Genome-Wide Association Study of Diabetic Kidney Disease Highlights Biology Involved in Glomerular Basement Membrane Collagen

SUMMIT Consortium, DCCT/EDIC Research Group, GENIE Consortium, Kretzler, M., Susztak, K., Colhoun, H. M., Krolewski, A. S., Paterson, A. D., ... Florez, J. C. (2019). Genome-Wide Association Study of Diabetic Kidney Disease Highlights Biology Involved in Glomerular Basement Membrane Collagen. Journal of the American Society of Nephrology, 30(10), 2000-2016. https://doi.org/10.1681/ASN.2019030218

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Genome-wide association study of diabetic kidney disease highlights biology involved in renal basement membrane collagen

Rany M. Salem[1*], Jennifer N. Todd[2, 3, 4*], Niina Sandholm[5, 6, 7*], Joanne B. Cole[2, 3, 4*], Wei-Min Chen[8], Darrel Andrews[9], Marcus G. Pezzolesi[10], Paul M. McKeigue[11], Linda T. Hiraki[12], Chengxiang Qiu[13], Viji Nair[14], Chen Di Liao[12], Jing Jing Cao[12], Erkka Valo[5, 6, 7], Suna Onengut-Gumuscu[8], Adam M. Smiles[15], Stuart J. McGurnaghan[16], Jani K. Haukka[5, 6, 7], Valma Harjutsalo[5, 6, 7,17], Eoin P. Brennan[9], Natalie van Zuydam[18, 19], Emma Ahlqvist[20], Ross Doyle[9], Tarunveer S. Ahluwalia[21], Maria Lajer[21], Maria F. Hughes[9], Jihwan Park[13], Jan Skupien[15], Athina Spiliopoulou[11], Andrew Liu[22], Rajasree Menon[14, 23], Carine M. Boustany-Kari[24], Hyun M. Kang[23, 25], Robert G. Nelson[26], Ronald Klein[27], Barbara E. Klein[27], Kristine E. Lee[27], Xiaoyu Gao[28], Michael Mauer[29], Silvia Maeastroni[30], Maria Luiza Caramori[29], Ian H. de Boer[31], Rachel G. Miller[32], Jingchuan Guo[32], Andrew P. Boright[12], David Tregouet[33, 34], Beata Gyorgy[33, 34], Janet K. Snell-Bergeon[35], David M. Maahs[36], Shelley B. Bull[37], Angelo J. Canty[38], Colin N.A. Palmer[39], Lars Stechemesser[40], Bernhard Paulweber[40], Raimund Weitgasser[41], Jelizaveta Sokolovska[42, 43], Vita Rovīte[44], Valdis Pīrāgs[42, 43], Edita Prakapiene[45], Lina Radzeviciene[46], Rasa Verkauskiene[46], Nicolae Mircea Panduru[6, 47], Leif C. Groop[48, 49], Mark I. McCarthy[18, 19, 50], Harvest F. Gu[51, 52], Anna Möllsten[53], Henrik Falhammar[54, 55], Kerstin Brismar[54, 55], GENIE Consortium, DCCT/EDIC Research Group, SUMMIT Consortium, Finian Martin[9], Peter Rossing[21, 56], Tina Costacou[32], Gianpaolo Zerbini[30], Michel Marre[57, 58, 59], Samy Hadjadj[60, 61], Amy J. McKnight[62], Carol Forsblom[5, 6], Gareth McKay[62], Catherine Godson[9], A. Peter Maxwell[62], Matthias Kretzler[14, 23], Katalin Susztak[13], Helen M. Colhoun[16], Andrzej Krolewski[15], Andrew D. Paterson[12], Per-Henrik Groop[5, 6, 7, 63], Stephen S. Rich[8], Joel N. Hirschhorn[2, 3], Jose C. Florez[3, 4, 64, 65]

AFFILIATIONS

1.	Family Medicine and Public Health, University of California San Diego, La Jolla, CA 92093 USA.
2.	Division of Endocrinology/ Department of Pediatrics, Boston Children's Hospital, Boston, MA 02115 USA.
3.	Programs in Metabolism and Medical and Population Genetics, Broad Institute, Cambridge, MA 02142 USA.
4.	Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA 02114 USA.
5.	Folkhälsan Institute of Genetics, Folkhälsan Research Center, Helsinki, Finland.
6.	Abdominal Center Nephrology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.
7.	Research Programs Unit, Diabetes and Obesity, University of Helsinki, Helsinki, Finland.
8.	University of Virginia, 3812B OMS, Charlottesville, VA 22908 USA.
9.	Diabetes Complications Research Centre/ Conway Institute, School of Medicine and Medical Sciences, University College Dublin, Dublin 4 Ireland.
10.	University of Utah Division of Nephrology and Hypertension Diabetes and Metabolism Center, University of Utah, Salt Lake City, UT 84132 USA.
11.	Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH16 4TJ UK.
12.	The Hospital for Sick Children, Toronto, ON M5G 0A4 Canada.
13.	Department of Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA.
14.	Division of Nephrology, Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109, USA.
15.	Joslin Diabetes Center, Boston, MA 02115 USA.
16.	Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU UK.
17.	The Chronic Disease Prevention Unit, National Institute for Health and Welfare, 00271, Helsinki, Finland.
18.	Wellcome Trust Centre for Human Genetics, University of Oxford, OX37BN, UK.
19.	Oxford Centre for Diabetes, Endocrinology and Metabolism, Oxford, OX37BN, UK.
20.	Lund University Diabetes Centre, Malmö, 20502, Sweden.
21.	Steno Diabetes Center Copenhagen, Niels Steensens Vej 2, DK 2820 Gentofte, Denmark.
22.	Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, MI, 48109 USA.
23.	Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA.
24.	Cardiometabolic Diseases Research, Boehringer Ingelheim, Ridgefield, CT 06877, USA.
25.	Center for Statistical Genetics, University of Michigan School of Public Health, Ann Arbor, MI, 48109 USA.
26.	Chronic Kidney Disease Section, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Health, Phoenix, AZ 85014, USA.
27.	University of Wisconsin-Madison, Madison, WI 53726 USA.
28.	The George Washington University, Washington, DC 20052 USA.
29.	University of Minnesota, Minneapolis, MN 55454 USA.
30.	Complications of Diabetes Unit, Diabetes Research Institute (DRI), Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Via Olgettina 60, 20132 Milano, Italy.
31.	University of Washington, Seattle, WA 98195 USA.
32.	University of Pittsburgh Public Health, Pittsburgh, PA, 15213, USA.
33.	Sorbonne Université, UPMC Univ Paris 06, INSERM UMR_S 1166, Paris, France.

