Blood UMOD and HER2 as causal mediators of CKD: supplementary material

Biomarker Assay Methodology in ORIGIN

At Myriad RBM Inc., the samples were thawed at room temperature (RT), vortexed, spun at 13,000g for 5 minutes for clarification. An aliquot was removed into a master microtiter plate for analysis. Using automated pipetting, an aliquot of each sample was introduced into one of the capture microsphere multiplexes of the Human DiscoveryMAP. The mixtures of sample and capture microspheres were thoroughly mixed and incubated at RT for 1 hour. Next, multiplexed cocktails of biotinylated reporter antibodies for each multiplex were added robotically. After thorough mixing, they were incubated for an additional hour at RT. Multiplexes were developed using an excess of streptavidin-phycoerythrin solution which was thoroughly mixed into each multiplex and incubated for 1 hour at RT. The volume of each multiplexed reaction was reduced by vacuum filtration and then increased by dilution into matrix buffer for analysis. Analysis was performed in Luminex 100 and 200 instruments and the resulting data stream was interpreted using proprietary data analysis software developed at RBM. For each multiplex, both calibrators and controls were included on each microtiter plate. Eight-point calibrators were run in the first and last column of each plate and 3-level quality controls were included in duplicate. Testing results were determined first for the high, medium and low controls for each multiplex to ensure proper assay performance. Unknown values for each of the analytes localized in a specific multiplex were determined using 4 and 5 parameter, weighted and non-weighted curve fitting algorithms included in the data analysis package.

Determining the Distribution of Each Biomarker in ORIGIN

Biomarkers were scrutinized in 5 steps. First, 26 biomarkers with undetectable levels in > 8409 (i.e. 99%) participants were excluded from further analyses. Second, scrutiny of the mean, median, and distribution of results and biologic literature pertaining to another 64 biomarkers with undetectable levels in > 1000 (i.e. 12%) participants led to exclusion of a further 21, leaving 237 biomarkers for analysis. Third, those biomarkers with levels below the level of quantification in < 10% of participants (n=850) were assigned a level corresponding to the lower limit of quantification. Fourth, biomarkers with levels below the level of quantification in > 10% of participants were identified and analyzed as ordinal variables as follows. A level of 1 was assigned to the participants with unquantifiable levels, dividing the remaining participants into 4 groups using quartiles. Values of 2, 3, 4 and 5 were assigned to participants within each progressively higher group. This approach was used to manage skewed biomarker distributions. Fifth, biomarkers with levels above the level of quantification were identified and those affected were assigned a level 1% above the upper limit of quantification. This approach led to 192 biomarkers for analysis as continuous variables and 45 biomarkers for analysis as 5-level ordinal variables.

The distributions of each of the 192 continuous biomarkers' levels were then scrutinized to identify extreme outliers with levels more than 4 standard deviations above or below the mean; levels that met those criteria were assigned the value corresponding to the mean plus or minus the 4th standard deviation respectively. Subsequently, the levels of 125 biomarkers with distributions that were not normally distributed were log-transformed using the natural logarithm. Biomarkers analyzed as continuous variables were then standardized to have mean 0 and standard deviation of 1. Finally, data from 93 participants in whom all 237 biomarkers were not analyzed due to insufficient volume of serum were excluded.

Genotyping Quality Control and Imputation in ORIGIN

SNPs were excluded on the basis of low call rate (<99%), deviation from Hardy-Weinberg ($p<10^{-6}$), and low minor allele frequency (MAF<0.01). Samples with low call rates (<99%), sex or ethnicity mismatches, or cryptic relatedness were also removed. We also removed ethnicities with small sample sizes (n<500). All quality control steps were performed using PLINK[1] and GCTA.[2] The 1000 Genomes Project[3] was used as the reference panel for ORIGIN imputation and was performed using the software IMPUTE2.[4,5] We removed SNPs imputed with low certainty (info<0.6, as defined by IMPUTE2).[5]

Biomarker Gene Identification

A distance of 300Kb was chosen based on observation of regional associations extending several hundred kilobases (Kb) away from known loci.[6] Genes falling within 300 Kb of each SNP were identified using the Reference Sequence gene list compiled by the UCSC (University of California Santa Cruz) Genome Table Browser.[7] Gene names were identified through the GeneCards Encyclopedia.[8]

Measurement of Biomarkers in Sample of Healthy Kidney Donors

To further explore the relationship of the novel markers found in the MR analysis, we assessed their concentration in ten healthy people before and after donor nephrectomy for renal transplantation at St. Joseph's Healthcare, Hamilton, ON. Blood and urine samples were collected within 3-months prior to nephrectomy (pre-op) and 3 months after the operation (post-op), once kidney function had returned to normal according to eGFR. Participant samples were stored at -80 °C until required for further analyses. Plasma samples were thawed at room temperature and biomarkers were measured using a solid-phase sandwich enzyme-linked immunosorbent assay (ELISA). 90mL urine collection containers were thawed completely and mixed thoroughly. Approximately 1mL of urine was aliquoted from the collection container into Eppendorf tubes. The Eppendorf tubes were centrifuged at room temp for 7 minutes at 16,100 xg. 10uL from each Eppendorf was aliquoted into a 2nd tube with 190uL Calibrator diluent (referred to as dilution #1). All "dilution #1" tubes were vortexed thoroughly. These tubes were <u>not</u> centrifuged. 10uL from each "dilution #1" Eppendorf was aliquoted into a 3rd tube with 490uL Calibrator diluent (referred to as dilution #2). All "dilution #2" tubes were vortexed thoroughly, then centrifuged at room temp for 7 minutes @ 16,100 xg. Urinary uromodulin was indexed to creatinine as uromodulin-to-creatinine ratio in order to account for differences in urine concentration. ELISA was performed using the Human Magnetic Luminex Screening Assay for Uromodulin (BR64) from R&D according to manufacturer's specifications. The biomarker concentration was determined using Bio-Rad Bio-Plex 200 system. The optical density was analyzed and exported for further analysis using BioPlex Manager 6.1.

