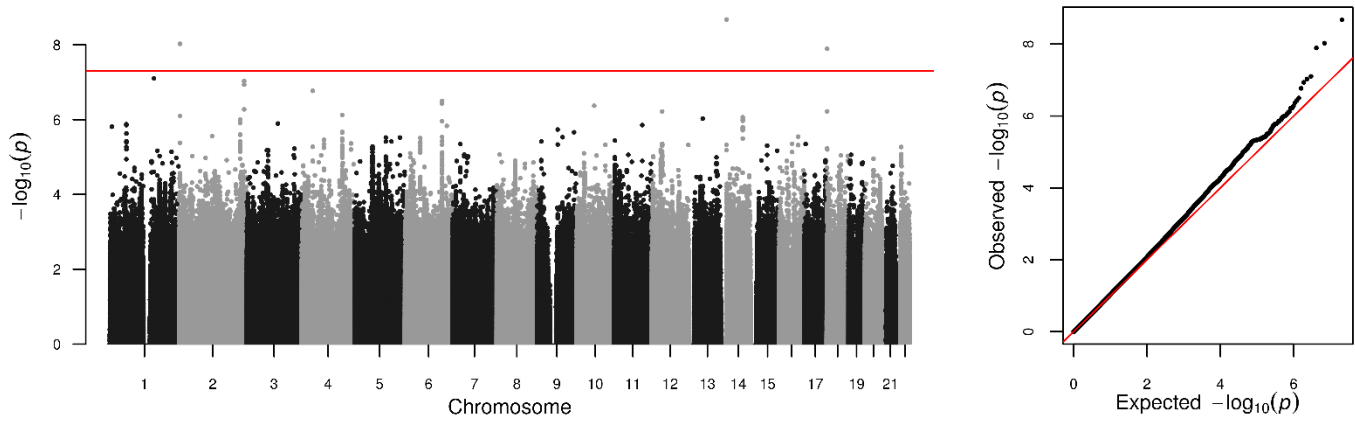
**F: All vs. Ctrl** (Micro- or Macroalbuminuria or ESRD vs. normal AER).  $\lambda_{GC} = 1.035$ , LDSR intercept = 1.005.



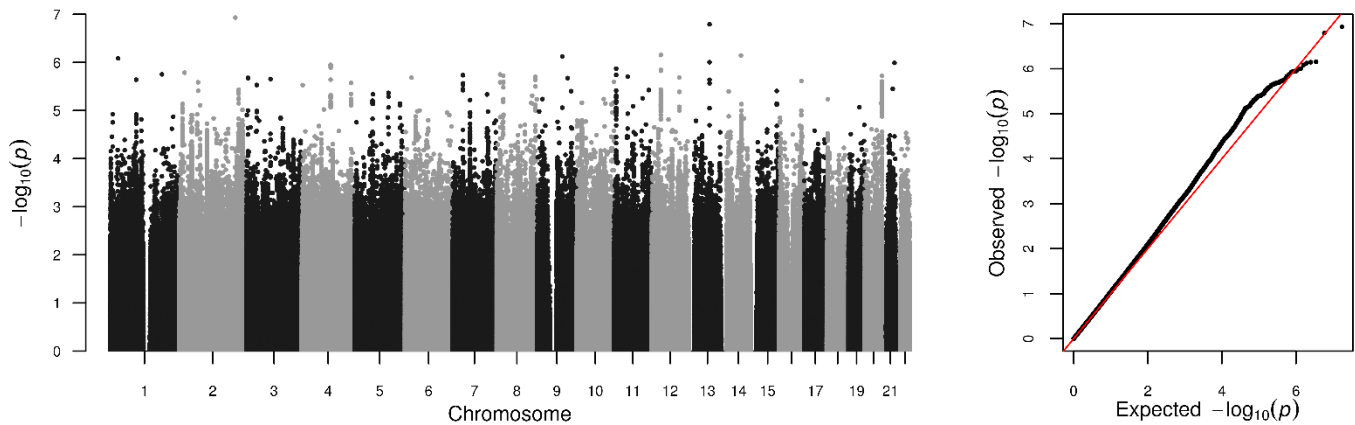
**G: Micro** (Microalbuminuria vs. normal AER).  $\lambda_{GC} = 1.008$ , LDSR intercept = 1.004.



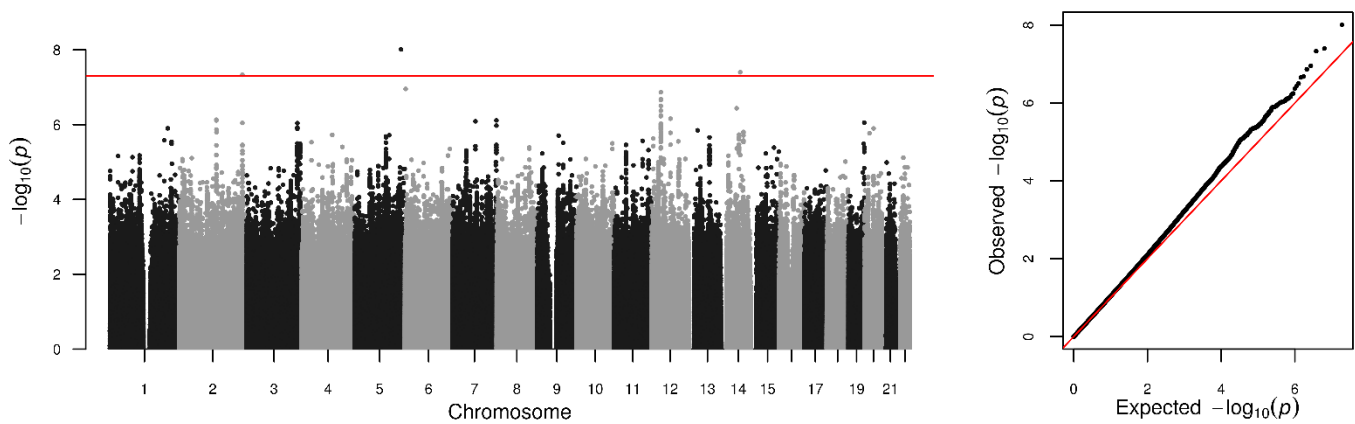
**H: CKD** (eGFR < 60 ml/min/1.73m<sup>2</sup> vs. eGFR ≥60 ml/min/1.73m<sup>2</sup>).  $\lambda_{GC} = 1.041$ , LDSR intercept = 1.028.



**I: CKD Extremes** (ESRD or eGFR < 15 ml/min/1.73m<sup>2</sup> vs. eGFR ≥ 60 ml/min/1.73m<sup>2</sup>).  $\lambda_{GC} = 1.023$ , LDSR intercept = 1.019.

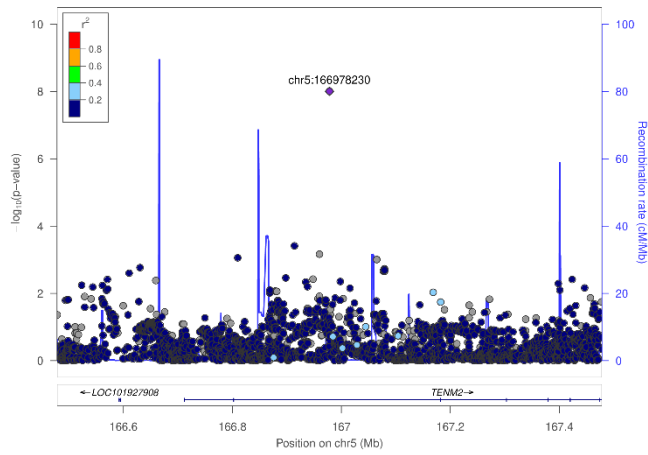


**J: CKD-DKD** (ESRD or eGFR < 45 ml/min/1.73m<sup>2</sup> and micro- or macroalbuminuria vs. eGFR ≥ 60ml/min/1.73m<sup>2</sup> and normal AER).  $\lambda_{GC} = 1.023$ , LDSR intercept = 1.031.