34. ICAN Institute for Cardiometabolism and Nutrition, Paris, France.
35. University of Colorado School of Medicine, Aurora, CO 80045 USA.
36. Department of Pediatrics-Endocrinology, Stanford University, Stanford, California 94305 USA.
37. The Lunenfeld-Tenenbaum Research Institute and University of Toronto, Toronto, ON M5G 0A4 Canada.
38. Dept. of Mathematics & Statistics, McMaster University, Hamilton, Ontario, L8S 4K1 Canada.
39. Pat Macpherson Centre for Pharmacogenetics and Pharmacogenomics, Ninewells Hospital and Medical School, University of Dundee, DD19SY UK.
40. First Department of Medicine, Paracelsus Medical University, Müllner Hauptstrasse 48, 5020 Salzburg, Austria.
41. Department of Medicine, Diakonissen-Wehrle Hospital, Salzburg, Austria.
42. University of Latvia, Raiņa 19, Riga, Latvia.
43. Pauls Stradins University Hospital, Pilsonu 13, Riga, Latvia.
44. Latvian Biomedical Research and Study Centre, Rātsupītes 1, Riga, Latvia.
45. Medical Academy, Lithuanian University of Health Sciences, Kaunas, LT 44307 Lithuania.
46. Institute of Endocrinology, Medical Academy, Lithuanian University of Health, Kaunas 44307 Lithuania Sciences.
47. 2nd Clinical Department, Carol Davila University of Medicine and Pharmacy, Bucharest, 020021 Romania.
48. Department of Clinical Sciences, Diabetes and Endocrinology, Lund University Diabetes Centre, Malmö, 205 02 Sweden.
49. Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, C329b, Finland.
50. Oxford NIHR Biomedical Research Centre, Oxford University Hospitals Trust, Oxford, OX37BN, UK.
51. Department of Clinical Science, Intervention and Technology, Karolinska University Hospital, Huddinge, Stockholm 14152 Sweden.
52. School of Basic Medicine and Clinical Pharmacy, China Pharmaceutical University, Nanjing 210009 China.
53. Umeå University, Department of Clinical Sciences, Pediatrics, 90187 Umeå, Sweden.
54. Rolf Luft Center for Diabetes Research and Endocrinology, Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm 17176 Sweden.
55. Department of Endocrinology, Diabetes and Metabolism, Karolinska University Hospital, Stockholm, 17176 Sweden.
56. University of Copenhagen, Copenhagen, Denmark.
57. Assistance Publique – Hôpitaux de Paris (AP-HP), Bichat Hospital, DHU FIRE, Department of Diabetology, Endocrinology and Nutrition, Paris, France.
58. Paris Diderot University, Sorbonne Paris Cité, UFR de Médecine, Paris, France; INSERM, UMRS 1138, Cordeliers Research Center, Paris, France.
59. Fondation Ophtalmologique Adolphe de Rothschild, Paris, France.
60. Centre Hospitalier Universitaire de Poitiers, Department of Endocrinology and Diabetology, Poitiers, France; INSERM CIC 1402, Poitiers, France.
61. L'institut du thorax, INSERM, CNRS, CHU Nantes, Nantes, France.
62. Centre for Public Health, Queens University of Belfast, Northern Ireland, BT7 1NN UK.
63. Department of Diabetes, Central Clinical School, Monash University, Melbourne, Victoria, Australia.
64. Diabetes Unit, Massachusetts General Hospital, Boston MA 02114 USA.
65. Department of Medicine, Harvard Medical School, Boston MA 02115

* These authors contributed equally

Corresponding author: Jose C. Florez jcflorez@mgh.harvard.edu

Lead author: Jose C. Florez

SIGNIFICANCE STATEMENT

Diabetic kidney disease (DKD) is a devastating microvascular complication of both type 1 and type 2 diabetes. It has been shown to have a heritable component, but prior searches for the genetic determinants of this condition have had limited success. In this study, a new international genomics consortium (the Diabetic Nephropathy Collaborative Research Initiative, funded by the JDRF) has coalesced to assemble nearly 20,000 samples from participants with type 1 diabetes, with and without kidney disease. We report 16 new signals at genome-wide significance and begin to describe the genetic architecture of this globally relevant phenotype. Our top signal centers on a protective missense coding variant at *COL4A3*, which encodes an integral component of the glomerular basement membrane that when mutated causes Alport syndrome.