Association of UMOD and HER2 concentration with CKD in ORIGIN

In addition to a minimally adjusted epidemiological model, we also tested models further adjusting for traditional CKD risk factors as a sensitivity analysis, namely prior type 2 diabetes, prior renal disease, BMI, estimate glomerular filtration rate (eGFR), current smoker, diagnosis of hypertension, and LDL (mmol/L). Consistent with our minimally adjusted models, we found that increased levels of blood HER2 was associated with an increased risk of incident CKD, while blood UMOD was associated with a decreased risk of incident CKD (HER2: odds ratio (OR)=1.07 per SD; 95% CI, 1.01 to 1.13; p=0.026 and UMOD: OR=0.86 per SD; 95% CI, 0.81 to 0.92; $p<9\text{x}10^{-7}$).

Additional MR analyses

In addition to the bootstrap method employed, we have also tested UMOD and HER2 in an IVW fixed-effect model using the 'MR-base' package in R, where consistent estimates were obtained (UMOD: OR=1.32 per SD; 95% CI 1.28 to 1.37; p=4.1x10-44; HER2: OR=1.30 per SD; 95% CI 1.15 to 1.48; p=4.4x10-5). We also performed a leave-one-out sensitivity analysis to ensure that no single variant was driving the observed causal effect and found consistent findings for each SNP excluded (Supplementary Table 3 and 4)

Supplementary Table 1: Summary of 5 independent (R^2<0.1) SNP associations at the *ERBB2* locus (+/- 300KB) located on chromosome 17. Association statistics provided for serum HER2 (in ORIGIN) and CKD (from CKDGen).

SNP	Position	MAF	Effect allele	Other allele	HER2 Beta	HER2 SE	HER2 p-value	CKD Beta	CKD SE	CKD p-value
rs11655584	38139024	0.080	T	C	0.1023	0.0393	0.00925	0.041	0.028	0.15
rs1565922	37831035	0.319	G	A	0.1386	0.0226	9.42E-10	0.022	0.016	0.19
rs17608925	38082831	0.148	C	T	0.1443	0.036	6.08E-05	0.066	0.024	0.0068
rs9892427	37804858	0.066	C	T	0.1949	0.0444	1.15E-05	0.067	0.028	0.018
rs9896218	37894463	0.085	C	A	0.2731	0.0484	1.83E-08	0.052	0.033	0.11

MAF: minor allele frequency (from CKDGen), SE: standard error. Beta coefficient corresponds to the risk coefficient for each SD increase in effect allele.

Supplementary Table 2: Summary of 17 independent (R^2<0.1) SNP associations at the *UMOD* locus (+/- 300KB) located on chromosome 16. Association statistics provided for serum UMOD (in ORIGIN) and CKD (from CKDGen).

SNP	Position	MAF	Effect allele	Other allele	UMOD Beta	UMOD SE	UMOD p-value	CKD Beta	CKD SE	CKD p-value
rs12917707	20367690	0.151	T	G	-0.7709	0.028	8.35E-154	-0.22	0.021	2.10E-25
rs1155876	20522779	0.084	C	T	-0.128	0.0331	0.000110	-0.088	0.027	0.0012
rs12446492	20408377	0.389	A	T	-0.3293	0.0235	1.05E-43	-0.100	0.016	3.20E-10
rs12708631	20365697	0.075	A	T	-0.1312	0.0357	0.000241	-0.130	0.034	7.00E-05
rs12920537	20421556	0.310	G	A	-0.236	0.0293	1.16E-15	-0.066	0.021	0.0017
rs12920708	20407237	0.230	C	A	-0.3083	0.0257	1.50E-32	-0.067	0.019	4.00E-04
rs12930599	20335325	0.219	A	G	-0.2183	0.0309	1.77E-12	-0.036	0.029	0.21
rs4380062	20252959	0.086	T	C	-0.1215	0.0387	0.00172	-0.012	0.028	0.67
rs4462596	20232265	0.261	G	A	-0.0935	0.0296	0.00159	-0.037	0.017	0.027
rs4558425	20309215	0.058	A	G	-0.1138	0.0342	0.000876	-0.046	0.034	0.17
rs6497445	20216617	0.133	C	T	0.0886	0.0302	0.00331	0.020	0.022	0.37
rs7187470	20419775	0.451	A	G	0.1134	0.0293	0.000110	0.0022	0.022	0.92
rs7192921	20129595	0.406	G	T	0.0665	0.0222	0.00274	0.017	0.015	0.28
rs7198000	20351937	0.124	A	G	0.1804	0.0324	2.82E-08	0.062	0.024	0.0088
rs7498776	20611149	0.080	C	T	0.104	0.0325	0.00140	-0.034	0.027	0.21
rs7499304	20564390	0.288	T	G	-0.1043	0.0284	0.000249	0.012	0.019	0.54
rs8060932	20344077	0.315	G	A	0.1586	0.0274	7.23E-09	0.044	0.019	0.017

MAF: minor allele frequency (from CKDGen), SE: standard error. Beta coefficient corresponds to the risk coefficient for each unit increase in effect allele.

Supplementary Table 3: Summary of leave-one-out sensitivity MR analysis for UMOD.

SNP excluded	OR	95% CI	P-value
rs1155876	1.31	(1.26, 1.37)	1.24E-42
rs12446492	1.31	(1.26, 1.37)	1.15E-35
rs12708631	1.31	(1.26, 1.36)	1.36E-42
rs12917707	1.30	(1.23, 1.38)	3.70E-20
rs12920537	1.32	(1.27, 1.37)	5.89E-42
rs12920708	1.33	(1.27, 1.38)	1.28E-41
rs12930599	1.32	(1.27, 1.37)	6.23E-44
rs4380062	1.32	(1.27, 1.37)	3.35E-44
rs4462596	1.32	(1.27, 1.37)	3.57E-43
rs4558425	1.32	(1.27, 1.37)	9.41E-44
rs6497445	1.32	(1.27, 1.37)	6.11E-44
rs7187470	1.32	(1.27, 1.37)	1.70E-44
rs7192921	1.32	(1.27, 1.37)	7.82E-44
rs7198000	1.32	(1.27, 1.37)	1.03E-42
rs7498776	1.32	(1.27, 1.37)	5.98E-45
rs7499304	1.32	(1.27, 1.38)	4.83E-45
rs8060932	1.32	(1.27, 1.37)	6.09E-43
Overall	1.32	(1.27, 1.37)	4.12E-44

Supplementary Table 4: Summary of leave-one-out sensitivity MR analysis for HER2.