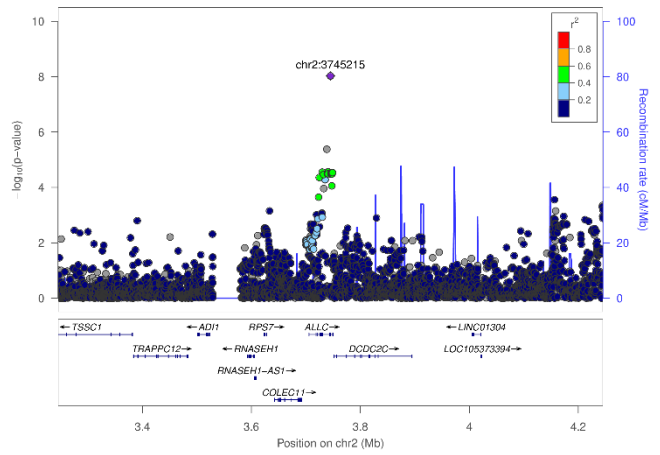


**Supplemental Figure 2: Regional association plots for the GWAS meta-analysis lead loci (A-K).**

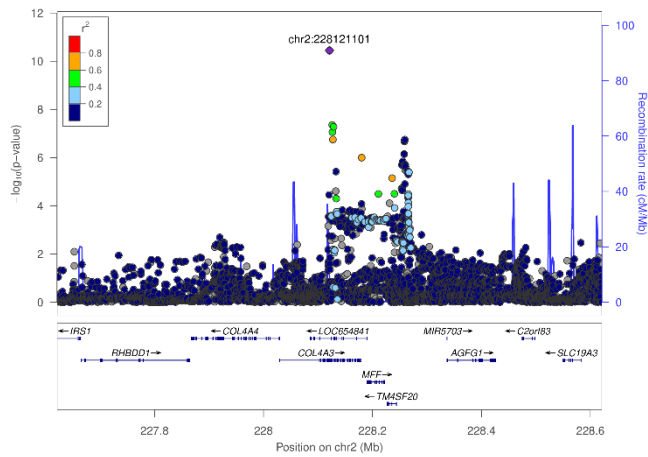
**A: CKD+DKD chr5:166978230 (rs72831309)**



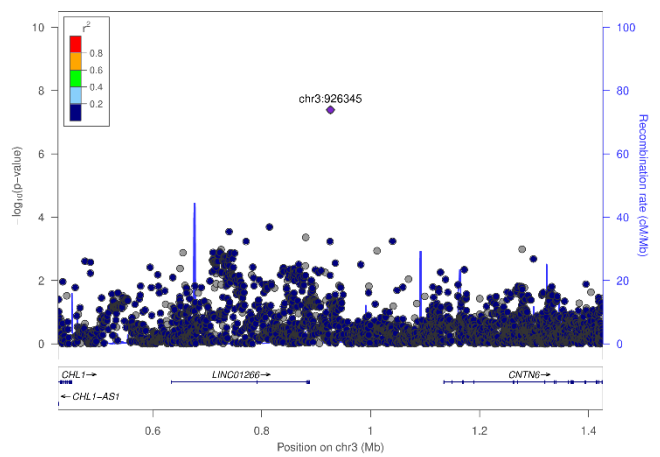
**B: CKD chr2:3745215 (rs12615970)**



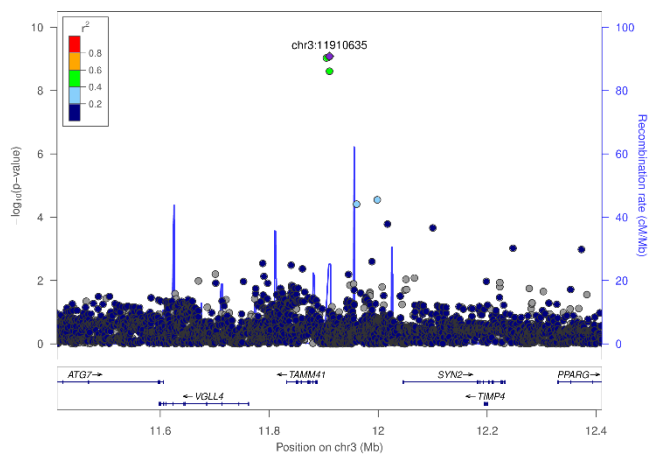
**C: Severe DKD chr2:228121101 (rs55703767)**