ABSTRACT

Background: Diabetic kidney disease (DKD) is a heritable but poorly understood complication of diabetes. **Methods**: To identify genetic variants predisposing to DKD, we performed genome-wide association analyses in 19,406 individuals with type 1 diabetes (T1D) using a spectrum of DKD definitions based on albuminuria and renal function. **Results**: We identified 16 genome-wide significant loci. The variant with the strongest association (rs55703767) is a common missense mutation in the collagen type IV alpha 3 chain (*COL4A3)* gene*,* which encodes a major structural component of the glomerular basement membrane (GBM) implicated in heritable nephropathies. The rs55703767 minor allele (Asp326Tyr) is protective against several definitions of DKD, including albuminuria and end-stage renal disease. Three other loci are in or near genes with known or suggestive involvement in DKD (*BMP7)* or renal biology (*COLEC11* and *DDR1*). **Conclusion**: The 16 DKD-associated loci provide novel insights into the pathogenesis of DKD, identifying potential biological targets for prevention and treatment.

INTRODUCTION

The devastating diabetic complication of diabetic kidney disease (DKD) is the major cause of end-stage renal disease (ESRD) worldwide[1,2]. Current treatment strategies at best slow the progression of DKD, and do not halt or reverse the disease. Although improved glycemic control influences the rate of diabetic complications, a large portion of the variation in DKD susceptibility remains unexplained: one third of people with type 1 diabetes (T1D) develop DKD despite adequate glycemic control, while others maintain normal renal function despite long-term severe chronic hyperglycemia[3].

Though DKD demonstrates both familial clustering[4-6] and single nucleotide polymorphism (SNP) heritability[7], the specific genetic factors influencing DKD risk remain largely unknown. Recent genome-wide association studies (GWAS) have only identified a handful of loci for DKD, albuminuria, or estimated glomerular filtration rate (eGFR) in individuals with diabetes[7-13]. Potential reasons for the limited success include small sample sizes, modest genetic effects, and lack of consistency of phenotype definitions and statistical analyses across studies. Through collaboration within the JDRF Diabetes Nephropathy Collaborative Research Initiative (DNCRI), we adopted three approaches to improve our ability to find new genetic risk factors for DKD: 1) assembling a large collection of T1D cohorts with harmonized DKD phenotypes, 2) creating a comprehensive set of detailed DKD definitions, and 3) augmenting genotype data with low frequency and exome array variants.

METHODS

Cohorts and Phenotype Definitions. The GWAS meta-analysis included up to 19,406 patients with T1D of European origin from 17 cohorts (for study list and details see **Table S1**). All participants gave informed consent and all studies were approved by ethics committees from

participating institutions. We defined a total of 10 different case-control outcomes to cover the different aspects of renal complications, using both albuminuria and eGFR (**Figure 1**). Five comparisons ("All vs. ctrl", "Micro", "DN", "Macro", and "ESRD vs. macro") were based on albuminuria, measured by albumin excretion rate (AER) from overnight or 24-h urine collection, or by albumin creatinine ratio (ACR). Two out of three consecutive collections were required (when available) to classify the renal status of subjects as either normoalbuminuria, microalbuminuria, macroalbuminuria, or ESRD; for detailed thresholds, see **Table S9**. Controls with normal AER were required to have a minimum diabetes duration of 15 years; subjects with microalbuminuria/macroalbuminuria/ESRD were required to have minimum diabetes duration of 5/10/10 years, respectively, to exclude renal complications of non-diabetic origins. Two comparisons ("ESRD vs. ctrl" and "ESRD vs. non-ESRD") were based on presence of ESRD as defined by eGFR<15 mL/min or dialysis or renal transplant. Two phenotypes ("CKD" and "CKD extreme") were defined based on eGFR estimated by the CKD-EPI formula: controls had eGFR ≥60ml/min/1.73m^2 for both phenotypes, and ≥15 years of diabetes duration; cases had eGFR <60ml/min/1.73m^2 for the "CKD" phenotype, and eGFR <15 ml/min/1.73m^2 or dialysis or renal transplant for the "CKD extreme" phenotype, and ≥10 years of diabetes duration. For the "CKD-DN" phenotype that combined both albuminuria and eGFR data, controls were required to have both eGFR ≥60ml/min/1.73m^2 and normoalbuminuria; cases had both eGFR <45ml/min/1.73m^2 and micro- or macroalbuminuria, or ESRD.

GWAS Genotyping, Quality Control and Imputation. All study samples underwent genotyping, quality control (QC) and imputation centrally at the University of Virginia. In brief, samples were genotyped on the HumanCore BeadChip array (Illumina, San Diego, CA, USA), which contains ~250,000 genome-wide tag SNPs and over 200,000 exome-focused variants. All samples were passed through a stringent QC protocol. Following initial genotype calling with Illumina software, all samples were re-called with zCall, a calling algorithm specifically designed for rare SNPs from

arrays. Variant orientation and position were aligned to hg19 (Genome Reference Consortium Human Build 37, GRCh37). Variant names were updated using 1000 Genomes as a reference. The data were then filtered for low quality variants (e.g. call rates <95% and excessive deviation from Hardy-Weinberg equilibrium) and samples (e.g. call rates <98%, gender mismatch, extreme heterozygosity). Principal Component Analysis (PCA) was performed separately for each cohort in order to empirically detect and exclude outliers with evidence of non-European ancestry (see supplement for full QC details, and **Figure S1** for trait specific Manhattan and QQ plots). Genotypes were expanded to a total of approximately 49 million by imputation, using the minimac imputation tool[14,15] and 1,000 Genomes Project (phase 3v5) as a reference.