SNP excluded	OR	95% CI	P-value
rs11655584	1.29	(1.13, 1.47)	0.00012
rs1565922	1.36	(1.17, 1.59)	6.38E-05
rs17608925	1.26	(1.1, 1.44)	0.0010
rs9892427	1.27	(1.11, 1.47)	0.00074
rs9896218	1.34	(1.15, 1.55)	0.00012
Overall	1.30	(1.15, 1.48)	4.39E-05

Supplementary Table 5: List of all biomarkers tested and their corresponding genes.

	Biomarker	Gene
1	6Ckine	*CCL21*
2	Adiponectin	*ADIPOQ*
3	Adrenomedullin	*ADM*
4	Agouti-Related Protein	*AGRP*
5	Aldose Reductase	*AKR1B1*
6	Alpha-1-acid glycoprotein 1	*ORM1*
7	Alpha-1-Antichymotrypsin	*SERPINA3*
8	Alpha-1-Antitrypsin	*SERPINA1*
9	Alpha-1-Microglobulin	*AMBP*
10	Alpha-2-Macroglobulin	*A2M*
11	Angiogenin	*ANG*
12	Angiopoietin-2	*ANGPT2*
13	Angiopoietin-related protein 3	*ANGPTL3*
14	Angiotensin-Converting Enzyme	*ACE*
15	Angiotensinogen	*AGT*
16	Antithrombin-III	*SERPINC1*
17	Apolipoprotein A-I	*APOA1*
18	Apolipoprotein A-II	*APOA2*
19	Apolipoprotein A-IV	*APOA4*
20	Apolipoprotein B	*APOB*
21	Apolipoprotein C-I	*APOC1*
22	Apolipoprotein C-III	*APOC3*
23	Apolipoprotein D	*APOD*
24	Apolipoprotein E	*APOE*
25	Apolipoprotein H	*APOH*
26	Apolipoprotein(a)	*LPA*
27	AXL Receptor Tyrosine Kinase	*AXL*
28	B cell-activating factor	*TNFSF13B*
29	B Lymphocyte Chemoattractant	*CXCL13*
30	Beta Amyloid 1-40	*APP*
31	Beta-2-Microglobulin	*B2M*
32	Brain-Derived Neurotrophic Factor	*BDNF*
33	C-Peptide	*INS*
34	C-Reactive Protein	*CRP*
35	Cathepsin D	*CTSD*
36	CD 40 antigen	*CD40*
37	CD163	*CD163*
38	CD40 Ligand	*CD40LG*
39	CD5 Antigen-like	*CD5L*
40	Cellular Fibronectin	*FN1*
41	Chemerin	*RARRES2*
42	Chemokine CC-4	*CCR4*
43	Chromogranin-A	*CHGA*
44	Clusterin	*CLU*
45	Collagen IV	*COL4A1,COL4A2,COL4A3,COL4A4,COL 4A5,COL4A6*
46	Complement C3	*C3*
47	Complement Factor H Related Protein 1	*CFHR1*
48	Cortisol	*NA*
49	Creatine Kinase-MB	*CKM,CKB*
50	Cystatin-C	*CST3*
51	E-Selectin	*SELE*
52	EN-RAGE	*S100A12*
53	Endoglin	*ENG*
54	Endostatin	*COL18A1*
55	Eotaxin-1	*CCL11*
56	Eotaxin-2	*CCL24*
57	Eotaxin-3	*CCL26*
58	Epithelial-Derived Neutrophil-Activating Protein 78	*CXCL5*
59	Erythropoietin	*EPO*
60	Ezrin	*EZR*
61	Factor VII	*F7*
62	Fas Ligand	*FASLG*
63	FASLG Receptor	*TNFRSF6B*
64	Fatty Acid-Binding Protein adipocyte	*FABP4*
65	Fatty Acid-Binding Protein liver	*FABP1*