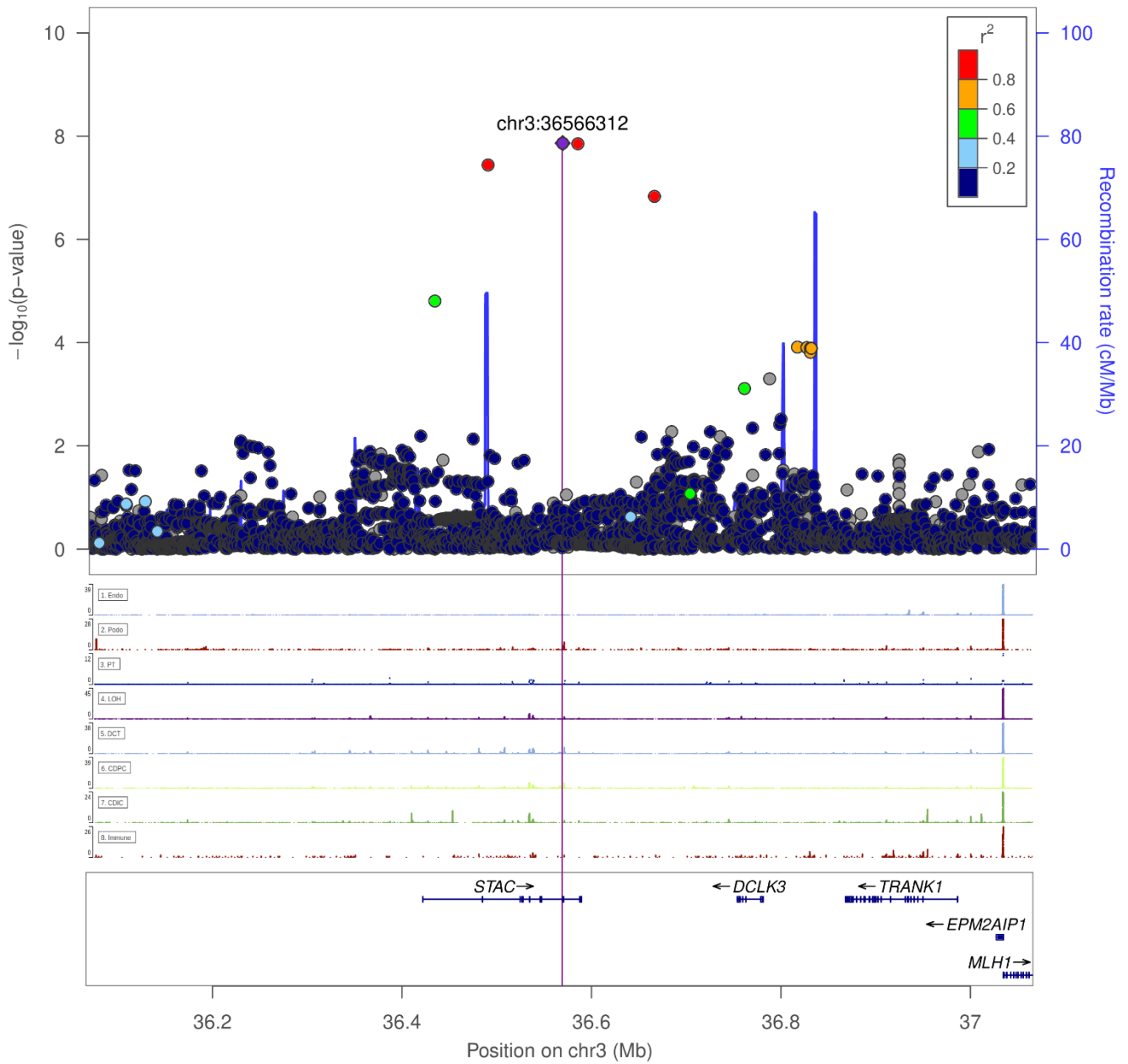
**D: ESRD chr3:926345 (rs115061173)**



**E: Micro chr3:11910635 (rs142823282)**

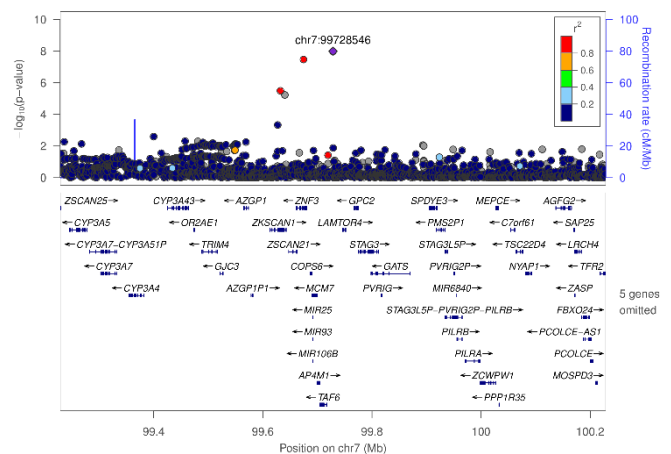
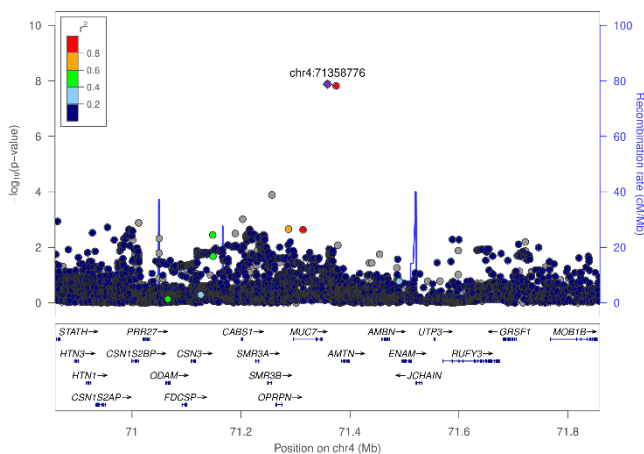


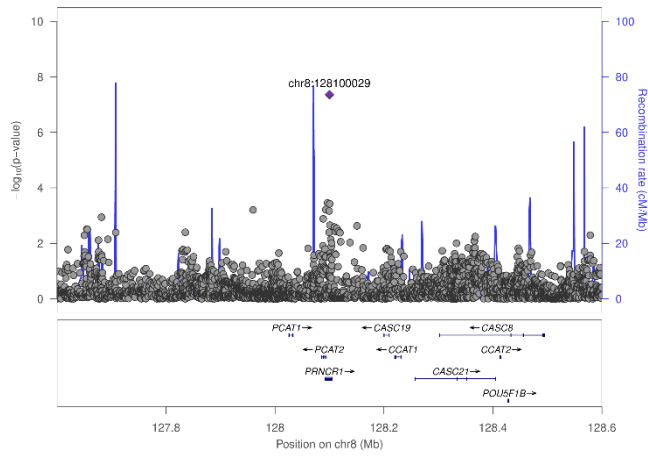
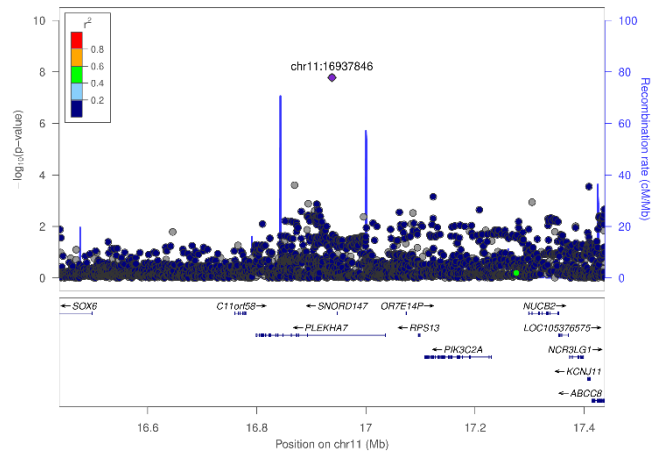
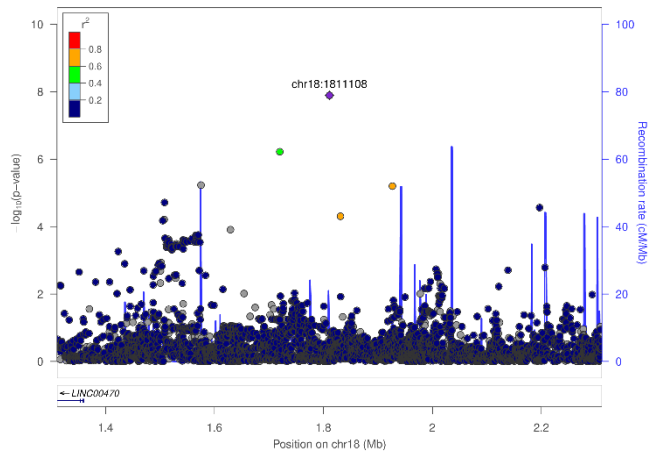
**F:** ESRD vs. All chr3:36566312 (rs116216059). The SNP rs116216059 is located on a single nucleus ATACseq (snATACseq) peak border in podocytes (PODO), peak value 1.1 (peak maximum value 7.0).