GWAS Analysis. A genome-wide association analysis was performed for each of the case-control definitions under an additive genetic model, adjusting for age, sex, diabetes duration, study site (where applicable) and principal components. We conducted a second set of analyses adjusting for BMI and HbA1c which we refer to as our fully adjusted covariate model. Allele dosages were used to account for imputation uncertainty. Inverse-variance fixed effects meta-analysis was performed using METAL and the following filters: INFO score >0.3, minor allele count >10 in both cases and controls, and presence of variant in at least two cohorts (Manhattan and QQ plots each trait and covariate model presented in **Figure S1**). The X chromosome was similarly analyzed for males and females both separately and in a combined analysis, with the exception of using hard call genotypes in place of allele dosages.[16] We estimated the percentage of variance explained for all genome-wide significant SNPs across all disease definitions using the McKelvey-Zavoina[17] pseudo-R^2 statistic predicting continuous latent variables underlying binary outcomes.

Glomerular basement membrane measurement in Renin-Angiotensin System Study (RASS). In brief, RASS was a double-blind placebo-controlled randomized trial of enalapril and losartan on renal pathology among 285 normoalbuminuric, normotensive subjects with T1D and

normal or increased measured glomerular filtration rate (>90 ml/min/1.73m^2).[18] Participants were followed for 5 years with percutaneous kidney biopsy completed prior to randomization and at 5 years. Structural parameters measured by electron microscopy on biopsy included GBM width, measured by the electron microscopic orthogonal intercept method[18]. All RASS participants contributed DNA for genotyping.

In silico **replication in SUMMIT consortium**. The SUMMIT consortium included up to 5,193 European Ancestry subjects with type 2 diabetes (T2D), with and without kidney disease. *In silico* replication was performed on previously published GWAS on DKD with harmonized trait definitions for seven of our primary T1D analyses: "DN", "Micro", "Macro", "ESRD", "ESRD vs. non-ESRD", "CKD", and "CKD-DN" under an additive model, adjusting for age, gender and duration of diabetes.[13]

RNAseq and microarray profiling of human kidney samples from the Pima cohort. Kidney biopsy samples from the Pima Indian cohort were manually micro-dissected into 119 glomerular and 100 tubule-interstitial tissues to generate gene expression profiles[19]. Expression profiling in the Pima Indian cohort kidney biopsies was carried out using Affymetrix GeneChip Human Genome U133 Array and U133Plus2 Array, as reported previously, and Affymetrix Human Gene ST Genechip 2.1[20,21], and on RNA-seq (Illumina). The libraries were prepared using the ClonTech SMARTSeq v4 Ultra Low Input polyA selection kit. Samples were sequenced on a HiSeq 4000, single end, 75bp. Mapping to human reference genome GRCh38.7 was performed with STAR 2.5.2b (https://github.com/alexdobin/STAR). For annotation and quantification of mapping results we used cufflinks, cuffquant and cuffnorm in version 2.2.1 (https://cole-trapnell-lab.github.io/cufflinks/). After mapping and quantification, PCA and Hierarchical Clustering was used to identify outliers and reiterated until no more outliers could be identified.

RNA-sequencing and cis-eQTL analysis in human kidney samples from University of Pennsylvania cohort. Human kidney samples were obtained from surgical nephrectomies for a

total of 455 subjects with pathological data and were manually microdissected under a microscope in RNAlater for glomerular and tubular compartments (433 tubule and 335 glomerulus samples). The local renal pathologist performed an unbiased review of the tissue section by scoring multiple parameters, and RNA were prepared using RNAeasy mini columns (Qiagen, Valencia, CA) according to manufacturer's instructions.

Whole kidney[22], tubular and glomerular[23] eQTL analyses have been described previously. Tubular and glomerular eQTL data sets were generated by 121 samples of tubules and 119 samples of glomeruli, respectively[23]. The cis window was defined as 1 megabase up- and down-stream of the transcriptional start site (±1Mb). Whole kidney cis-eQTL (further just referred to as eQTL) data set was generated from 96 human samples obtained from The Cancer Genome Atlas (TCGA) through the TCGA Data portal[22].

Mouse kidney single cell RNA-sequencing. Kidneys were harvested from 4 to 8-week-old male mice with C57BL/6 background and dissociated into single cell suspension as described in our previous study[24]. The single cell sequencing libraries were sequenced on an Illumina HiSeq with 2x150 paired-end kit. The sequencing reads were demultiplexed, aligned to the mouse genome (mm10) and processed to generate gene-cell data matrix using Cell Ranger 1.3 (http://10xgenomics.com)[24].

Genomic features of human kidney. Human kidney-specific chromatin immunoprecipitation followed by sequencing (ChIP-seq) data can be found at GEO: GSM621634, GSM670025, GSM621648, GSM772811, GSM621651, GSM1112806, GSM621638. Different histone markers were combined into chromatin states using ChromHMM[25].

Gene and gene set analysis. PASCAL and MAGMA (v1.06) gene and pathway scores were conducted on all 20 sets of GWAS summary statistics using default pathway libraries from BioCarta, REACTOME, and KEGG. MAGENTA (vs2, July 2011) pathway analysis included 4725 pathways with a minimum of five genes within the gene set for the 10 standard adjustment models.

We conducted DEPICT individually on all 20 sets of GWAS summary statistics with $P<10^{-5}$ and additional pooled analyses using genome-wide minimum P-values from: 1) All 20 analyses (10 phenotypes and 2 covariate models) and 2) 16 analyses of the 8 most related phenotypes which excluded ESRD vs Macro and Micro.