66	Ferritin	*FTL,FTH1*
67	Fetuin-A	*AHSG*
68	Fibroblast Growth Factor 21	*FGF21*
69	Fibroblast growth factor 23	*FGF23*
70	Fibulin-1C	*FBLN1*
71	Ficolin-3	*FCN3*
72	Follicle-Stimulating Hormone	*FSHB,CGA*
73	Galectin-3	*LGALS3*
74	Gastric inhibitory polypeptide	*GIP*
75	Gelsolin	*GSN*
76	Glucagon-like Peptide 1 total	*GCG*
77	Glucose-6-phosphate Isomerase	*GPI*
78	Glutathione S-Transferase alpha	*GSTA1,GSTA2,GSTA3,GSTA4,GSTA5*
79	Glycogen phosphorylase isoenzyme BB	*PYGB*
80	Granulocyte Colony-Stimulating Factor	*CSF3*
81	Growth differentiation factor 15	*GDF15*
82	Growth Hormone	*GH1,GH2*
83	Growth-Regulated alpha protein	*CXCL1*
84	Haptoglobin	*HP*
85	Heat-Shock protein 70	*HSPA1A,HSPA1B,HSPA1L,HSPA2,HSPA4,HSPA4L,HSPA5,HSPA6,HSPA8,HSPA9,HSPA12A,HSPA12B,HSPA13,HSPA14*
86	Hemopexin	*HPX*
87	Hepatocyte Growth Factor	*HGF*
88	Hepatocyte Growth Factor receptor	*MET*
89	Hepsin	*HPN*
90	Human Epidermal Growth Factor Receptor 2	*ERBB2*
91	Immunoglobulin A	*IGH*
92	Immunoglobulin E	*IGH*
93	Immunoglobulin M	*IGH*
94	Insulin	*INS*
95	Insulin-like Growth Factor Binding Protein 4	*IGFBP4*
96	Insulin-like Growth Factor Binding Protein 5	*IGFBP5*
97	Insulin-like Growth Factor Binding Protein 6	*IGFBP6*
98	Insulin-like Growth Factor I	*IGF1*
99	Insulin-like Growth Factor-Binding Protein 1	*IGFBP1*
100	Insulin-like Growth Factor-Binding Protein 2	*IGFBP2*
101	Insulin-like Growth Factor-Binding Protein 3	*IGFBP3*
102	Intercellular Adhesion Molecule 1	*ICAM1*
103	Interferon gamma	*IFNG*
104	Interferon gamma Induced Protein 10	*CXCL10*
105	Interferon-inducible T-cell alpha chemoattractant	*CXCL11*
106	Interleukin-1 beta	*IL1B*
107	Interleukin-1 receptor antagonist	*IL1RN*
108	Interleukin-10	*IL10*
109	Interleukin-12 Subunit p40	*IL12B*
110	Interleukin-16	*IL16*
111	Interleukin-17	*IL17A*
112	Interleukin-18	*IL18*
113	Interleukin-2	*IL2*
114	Interleukin-2 receptor alpha	*IL2RA*
115	Interleukin-23	*IL23A,IL12B*
116	Interleukin-6	*IL6*
117	Interleukin-6 receptor	*IL6R*
118	Interleukin-6 receptor subunit beta	*IL6ST*
119	Interleukin-7	*IL7*
120	Interleukin-8	*CXCL8*
121	Kallikrein 5	*KLK5*
122	Kidney Injury Molecule-1	*HAVCR1*
123	Lactoferrin	*LTF*
124	Lactoylglutathione lyase	*GLO1*
125	Latency-Associated Peptide of Transforming Growth Factor beta 1	*LTBP1*
126	Lectin-Like Oxidized LDL Receptor 1	*OLR1*
127	Leptin	*LEP*
128	Leptin Receptor	*LEPR*
129	Leucine-rich alpha-2-glycoprotein	*LRG1*
130	Luteinizing Hormone	*LHB,CGA*
131	Macrophage Colony-Stimulating Factor 1	*CSF1*
132	Macrophage inflammatory protein 3 beta	*CCL19*
133	Macrophage Inflammatory Protein-1 alpha	*CCL3*

134	Macrophage Inflammatory Protein-1 beta	*CCL4*
135	Macrophage Inflammatory Protein-3 alpha	*CCL20*
136	Macrophage Migration Inhibitory Factor	*MIF*
137	Macrophage-Derived Chemokine	*CCL22*
138	Macrophage-Stimulating Protein	*MST1*
139	Matrix Metalloproteinase-1	*MMP1*
140	Matrix Metalloproteinase-10	*MMP10*
141	Matrix Metalloproteinase-3	*MMP3*
142	Matrix Metalloproteinase-7	*MMP7*
143	Matrix Metalloproteinase-9	*MMP9*
144	Matrix Metalloproteinase-9 total	*MMP9*
145	Mesothelin	*MSLN*
146	Methylglyoxal	*NA*
147	MHC class I chain-related protein A	*MICA*
148	Monocyte Chemotactic Protein 1	*CCL2*
149	Monocyte Chemotactic Protein 2	*CCL8*
150	Monocyte Chemotactic Protein 3	*CCL7*
151	Monocyte Chemotactic Protein 4	*CCL13*
152	Monokine Induced by Gamma Interferon	*CXCL9*
153	Myeloid Progenitor Inhibitory Factor 1	*CCL23*
154	Myeloperoxidase	*MPO*
155	Myoglobin	*MB*
156	N-terminal prohormone of brain natriuretic peptide	*NPPB*
157	Neuronal Cell Adhesion Molecule	*NRCAM*
158	Neuropilin-1	*NRP1*
159	Neutrophil Activating Peptide 2	*PPBP*
160	Neutrophil Gelatinase-Associated Lipocalin	*LCN2*
161	Omentin	*ITLN1*
162	Osteocalcin	*BGLAP*
163	Osteopontin	*SPP1*
164	Osteoprotegerin	*TNFRSF11B*
165	P-Selectin	*SELP*
166	Pancreatic Polypeptide	*PPY*
167	Paraoxanase-1	*PON1*
168	Pentraxin-3	*PTX3*
169	Pepsinogen I	*NA*
170	Peptide YY	*PYY*
171	Periostin	*POSTN*
172	Peroxiredoxin-4	*PRDX4*
173	Phosphoserine Aminotransferase	*PSAT1*
174	Pigment Epithelium Derived Factor	*SERPINF1*
175	Plasminogen Activator Inhibitor 1	*SERPINE1*
176	Platelet-Derived Growth Factor BB	*PDGFB*
177	Progesterone	*NA*
178	Progranulin	*GRN*
179	Proinsulin Intact	*INS*
180	Proinsulin Total	*INS*
181	Prolactin	*PRL*
182	Prostasin	*PRSS8*
183	Prostatic Acid Phosphatase	*ACPP*
184	Protein S100-A4	*S100A4*
185	Protein S100-A6	*S100A6*
186	Pulmonary and Activation-Regulated Chemokine	*CCL18*
187	Receptor for advanced glycosylation end products	*AGER*
188	Receptor tyrosine-protein kinase erbB-3	*ERBB3*
189	Resistin	*RETN*
190	Retinol-binding protein 4	*RBP4*
191	Secreted frizzled-related protein 4	*SFRP4*
192	Selenoprotein P	*SEPP1*
193	Serotransferrin	*TF*
194	Serum Amyloid A Protein	*SAA1,SAA2,SAA4,SAA3P*
195	Serum Amyloid P-Component	*APCS*
196	Serum Glutamic Oxaloacetic Transaminase	*GOT1,GOT2*
197	Sex Hormone-Binding Globulin	*SHBG*
198	Sortilin	*SORT1*
199	ST2	*IL1RL1*
200	Stem Cell Factor	*KITLG*
201	Stromal cell-derived factor-1	*CXCL12*
202	Superoxide Dismutase 1 soluble	*SOD1*
203	T Lymphocyte-Secreted Protein I-309	*CCL1*