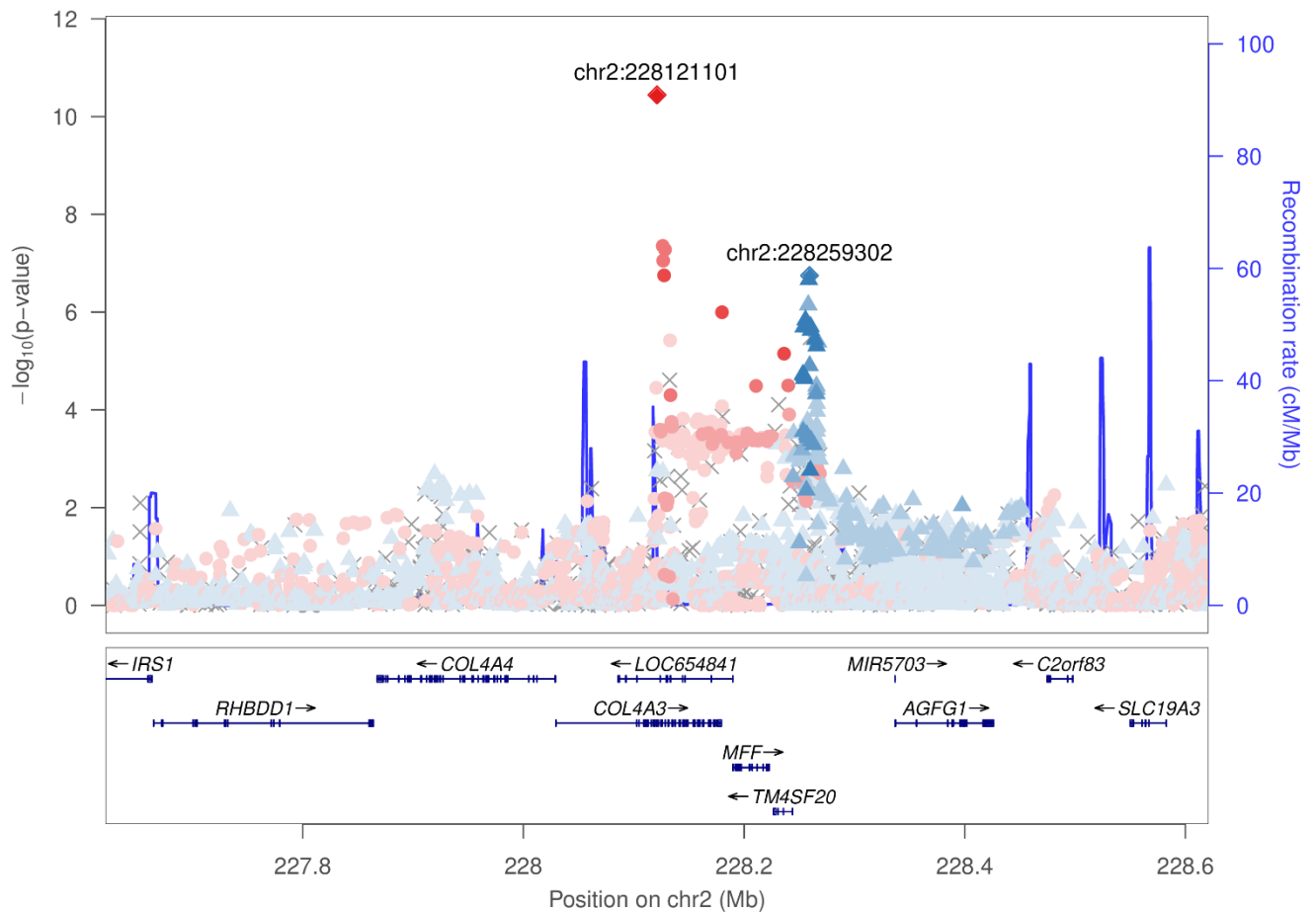
**G:** Severe DKD chr4:71358776 (rs191449639)

**H:** Micro chr7:99728546 (rs77273076)



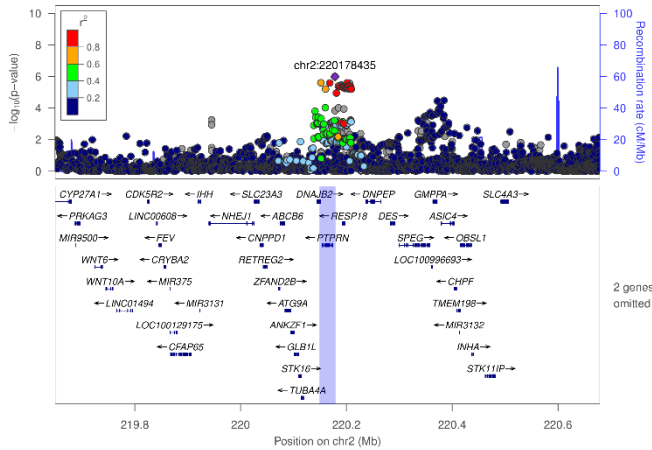
**I: ESRD vs. macro chr8:128100029 (rs551191707)****J: Micro chr11:16937846 (rs183937294)****K: CKD chr18:1811108 (rs185299109)**

**Supplemental Figure 3: Regional association plot for the *COL4A3* gene region associated with Severe DKD, indicating a secondary association peak at chr2:228259302 (rs6436688, effect allele (A) frequency 56%, OR = 1.13 (95% confidence interval 1.08 – 1.19), p-value  $1.79 \times 10^{-7}$ ). SNP rs6436688 is in partial LD ( $D' = 0.51$ ,  $r^2 = 0.08$ , 1000Genomes European ancestry populations) with the original *COL4A3* lead variant rs55703767. Variants are colored according to their LD correlation with the primary signal (chr2:228121101 rs55703767) in red, or with the secondary peak in blue; stronger color indicates stronger correlation.**

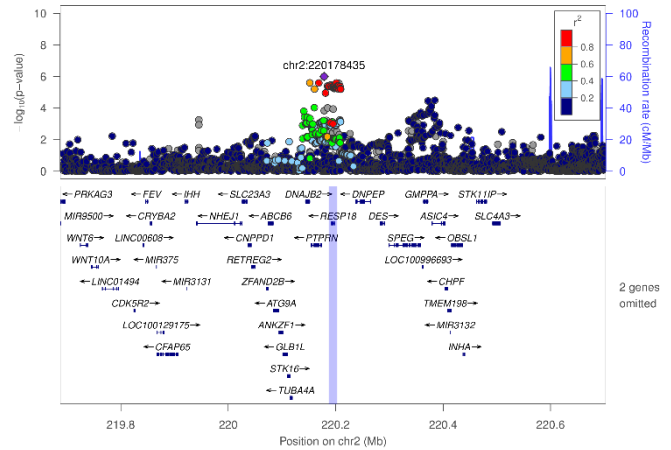


**Supplemental Figure 4: Regional association plots for the gene-level analysis results from MAGMA and PASCAL analysis.** The implicated gene region is highlighted in light blue. If the same gene was significant in both analyses, only MAGMA region is highlighted (gene flanking  $\pm 50$  kbp, vs.  $\pm 5$  kbp for PASCAL).