Transcriptome-wide association study (TWAS): TWAS of kidney glomeruli and tubuli was performed using MetaXcan with default parameters,[26] based on eQTL data for human glomerular and tubular cells.[23]

RESULTS

Phenotypic comparisons

We investigated a broad spectrum of DKD definitions based on albuminuria and renal function criteria, defining a total of 10 different case-control comparisons to cover the different aspects of disease progression (**Figure 1**). Seven comparisons were based on albuminuria and/or ESRD (including diabetic nephropathy [DN], defined as either macroalbuminuria or ESRD); two were defined based on eGFR (used to classify severity of chronic kidney disease [CKD]); and one combined both albuminuria and eGFR data ("CKD-DN"). Each phenotypic definition was analyzed separately in GWAS; to account for the 10 definitions each analyzed under two covariate adjustment models, we estimated[16] the total effectively independent tests as 7.4, allowing us to compute a more conservative study-wide significance threshold ($P<6.76\times10^{-9}$), based on genome-wide significance ($P<5\times10^{-8}$) and Bonferroni correction for 7.4 effective tests.

Top genome-wide association results highlight *COL4A3*

GWAS meta-analysis included association results for up to 19,406 individuals with T1D of European descent from 17 cohorts for the 10 case-control definitions **(Table S1)**. We identified 16 novel independent loci that achieved genome-wide significance ($P<5\times10^{-8}$) in either the minimal or fully adjusted models, in which four lead SNPs also surpassed our more conservative

study-wide significance threshold (**Table 1; Figure 2,** Manhattan plot**; Figures S2a-p**, regional association and forest plots). None of the loci reaching genome-wide significance have been previously identified in GWAS or candidate gene studies for DKD or closely related traits. All SNPs with minor allele frequency greater than 1% explain 2.5% and 3.0% of the total variance (McKelvey-Zavoina[17] pseudo-R^2) of DN after adjusting for covariates in the minimal and full covariate models, respectively (**Table S12**).

The strongest signal was rs55703767 (minor allele frequency [MAF]=0.21), a common missense variant (G>T; Asp326Tyr) in exon 17 of *COL4A3*. This SNP was associated with protection from DN (odds ratio [OR]=0.79, P=5.34×10^{-12}), any albuminuria (OR=0.84, P=3.88×10^{-10}), the combined CKD-DN phenotype (OR =0.77, P=5.30×10^{-9}), and macroalbuminuria (OR=0.79, P=9.28×10^{-9}). Interestingly, we found that rs55703767 in *COL4A3* was more strongly associated in men (OR=0.73, P=1.29×10^{-11}) than in women (OR=0.85, P=1.39×10^{-3}; P_{het}=1.58×10^{-2}). *COL4A3* encodes the alpha 3 chain of collagen type IV, a major structural component of the GBM[27].

COL4A3 variation and kidney phenotypes

In persons with T1D and normoalbuminuria, GBM width predicts progression to proteinuria and ESRD independently of glycated hemoglobin (HbA1c)[28]. We examined the influence of the *COL4A3* variant on GBM width in 253 RASS[18] participants with T1D and normal AER, eGFR (>90 ml/min/1.73 m^2) and blood pressure, who had biopsy and genetic data (**Table S2**). The DKD-protective minor T allele was associated with 19.7 nm lower GBM width (standard error (SE) 8.2 nm, P=0.02), with the lowest mean GBM width among TT homozygotes (**Figure 3; Table S3**), after adjusting for age, sex, and diabetes duration, and without detectable interactions with T1D duration or mean HbA1c. Thus, the protective T allele carriers had thinner GBM prior to any renal complications.

We did not detect any eQTL association between rs55703767 and *COL4A3* expression in mouse glomeruli or in human tissues, and thus assume that the variant affects the *COL4A3* structure rather than expression levels. Nevertheless in a Pima Indian cohort of 97 subjects with DKD with morphometric and expression data from renal biopsies, *COL4A3* expression was negatively correlated with the GBM surface density (filtration surface density) (β=-0.27, *P*=0.02), which is associated with eGFR in DKD in both T1D and T2D[29,30]. Furthermore, in 335 micro-dissected human glomerulus samples, expression of *COL4A3* was negatively correlated with glomerulosclerosis, potentially reflecting podocyte depletion in sclerotic glomeruli (corr=0.16, *P*=4.8×10^{-3}; **Figure S3**). *COL4A3* expression in glomeruli, but not in tubules, was also nominally correlated with eGFR (corr 0.108, *P*=0.047; **Figure S3).**

Evidence for hyperglycemia specificity

Hyperglylcemia promotes the development of diabetic complications. If a genetic variant exerts a stronger effect in the setting of hyperglycemia, 1) it might not be detected in general CKD, 2) it may be detected whether hyperglycemia is conferred by T1D or T2D, 3) its effect may be stronger at higher glycemic strata, and 4) interventions that reduce glycemia may attenuate the association signal. *COL4A3* rs55703767 was not associated with eGFR in a general population sample of 110,517 mainly non-diabetic participants of European ancestry[31] (**Table S4**). However, in a smaller cohort of 5,190 participants with T2D and DKD phenotypes in the SUrrogate markers for Micro- and Macro-vascular hard endpoints for Innovative diabetes Tools (SUMMIT) consortium, we detected a directionally consistent suggestive association of *COL4A3* rs55703767 with DN (2-tailed *P*=0.08; **Table S5**).