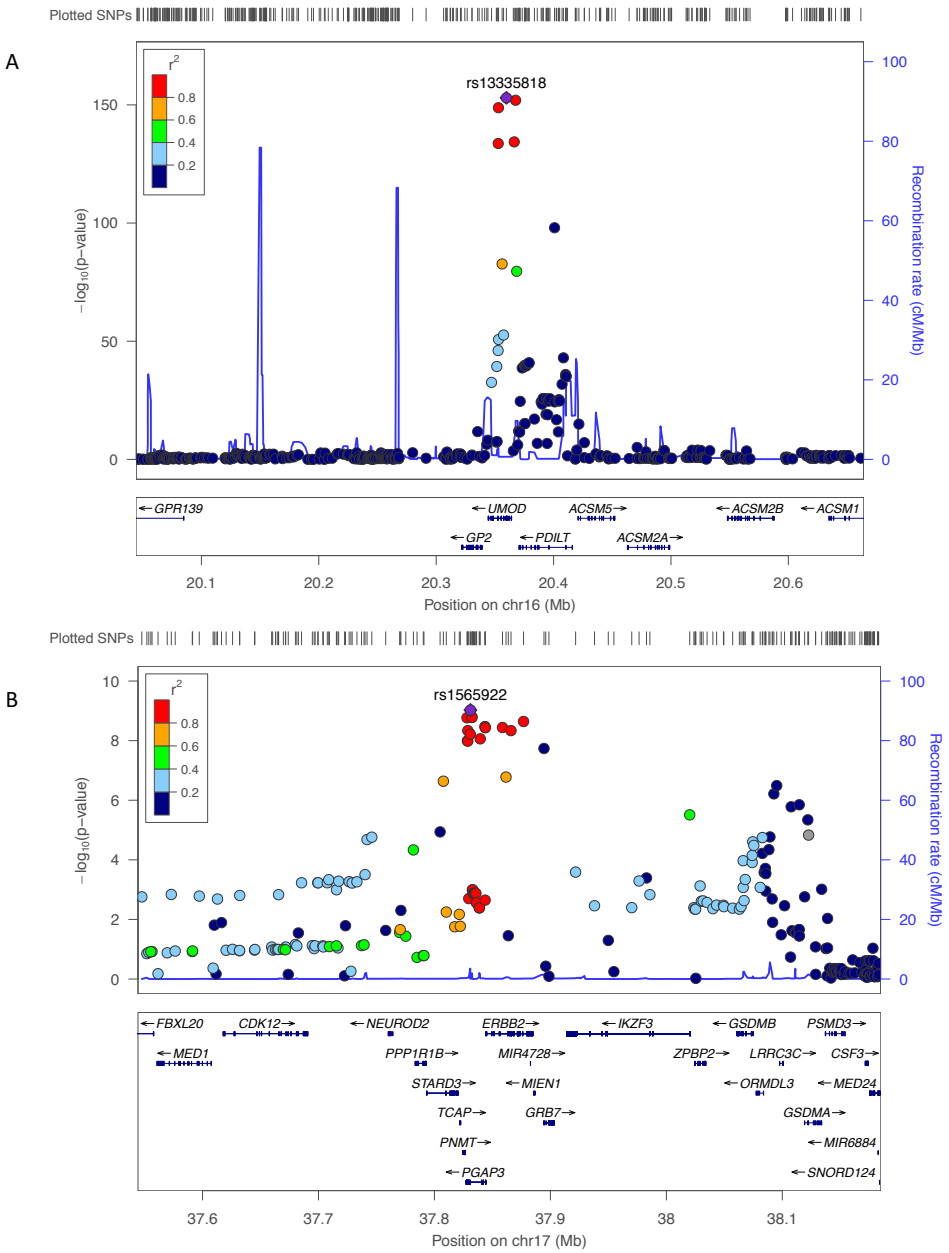
204	T-Cell-Specific Protein RANTES	*CCL5*
205	Tamm-Horsfall Urinary Glycoprotein (Uromodoulin)	*UMOD*
206	Tenascin-C	*TNC*
207	Testosterone Total	*NA*
208	Tetranectin	*CLEC3B*
209	Thrombin-activable fibrinolysis inhibitor	*CPB2*
210	Thrombomodulin	*THBD*
211	Thrombospondin-1	*THBS1*
212	Thyroid-Stimulating Hormone	*TSHB,CGA*
213	Thyroxine-Binding Globulin	*SERPINA7*
214	Tissue Inhibitor of Metalloproteinases 1	*TIMP1*
215	Tissue type Plasminogen activator	*PLAT*
216	TNF-Related Apoptosis-Inducing Ligand Receptor 3	*TNFRSF10C*
217	Transthyretin	*TTR*
218	Trefoil Factor 3	*TFF3*
219	Tumor Necrosis Factor alpha	*TNF*
220	Tumor necrosis factor receptor 2	*TNFRSF1B*
221	Tumor Necrosis Factor Receptor I	*TNFRSF1A*
222	Tyrosine kinase with Ig and EGF homology domains 2	*TIE1*
223	Urokinase-type Plasminogen Activator	*PLAU*
224	Urokinase-type plasminogen activator receptor	*PLAUR*
225	Vascular Cell Adhesion Molecule-1	*VCAM1*
226	Vascular Endothelial Growth Factor	*VEGFA*
227	Vascular Endothelial Growth Factor C	*VEGFC*
228	Vascular endothelial growth factor D	*FIGF*
229	Vascular Endothelial Growth Factor Receptor 2	*FLT1*
230	Vascular endothelial growth factor receptor 3	*FLT4*
231	Visceral adipose tissue derived serpin A12	*SERPINA12*
232	Visfatin	*NAMPT*
233	Vitamin D-Binding Protein	*GC*
234	Vitamin K-Dependent Protein S	*PROS1*
235	Vitronectin	*VTN*
236	von Willebrand Factor	*VWF*
237	YKL-40	*CHI3L1*

Supplementary Table 6: Participant characteristics for the genetic and biomarker sub-study subsets of the ORIGIN study.