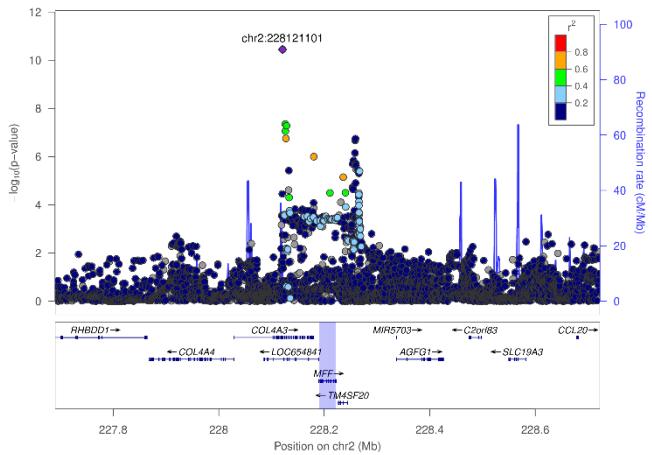
**A: CKD, *PTPRN***



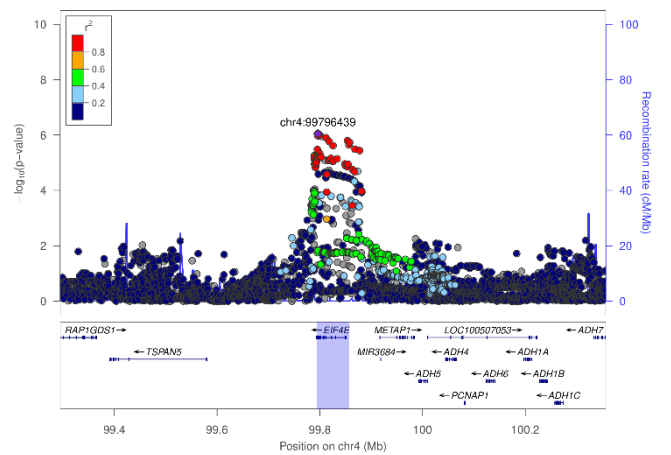
**B: CKD, *RESP18***



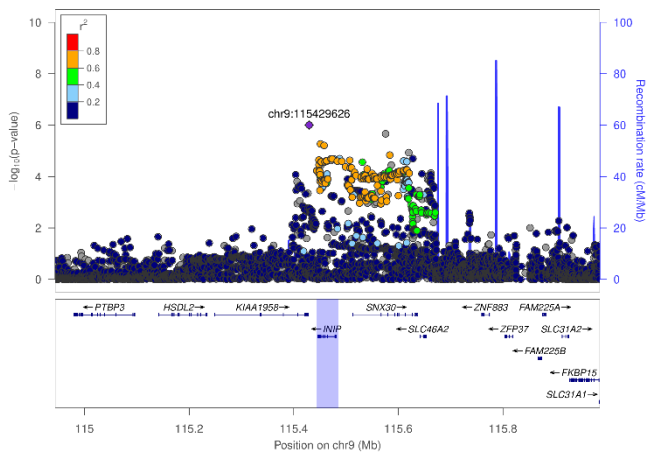
**C: Severe DKD, *MFF***



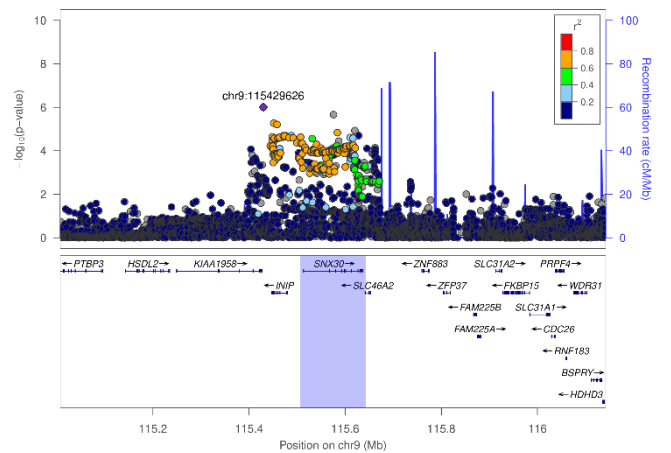
**D: ESRD vs. macro, *EIF4E***



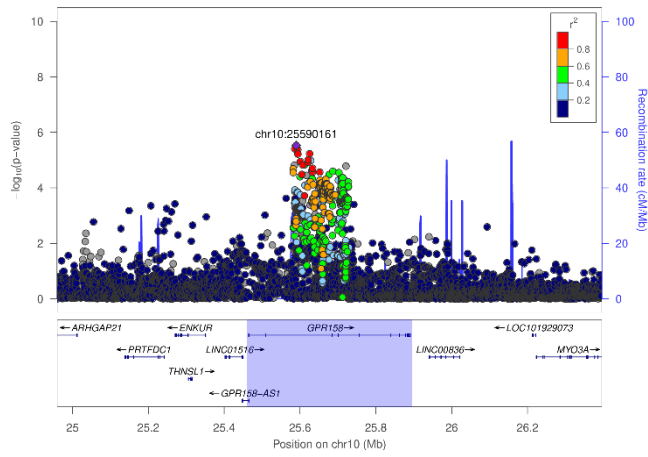
**E: DKD, *INIP***



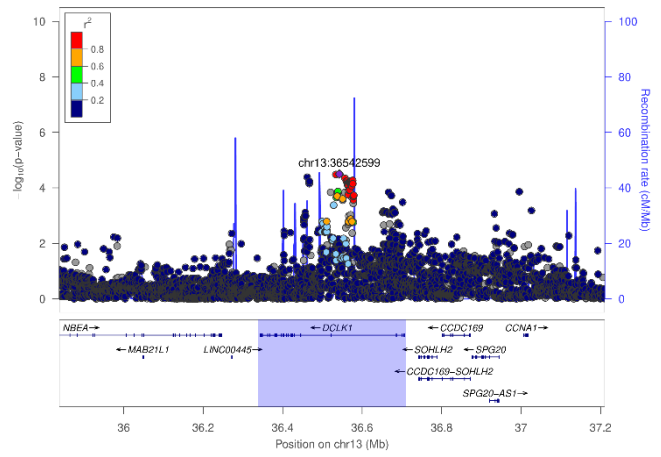
**F: DKD, *SNX30***



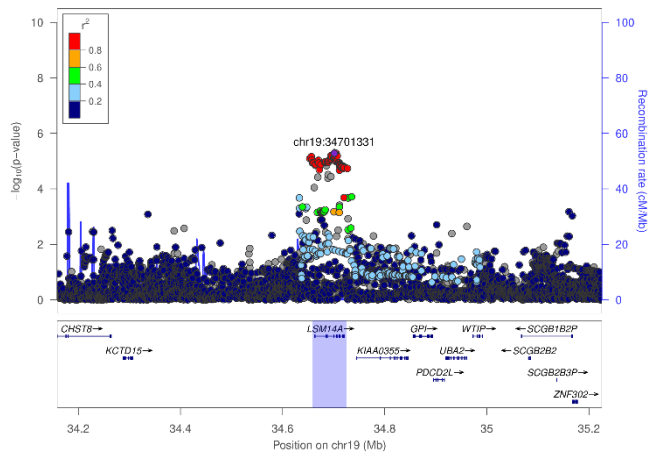
**G: Severe DKD, *GPR158***



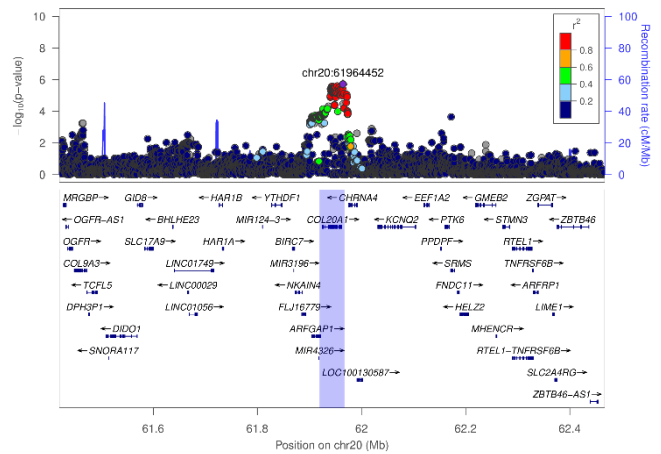
**H: ESRD vs. macro, *DCLK1***



**I: Severe DKD, *LSM14A***

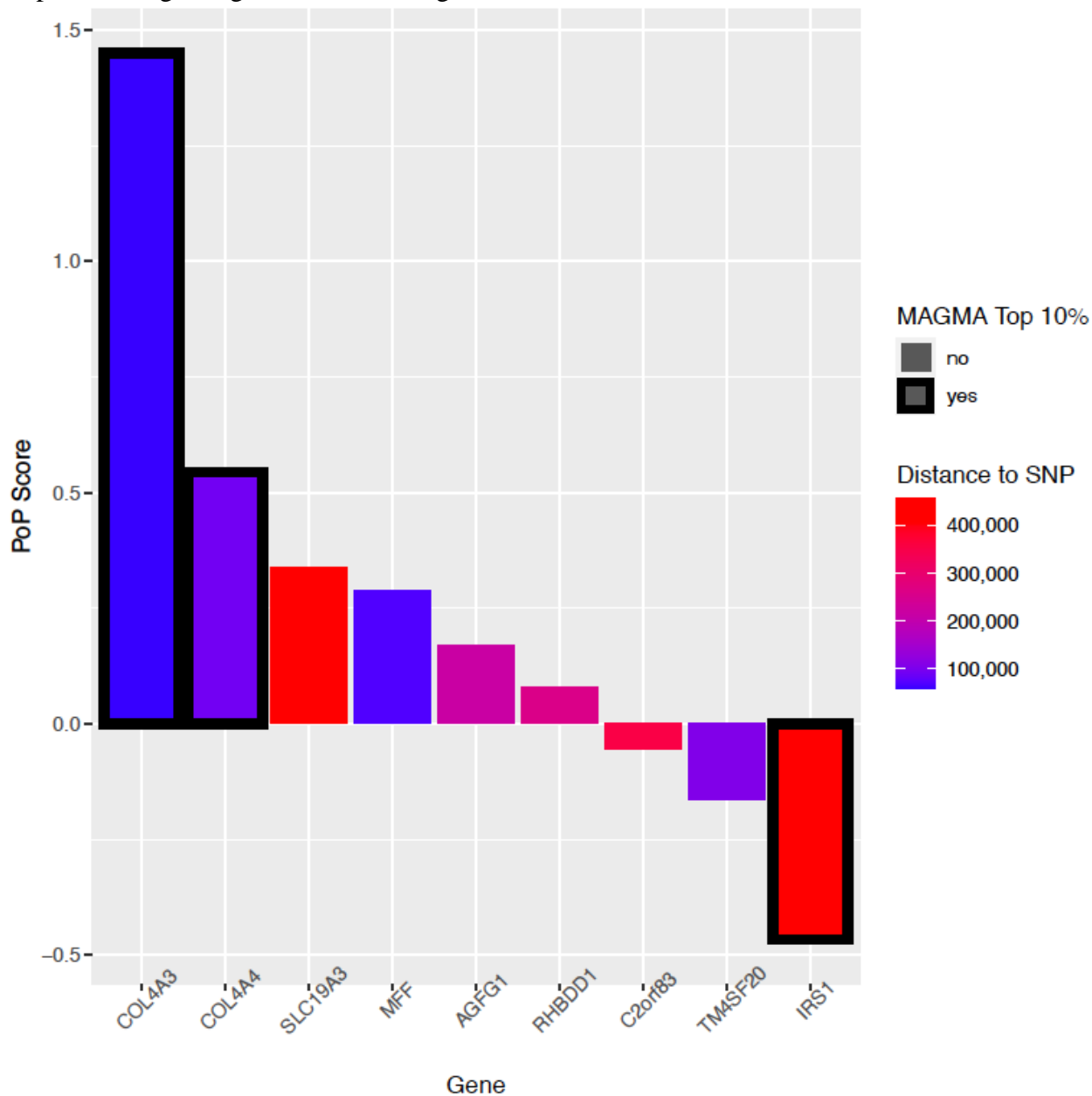