We further stratified the association analyses by HbA1c in the Finnish Diabetic Nephropathy (FinnDiane) Study, a T1D cohort study with extensive longitudinal phenotypic data[32]. Based on the time-weighted mean of all available HbA1c measurements for each individual, 1,344 individuals had mean HbA1c <7.5% (58 mmol/mol), and 2,977 with mean HbA1c ≥7.5%. *COL4A3*

rs55703767 was nominally significant ($P<0.05$) only in individuals with HbA1c ≥7.5% **(Figure 4, Table S6, Figure S4)**. However, the interaction between HbA1c and *COL4A3* rs55703767 was not significant ($P=0.83$). Upon further examination, the genetic effect was diminished also in the highest HbA1c quartile (>9.3%), as the environmental effect of HbA1c seems to overwhelm any potential genetic effects at *COL4A3* (test of heterogeneity non-significant). In a similar setting of individuals with T2D from the GoDARTS study (N=3226)[13,33], no difference was observed for *COL4A3* rs55703767 between HbA1c strata below or above 7.5% **(Figure S5)**. We performed a similar HbA1c stratified analysis in the Diabetes Control and Complications Trial (DCCT), whose participants, all with T1D, continue to be followed in the Epidemiology of Diabetes Interventions and Complications (EDIC)[34,35]. In DCCT-EDIC the effect of *COL4A3* rs55703767 was stronger among those recruited in the secondary cohort (mild retinopathy and longer diabetes duration at baseline) who were originally randomized to conventional treatment and therefore had higher HbA1c than the intensive treatment group (**Table S7**). Taken together, these independent lines of evidence strongly suggest that the *COL4A3* variant effect on DKD risk is amplified by poor glycemic control.

Other association signals

A comprehensive list of all loci that achieved genome-wide significance from either the minimal or the fully adjusted covariate models is reported in Table 1. The fewer covariates required for the minimal model results in improved statistical power due to fewer subjects with missing data, whereas the fully adjusted model allows the identification of associations potentially mediated by covariates. Comparison of the adjustment models revealed strong consistency between the two models (Figure S6). Table 1 is stratified into two sets of loci: common and/or known kidney biology loci (top half, n=8) and uncommon and no known kidney biology loci (bottom half, n=8). We focus on loci in the top half of the table, common and/or in/near genes with relevant kidney biology.

Two other genome-wide significant signals were near genes encoding proteins related to collagen. Variant rs12615970 (MAF=0.13), located 53 kb downstream of *COLEC11*, was associated with CKD (OR=1.31, P=9.43×10^{-9}), and rs116772905 (MAF=0.011) in exon 14 of *DDR1* was associated with microalbuminuria (OR=3.78, P=4.40×10^{-8}). rs116772905 is in perfect linkage disequilibrium with rs118124843, the lead association with microalbuminuria for this locus under the full adjustment model (taking into account both BMI and HbA1c), located 29 kb downstream of *DDR1* (OR=3.97, P=3.37×10^{-8}). *COLEC11* encodes a collectin protein containing both a collagen-like domain and a carbohydrate recognition domain for binding sugars, and *DDR1* encodes the discoidin domain-containing receptor 1, which binds collagens including type IV collagen.

In addition to *COL4A3* rs55703767, three other low-frequency variants achieved study-wide significance (P<6.76×10^{-9}), each associated with microalbuminuria: rs142823282 (MAF=0.017), 22 kb upstream of *TAMM41* encoding a mitochondrial translocator assembly and maintenance protein[36,37] (OR=6.75, P=1.13×10^{-11}), rs144434404 (MAF=0.011), in intron 1 of *BMP7* encoding the bone morphogenetic protein 7 previously implicated in DKD[38] (OR=6.75, P=2.67×10^{-9}), and rs145681168 (MAF=0.014), in intron 3 of two transcripts of *HAND2* antisense RNA 1 (*HAND2-AS1*; OR=5.53, P=5.40×10^{-9}) and 50 kb upstream of *HAND2*, encoding a heart and neural crest derivatives transcription factor.

Two additional common variants achieved genome-wide significance: rs551191707 (MAF=0.122) in *PRNCR1* associated with ESRD when compared with macroalbuminuria (OR=1.70, P=4.39×10^{-8}) and rs61983410 (MAF=0.213) in an intergenic region on chromosome 14 associated with microalbuminuria (full model OR=0.78, P=3.06×10^{-8}). The remaining seven variants associated with features of DKD had lower allele frequencies (four with 0.01≤MAF≤0.05 and four with MAF<0.01) and did not achieve study-wide significance.

As we had done for *COL4A3* rs55703767, we tested whether the associations of the 15 other variants were amplified by hyperglycemia. None of the 15 variants were significantly associated with eGFR in the general population (**Table S4**). In the smaller SUMMIT T2D cohort[13] we were able to interrogate seven loci with comparable trait definitions. The odds ratios were directionally consistent in six of them (binomial sign test: *P*=0.0625, **Table S5**). In FinnDiane seven of the remaining 15 loci were observed with sufficient frequency (minor allele counts >10) to allow subgroup analysis. Two additional SNPs (rs149641852 in *SNCAIP* and rs12615970 near *COLEC11)* were nominally significant (*P*<0.05) only in individuals with HbA1c ≥7.5%, however the genotype × HbA1c interaction term was non-significant (**Table S6, Figure S4**).