Variable	Genetic Study participants (n=4,147)	All biomarker participants (n=8,197)
Age (years), mean (SD)	63.45 (7.98)	63.72 (7.94)
Gender (% male)	64.14	66.11
Ethnicity (%)		
European	46.56	55.41
Black	0	4.36
South Asian	0	5.49
South East Asian	0	0.46
Latin	53.44	34.28
Current smoker (% yes)	55.79	57.79
LDL (mmol/L)	3.07 (1.05)	2.89 (1.03)
HDL (mmol/L)	1.17 (0.32)	1.18 (0.32)
Fasting plasma glucose (mmol/L)	7.58 (2.17)	7.33 (2.02)
Hypertension (% yes)	82.90	78.91
Hypercholesterolemia (% yes)[a]	8.85	6.8
Body mass index (kg/m^2)	30.45 (5.33)	30.04 (5.27)
Prior diabetes (% yes)	87.56	81.66
EGFR (mL/min/1.73 m^2)	75.91 (21.07)	77.51 (21.86)
Prior renal disease (% yes)	6.75	5.83
Prospective CKD (% with event)	21.41	20.47

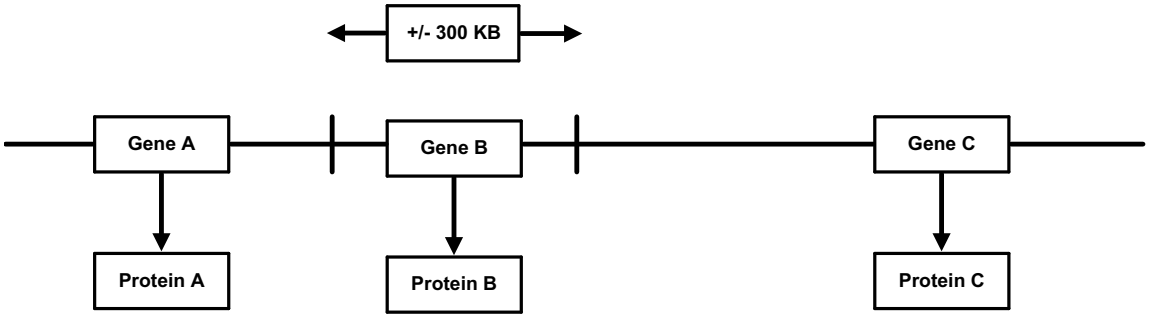
Data are presented as mean (SD) unless stated otherwise.

Supplementary Figure 1: Regional plots for UMOD and HER2 associations.