**J: CKD extremes, *COL20A1***

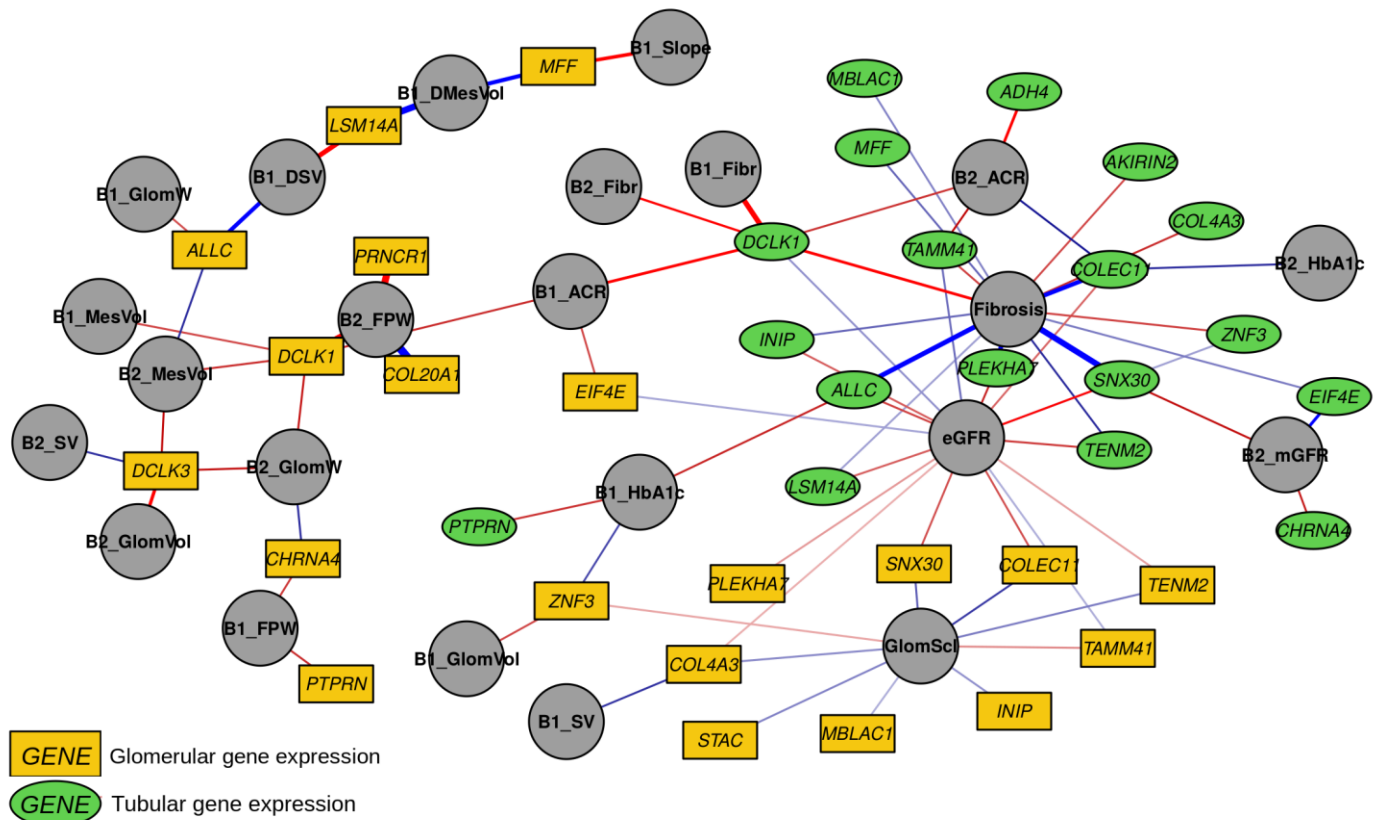




**Supplemental Figure 5: Gene prioritization for the COL4A3 gene at lead SNP rs5570367 associated with Severe DKD using multiple intersecting gene prioritization approaches (PoPS, nearest gene, and MAGMA). Plotted is the PoP Score (y-axis) versus the genes within a 500kb flanking window surrounding the lead SNP (x-axis), colored by distance to the lead SNP, and bolded if the gene was also within the top 10% of prioritized genes genome-wide using MAGMA.**