Variants previously associated with DKD

We investigated the effect of variants previously associated at genome-wide significance with renal complications in individuals with diabetes[8-13,39]. Across the ten sub-phenotypes in our meta-analysis, we found evidence of association for seven of nine examined loci (*P*<0.05, **Figure S7**): We replicated two loci that were previously discovered without overlapping individuals with the current study: *SCAF8/CNKSR3* rs12523822, originally associated with DKD (*P*=6.8×10^{-4} for "All vs ctrl") [8]; and *UMOD* rs77924615, originally associated with eGFR in both individuals with and without diabetes (*P*=5.2×10^{-4} for "CKD")[31]. Associations at the *AFF3, RGMA-MCTP2*, and *ERBB4* loci, identified in the GEnetics of Nephropathy—an International Effort (GENIE) consortium[12], comprised of a subset of studies included in this current effort, remained associated with DKD, though the associations were attenuated in this larger dataset (*RGMA-MCTP2* rs12437854 *P*=2.97×10^{-5}; *AFF3* rs7583877 *P*=5.97×10^{-4}; *ERBB4* rs7588550 *P*=3.53×10^{-5}; **Figure S8**). Associations were also observed at the *CDCA7/SP3* (rs4972593, *P*=0.020 for "CKD-DN", originally for ESRD exclusively in women[11]) and *GLRA3* (rs1564939, *P*=0.016 for "CKD extremes", originally for AER[10,39]), but these analyses also include individuals that overlap with the original studies. Apart from the *UMOD* locus, none of the 63 loci associated with eGFR in the

general population[31] were associated with DKD after correction for multiple testing ($P<7.0\times10^{-4}$, **Figure S9**).

Gene and gene set analysis

We conducted gene-level analyses by employing two methods that aggregate SNP summary statistics over a gene region while accounting for linkage disequilibrium, MAGMA and PASCAL[40,41]. MAGMA identified five genes at a Bonferroni-corrected threshold ($P<0.05/18,222$ genes tested = 2.74×10^{-6}): the collagen gene *COL20A1* associated with "CKD extreme" (full model $P=5.77\times10^{-7}$) and "ESRD vs. non-ESRD" (full model P=9.53×10^{-7}), *SLC46A2* associated with "All vs. ctrl" ($P=7.38\times10^{-7}$), *SFXN4* associated with "Macro" (full model $P=1.65\times10^{-7}$), *GLT6D1* associated with "ESRD vs. macro" ($P=1.49\times10^{-6}$), and *SNX30* associated with "All vs ctrl" ($P=2.49\times10^{-6}$) (**Table S8**). Although PASCAL did not identify any significant gene level associations, the five MAGMA-identified genes had $P<5.0\times10^{-4}$ in PASCAL (**Table S9).** Both *SFXN4* and *CBX8* have been reported to be differentially methylated in patients with diabetes with and without nephropathy[42,43].

Additionally, we used MAGMA, PASCAL, DEPICT, and MAGENTA to conduct gene-set analysis in our GWAS dataset. The four methods identified 12 significantly enriched gene sets (**Table S10**). One gene set, "negative regulators of RIG-I MDA5 signaling" was identified in two different pathway analyses (MAGMA and PASCAL) of our fully adjusted GWAS of ESRD *vs.* Macro. Several additional related and overlapping gene sets were identified, including "RIGI MDA5 mediated induction of IFN alpha beta pathways", "TRAF3 dependent IRF activation pathway", and "TRAF6 mediated IRF activation" (PASCAL) and "activated TLR4 signaling" (MAGENTA). RIG-I, MDA5 and the toll-like receptor TLR4 are members of the innate immune response system that respond to both cellular injury and infection[44,45] and transduce highly intertwined signaling cascades. These include the signaling molecules TRAF3 and TRAF6, which induce expression of type I interferons and proinflammatory cytokines implicated in the progression of DKD[46,47].

Specifically, the TLR4 receptor and several of its ligands and downstream cytokines display differential levels of expression in DKD renal tubules *vs.* normal kidneys and *vs.* non-diabetic kidney disease controls[48], and TLR4 knockout mice are protected from DKD and display marked reductions in interstitial collagen deposition in the kidney[49]. Other pathways of interest include "other lipid, fatty acid and steroid metabolism", "nitric oxide signaling in the cardiovascular system", and "Tumor necrosis factor (TNF) family member", with both nitric oxide and TNF-α implicated in DKD[50,51,52].

Expression and epigenetic analyses

We interrogated gene expression datasets in relevant tissues to determine whether our top signals underlie expression quantitative trait loci (eQTL). We first analyzed genotype and RNAseq gene expression data from 96 whole human kidney cortical samples[22] and micro-dissected human kidney samples (121 tubule and 119 glomerular samples)[23] from subjects of European descent without any evidence of renal disease (**Figure S10**). No findings in this data set achieved significance after correction for multiple testing. In the GTEx and eQTLgen datasets, *COL4A3* rs55703767 had a significant eQTL (P=5.63×10^{-38}) with the *MFF* gene in blood, but is most likely due to modest LD with other nearby strong eQTLs in the region. rs118124843 near *DDR1* and *VARS2* had multiple significant eQTLs in blood besides *VARS2* (P=1.71×10^{-5}; **Table S11**). Interestingly, rs142823282 near *TAMM41* was a *cis*-eQTL for *PPARG* (P=4.60×10^{-7})*,* a transcription factor regulating adipocyte development, glucose and lipid metabolism; PPARγ agonists have been suggested to prevent DKD[53].

To ascertain the potential functional role of our top non-coding signals, we mined ChIP-seq data derived from healthy adult human kidney samples[25]. SNP rs142823282 near *TAMM41* was located close to kidney histone marks H3K27ac, H3K9ac, H3K4me1, and H3K4me3, suggesting that this is an active regulator of *TAMM41* or another nearby gene (**Figure S11**). Interestingly, in

recent work we have shown that DNA methylation profiles in participants with T1D with/without kidney disease show the greatest differences in methylation sites near *TAMM41*[54].