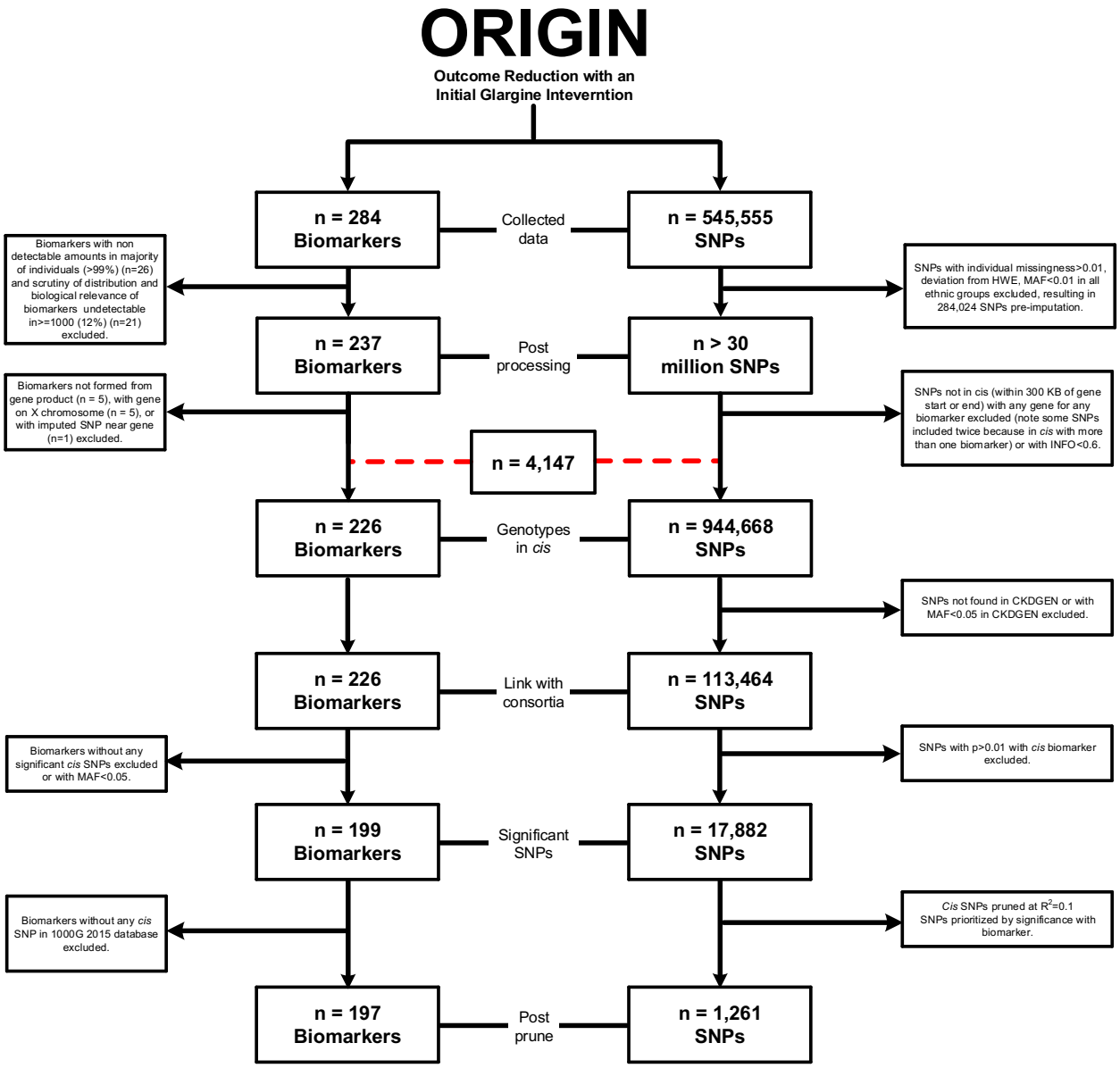


Plots show association of SNPs prior to MR instrument selection with serum UMOD levels at the *UMOD* locus (A) and serum HER2 levels at the *ERBB2* locus (B) +/- 300 KB along with recombination rates. $-\log_{10} P$ values (y axis) of the SNPs are shown according to their chromosomal positions (x axis). The most significant SNP in the analysis is labeled as a purple triangle. The color intensity of each symbol reflects the extent of LD with the top SNP, colored red ($r^2 > 0.8$) through to blue ($r^2 < 0.2$). SNPs with missing LD information are labeled grey. Genetic recombination rates (cM/Mb), estimated using 1000 Genomes European samples, are shown with a light blue line. Physical positions are based on build hg19 of the human genome. Also shown are the relative positions of genes mapping to the region of association.

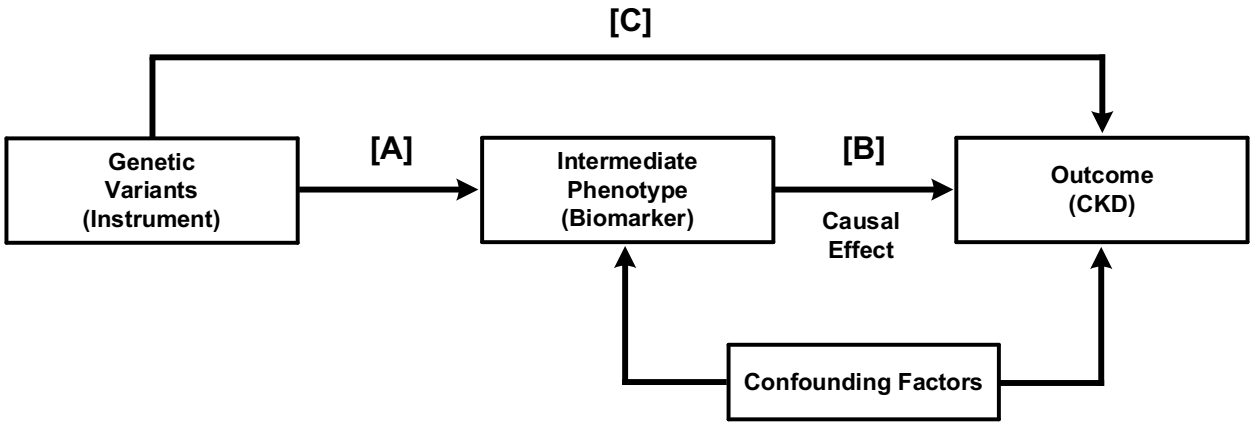
Supplementary Figure 2: Schematic representation of *cis* association.



Supplementary Figure 3: Overview of SNP and biomarker selection.



Supplementary Figure 4: Schematic representation of the instrumental variable assumptions of Mendelian randomization study. Instrumental variable analyses use associations of A and B to estimate the causal effect of an intermediate phenotype (biomarker) on an outcome (CKD), represented by association C.



REFERENCES

1. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M a R, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575.

2. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82. Available from: http://dx.doi.org/10.1016/j.ajhg.2010.11.011

3. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, Gabriel SB, Gibbs RA, Green ED, Hurles ME, Knoppers BM, Korbel JO, Lander ES, Lee C, Lehrach H, Mardis ER, Marth GT, McVean GA, Nickerson DA, Schmidt JP, Sherry ST, Wang J, Wilson RK, Boerwinkle E, Doddapaneni H, Han Y, Korchina V, Kovar C, Lee S, Muzny D, Reid JG, Zhu Y, Chang Y, Feng Q, Fang X, Guo X, Jian M, Jiang H, Jin X, Lan T, Li G, Li J, Li Y, Liu S, Liu X, Lu Y, Ma X, Tang M, Wang B, Wang G, Wu H, Wu R, Xu X, Yin Y, Zhang D, Zhang W, Zhao J, Zhao M, Zheng X, Gupta N, Gharani N, Toji LH, Gerry NP, Resch AM, Barker J, Clarke L, Gil L, Hunt SE, Kelman G, Kulesha E, Leinonen R, McLaren WM, Radhakrishnan R, Roa A, Smirnov D, Smith RE, Streeter I, Thormann A, Toneva I, Vaughan B, Zheng-Bradley X, Grocock R, Humphray S, James T, Kingsbury Z, Sudbrak R, Albrecht MW, Amstislavskiy VS, Borodina TA, Lienhard M, Mertes F, Sultan M, Timmermann B, Yaspo M-L, Fulton L, Fulton R, Ananiev V, Belaia Z, Beloslyudtsev D, Bouk N, Chen C, Church D, Cohen R, Cook C, Garner J, Hefferon T, Kimelman M, Liu C, Lopez J, Meric P, O'Sullivan C, Ostapchuk Y, Phan L, Ponomarov S, Schneider V, Shekhtman E, Sirotkin K, Slotta D, Zhang H, Balasubramaniam S, Burton J, Danecek P, Keane TM, Kolb-Kokocinski A, McCarthy S, Stalker J, Quail M, Davies CJ, Gollub J, Webster T, Wong B, Zhan Y, Campbell CL, Kong Y, Marcketta A, Yu F, Antunes L, Bainbridge M, Sabo A, Huang Z, Coin LJM, Fang L, Li Q, Li Z, Lin H, Liu B, Luo R, Shao H, Xie Y, Ye C, Yu C, Zhang F, Zheng H, Zhu H, Alkan C, Dal E, Kahveci F, Garrison EP, Kural D, Lee W-P, Fung Leong W, Stromberg M, Ward AN, Wu J, Zhang M, Daly MJ, DePristo MA, Handsaker RE, Banks E, Bhatia G, del Angel G, Genovese G, Li H, Kashin S, McCarroll SA, Nemesh JC, Poplin RE, Yoon SC, Lihm J, Makarov V, Gottipati S, Keinan A, Rodriguez-Flores JL, Rausch T, Fritz MH, Stütz AM, Beal K, Datta A, Herrero J, Ritchie GRS, Zerbino D, Sabeti PC, Shlyakhter I, Schaffner SF, Vitti J, Cooper DN, Ball E V., Stenson PD, Barnes B, Bauer M, Keira Cheetham R, Cox A, Eberle M, Kahn S, Murray L, Peden J, Shaw R, Kenny EE, Batzer MA, Konkel MK, Walker JA, MacArthur DG, Lek M, Herwig R, Ding L, Koboldt DC, Larson D, Ye K, Gravel S, Swaroop A, Chew E, Lappalainen T, Erlich Y, Gymrek M, Frederick Willems T, Simpson JT, Shriver MD, Rosenfeld JA, Bustamante CD, Montgomery SB, De La Vega FM, Byrnes JK, Carroll AW, DeGorter MK, Lacroute P, Maples BK, Martin AR, Moreno-Estrada A, Shringarpure SS, Zakharia F, Halperin E, Baran Y, Cerveira E, Hwang J, Malhotra A, Plewczynski D, Radew K, Romanovitch M, Zhang C, Hyland FCL, Craig DW, Christoforides A, Homer N, Izatt T, Kurdoglu AA, Sinari SA, Squire K, Xiao C, Sebat J, Antaki D, Gujral M, Noor A, Ye K, Burchard EG, Hernandez RD, Gignoux CR, Haussler D, Katzman SJ, James Kent W, Howie B, Ruiz-Linares A, Dermitzakis ET, Devine SE, Min Kang H, Kidd JM, Blackwell T, Caron S, Chen W, Emery S, Fritsche L, Fuchsberger C, Jun G, Li B, Lyons R, Scheller C, Sidore C, Song S, Sliwerska E, Taliun D, Tan A, Welch R, Kate Wing M, Zhan X, Awadalla P, Hodgkinson A, Li Y, Shi X, Quitadamo A, Lunter G, Marchini JL, Myers S, Churchhouse C, Delaneau O, Gupta-Hinch A, Kretzschmar W, Iqbal Z, Mathieson I, Menelaou A, Rimmer A, Xifara DK, Oleksyk TK, Fu Y, Liu X, Xiong M, Jorde L, Witherspoon D, Xing J, Browning BL, Browning SR, Hormozdiari F, Sudmant PH, Khurana E, Tyler-Smith C, Albers CA, Ayub Q, Chen Y, Colonna V, Jostins L, Walter K, Xue Y, Gerstein MB, Abyzov A, Balasubramanian S, Chen J, Clarke D, Fu Y, Harmanci AO, Jin M, Lee D, Liu J, Jasmine Mu X, Zhang J, Zhang Y, Hartl C, Shakir K, Degenhardt J, Meiers S, Raeder B, Paolo Casale F, Stegle O, Lameijer E-W, Hall I, Bafna V, Michaelson J, Gardner EJ, Mills RE, Dayama G, Chen K, Fan X, Chong Z, Chen T, Chaisson MJ, Huddleston J, Malig M, Nelson BJ, Parrish NF, Blackburne B, Lindsay SJ, Ning Z, Zhang Y, Lam H, Sisu C, Challis D, Evani US, Lu J, Nagaswamy U, Yu J, Li W, Habegger L, Yu H, Cunningham F, Dunham I, Lage K, Berg Jespersen J, Horn H, Kim D, Desalle R, Narechania A, Wilson Sayres MA, Mendez FL, David Poznik G, Underhill PA, Coin L, Mittelman D, Banerjee R, Cerezo M, Fitzgerald TW, Louzada S, Massaia A, Ritchie GR, Yang F, Kalra D, Hale W, Dan X, Barnes KC, Beiswanger C, Cai H, Cao H, Henn B, Jones D, Kaye JS, Kent A, Kerasidou A, Mathias R, Ossorio PN, Parker M, Rotimi CN, Royal CD, Sandoval K, Su Y, Tian Z, Tishkoff S, Via M, Wang Y, Yang H, Yang L, Zhu J, Bodmer W, Bedoya G, Cai Z, Gao Y, Chu J, Peltonen L, Garcia-Montero A, Orfao A, Dutil J,

Martinez-Cruzado JC, Mathias RA, Hennis A, Watson H, McKenzie C, Qadri F, LaRocque R, Deng X, Asogun D, Folarin O, Happi C, Omoniwa O, Stremlau M, Tariyal R, Jallow M, Sisay Joof F, Corrah T, Rockett K, Kwiatkowski D, Kooner J, Tịnh Hiê`n T, Dunstan SJ, Thuy Hang N, Fonnie R, Garry R, Kanneh L, Moses L, Schieffelin J, Grant DS, Gallo C, Poletti G, Saleheen D, Rasheed A, Brooks LD, Felsenfeld AL, McEwen JE, Vaydylevich Y, Duncanson A, Dunn M, Schloss JA. A global reference for human genetic variation. Nature 2015;526:68–74. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26432245

4. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet 2012;44:955–959. Available from: http://www.nature.com/ng/journal/v44/n8/full/ng.2354.html%5Cnhttp://www.nature.com/ng/journal/v44/n8/pdf/ng.2354.pdf

5. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009;5:e1000529.

6. Paré G, Asma S, Deng WQ. Contribution of large region joint associations to complex traits genetics. PLoS Genet 2015;11:e1005103. Available from: http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1005103

7. Hsu F, Kent JW, Clawson H, Kuhn RM, Diekhans M, Haussler D. The UCSC known genes. Bioinformatics 2006;22:1036–1046.

8. Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. Bioinformatics 1998;14:656–664.