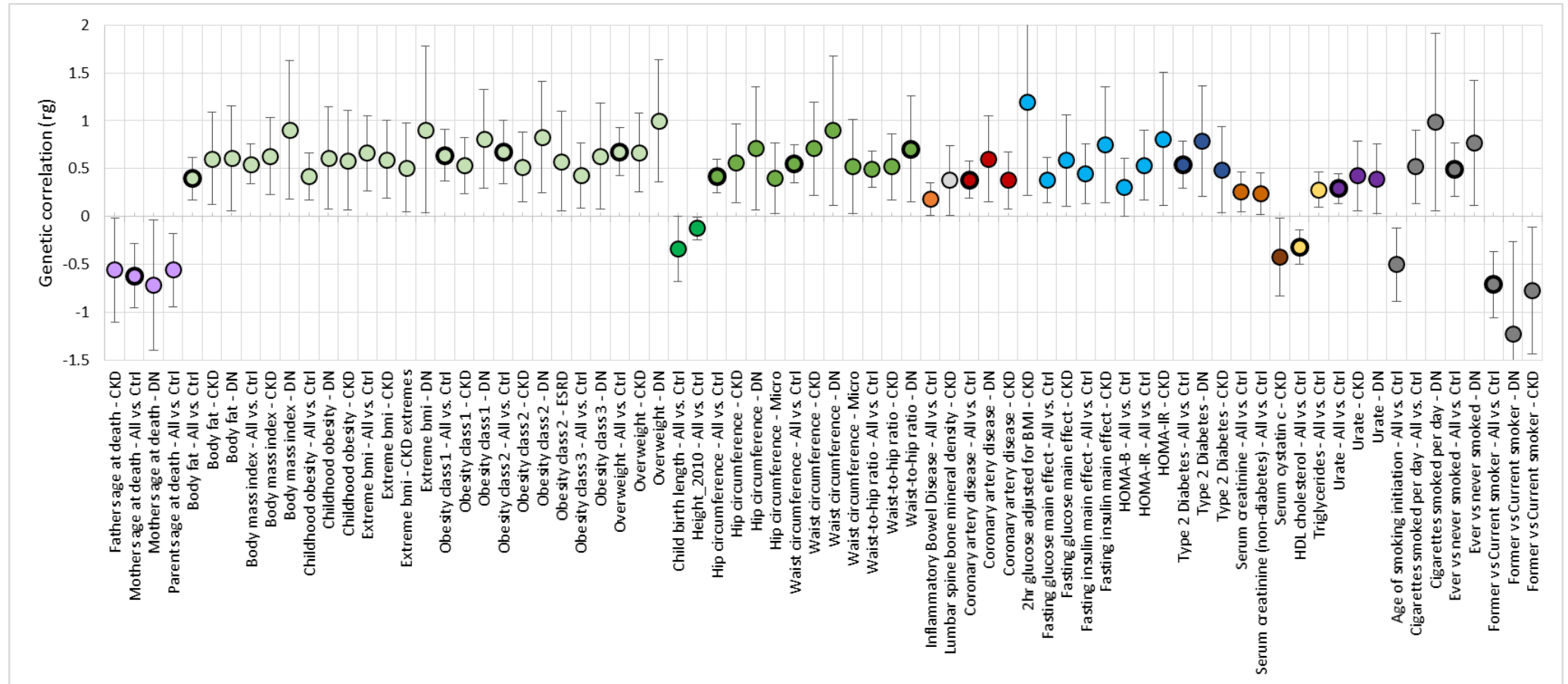


**Supplemental Figure 6: Tubular and glomerular gene expression of the lead genes correlate with multiple morphological and pathological renal parameters.** Golden rectangles indicate glomerular gene expression, green ellipses tubular gene expression, and gray circles the morphological phenotypes. All nominally significant correlations are shown. Blue edges indicate negative correlation, red edges positive correlation. Correlation with fibrosis, Glomerulosclerosis (GlomScl), and eGFR are measured in the nephrectomy samples. B1\_ and B2\_ indicate phenotypes from the first and second renal biopsies (B1, B2, respectively) from the Pima Indians, correlated with gene expression in transcriptomic data from the corresponding time point. B1/2\_GlomVol: glomerular volume; B1/2\_GlomW: glomerular width; B1/2\_FPW: podocyte foot process width. B1/2\_ACR: albumin creatinine ratio; B1/2\_Fibr: fibrosis; B1/2\_HbA1c: HbA1c B1/2\_MesVol: mesangial volume; B1/2\_SV: Surface volume of peripheral glomerular basement membrane per glomerulus; B1\_Slope: measured GFR (mGFR) slope between the B1 and B2. B1\_DMesVol: change in mesangial volume between B1 and B2. B1\_DSV: change in SV between B1 and B2.

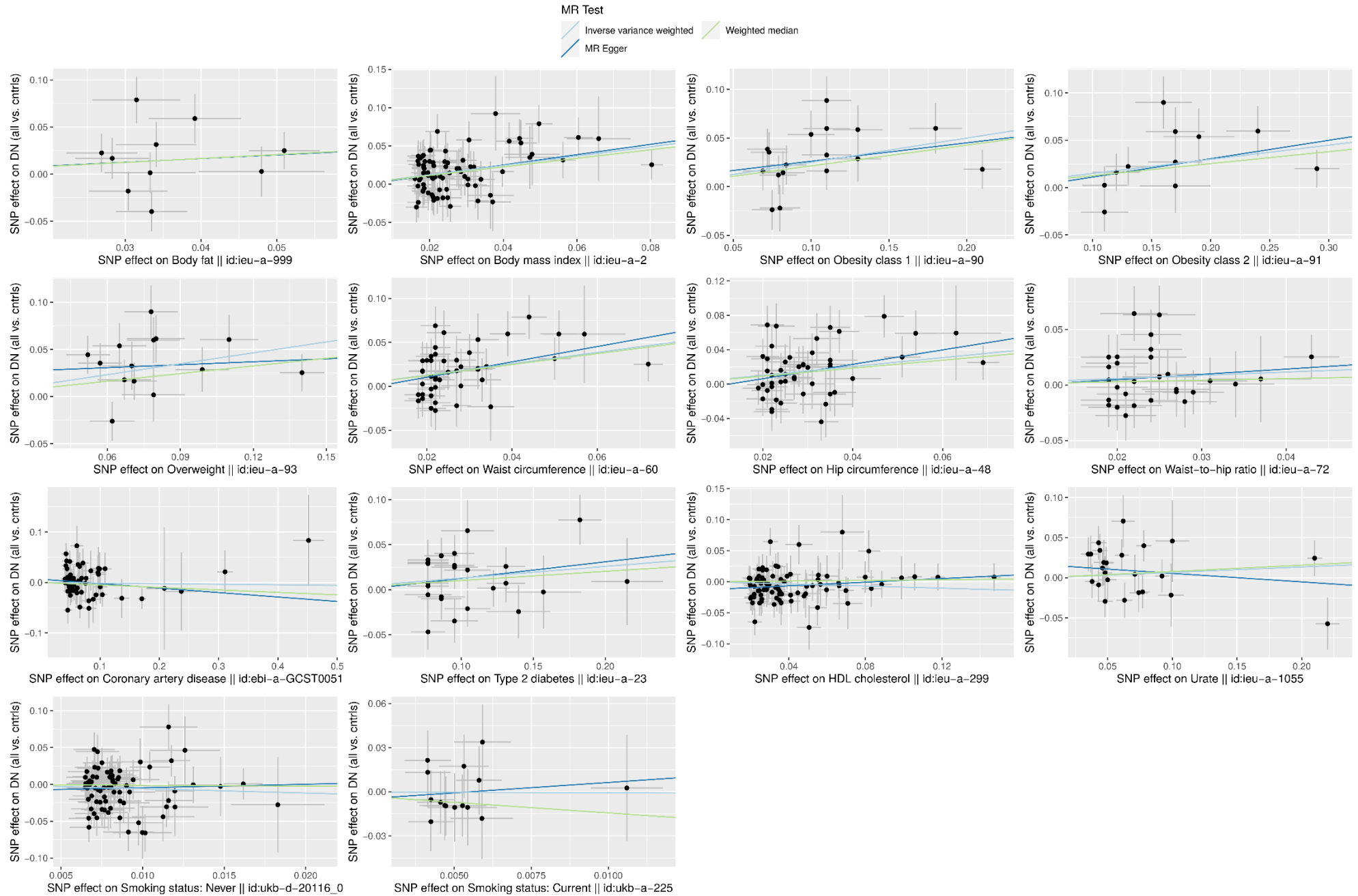


**Supplemental Figure 7: Genetic correlation between DKD and related traits based on LD score regression.**

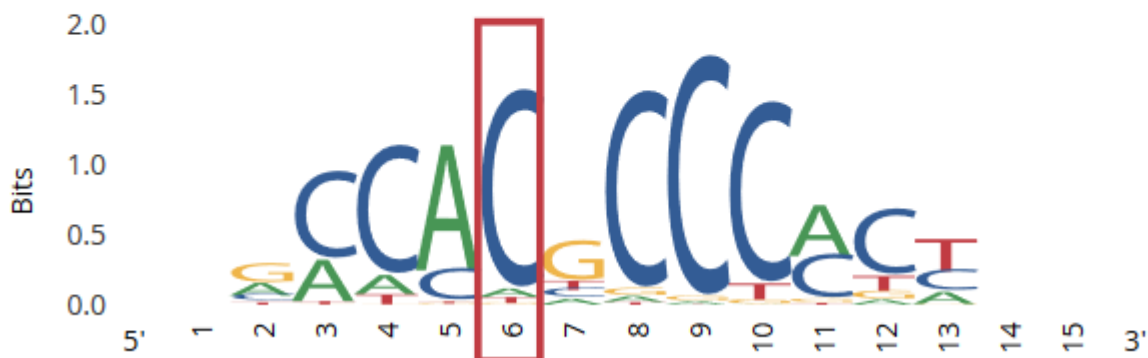
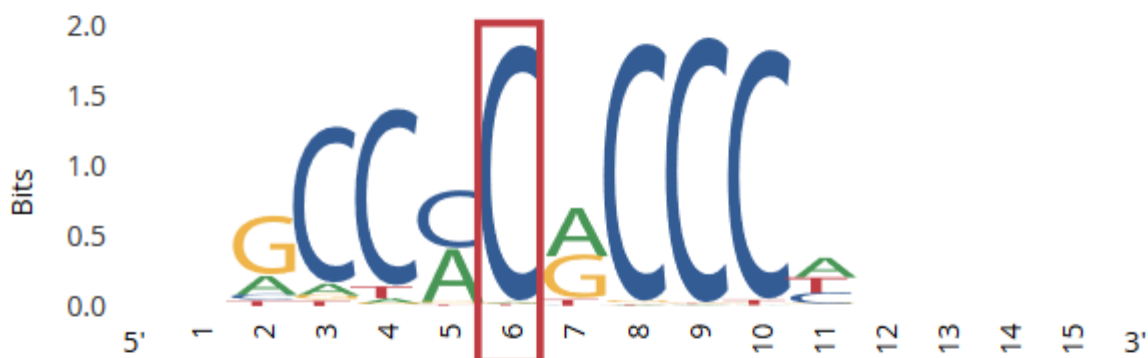
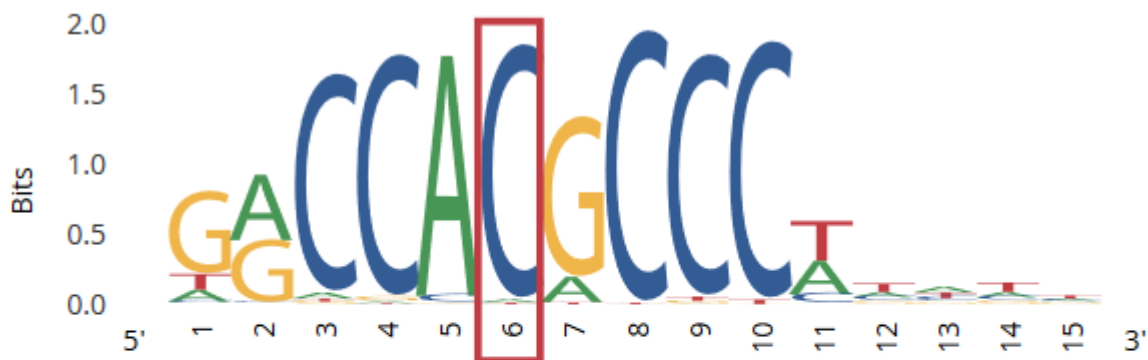
Only trait combinations with  $p < 0.05$  are shown, and traits that remained significant after correcting for 78 studied traits ( $p\text{-value} < 0.05/78 = 6.4 \times 10^{-4}$ ) are indicated with dark dot borders. Dot colors indicate aging related (light purple), anthropometric (green; including BMI and obesity related (light green), height (dark green), and waist and/or hip related (pale green)), inflammatory bowel disease (orange), bone mineral density (light gray), coronary artery disease (red), glycemic (light blue), type 2 diabetes (dark blue), serum creatinine and cystatin C (brown), lipids (yellow), uric acid (purple), and smoking related (dark gray) traits.



**Supplemental Figure 8: Mendelian Randomization scatter plots for SNP effects for the metabolic traits vs. DKD (All vs. Ctrl).** Lines indicate IVW, Weighted median, and MR Egger coefficients.



**Supplemental Figure 9: rs1260634 intronic in the ALLC gene affects the predicted binding motifs for KLF12, KLF4, and SP8 (top to bottom). Images obtained from the RegulomeDB (www.regulomedb.org)**



STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
<b>Title and abstract</b>	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	1
<b>Introduction</b>			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	3
<b>Methods</b>			
Study design	4	Present key elements of study design early in the paper	Fig 1
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	4
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Table S1
Bias	9	Describe any efforts to address potential sources of bias	10
Study size	10	Explain how the study size was arrived at	Table S2
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Table S1
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	5-10
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	
<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Table S2
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	ref to orig papers
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	

Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	Table S2
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Table 1; Bonf thresholds are given
		(b) Report category boundaries when continuous variables were categorized	
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
<b>Discussion</b>			
Key results	18	Summarise key results with reference to study objectives	15
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	19
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15-16
Generalisability	21	Discuss the generalisability (external validity) of the study results	19
<b>Other information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	20

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).