To establish whether the expression of our top genes shows enrichment in a specific kidney cell type, we queried an expression dataset of ~50,000 single cells obtained from mouse kidneys[24]. Expression was detected for six genes in the mouse kidney atlas: three (*COL4A3*, *SNCAIP*, and *BMP7*) were almost exclusively expressed in podocytes (**Figure 5**), supporting the significant role for podocytes in DKD.

Gene expression levels in kidneys in cases vs. controls were predicted with TWAS based on the GWAS summary statistics and eQTL data of kidney glomeruli and tubuli.[23] While none of the genes survived correction for multiple testing, analysis suggested 18 genes with differential expression in cases and controls with $P<1\times10^{-4}$, including the *NPNT*, *PRRC2C*, and *VPS33B* genes (Table S12). *NPNT* encodes for nephronectin, an extra-cellular matrix protein on GBM. Knocking out *NPNT* or decreasing *NPNT* expression levels have been shown to induce podocyte injury related to GBM.[55] On the contrary, TWAS predicted higher *NPNT* expression within DN cases vs. normal AER. In line with our TWAS finding, NPNT is significantly upregulated in glomeruli of diabetic nephropathy mouse model vs. non-diabetic mouse ($p=6.4\times10^{-4}$, fold change 1.3, in top 2%, accessed through www.neproseq.org).[56] Furthermore, a variant near *PRRC2C* was recently associated with albuminuria in the UK Biobank,[57] and rare mutations in *VPS33B* gene cause arthrogryposis, renal dysfunction, and cholestasis-1 (ARCS1) syndrome involving proximal–tubular dysfunction and usually death by the age of 1.[58]

DISCUSSION

Our genome-wide analysis of 19,406 participants with T1D identified 16 genome-wide significant loci associated with DKD, four of which remained significant after a conservative correction for multiple testing. Four of the 16 genome-wide significant signals are in or near genes with known

or suggestive biology related to renal function/collagen (*COL4A3*, *BMP7*, *COLEC11*, and *DDR1*), but this is the first time that naturally occurring variation (MAF > 1%) in these loci has been associated with DKD. Our most significant signal was a protective missense variant in *COL4A3*, rs55703767, reaching both genome-wide and study-wide significance with multiple definitions of DKD. Moreover, this variant demonstrated a significant association with GBM width such that protective allele carriers had thinner GBM before any signs of kidney disease, and its effect was dependent on glycemia.

COL4A3, with COL4A4 and COL4A5, make up the so-called "novel chains" of type IV collagen[59], which together play both structural and signaling roles in the GBM. Specifically, COL4A3 is known to bind a number of molecules including integrins, heparin and heparin sulfate proteoglycans, and other components of the GBM such as laminin and nidogen. These interactions mediate the contact between cells and the underlying collagen IV basement membrane, and regulate various processes essential to embryonic development and normal physiology including cell adhesion, proliferation, survival and differentiation. Dysregulation of these interactions has been implicated in several pathological conditions including CKD[60].

Mutations in *COL4A3* are responsible for the autosomal recessive form of Alport syndrome, a progressive inherited nephropathy, as well as benign familial hematuria, characterized by thin (or variable width) GBM, and thought to be a milder form of Alport syndrome[61]. Furthermore, mutations in *COL4A3* have also been identified in patients with focal segmental glomerulosclerosis (FSGS), often leading to proteinuria and renal failure. Some of these patients with FSGS presented with segmental GBM thinning.[62] Of note, the common rs55703767 (*COL4A3* Asp326Tyr) variant, protecting from DKD, was also associated with thinner GBM in individuals with diabetes but without renal complications, a feature that seems to be beneficial in the context of diabetes. The rs55703767 SNP is predicted to alter the third amino acid of the canonical triple-helical domain sequence of Glycine (G)-X-Y (where X and Y are often proline (P)

and hydroxyproline (Y), respectively) from G-E-D (D=Aspartic) to G-E-Y[63], potentially impacting the structure of the collagen complex. In addition, a recent study[64] of candidate genes involved in renal structure reported rs34505188 in *COL4A3* (not in linkage disequilibrium with rs55703767, r^2=0.0006) to be associated with ESRD in African Americans with T2D (MAF=2%, OR=1.55, P=5×10^{-4}). Together with the trend towards association we have seen in SUMMIT and the glycemic interaction we have reported here, these findings suggest variation in *COL4A3* may be associated with DKD in T2D as well.

Given its association as a protective SNP, we can speculate that the rs55703767 variant may confer tensile strength or flexibility to the GBM, which may be of particular relevance in the glomerular hypertension associated with DKD. Alternatively, *COL4A3* may regulate the rates of production and/or turnover of other GBM components, affecting GBM width changes in diabetes. How these effects might confer protection in a manner dependent on ambient glucose concentrations is unknown. Future mechanistic studies will be required to determine the precise role of this variant in DKD; elucidation of its interaction with glycemia in providing protection might be relevant to other molecules implicated in diabetic complications.

In keeping with the collagen pathway, the synonymous exonic variant rs118124843, which reached genome-wide significance for the "Micro" phenotype, is located near *DDR1*, the gene encoding the discoidin domain-containing receptor 1. Based on chromatin conformation interaction data from Capture HiC Plotter (CHiCP),[65] the rs118124843 containing fragment interacts with six gene promoter regions, including *DDR1*, suggesting that the variant regulates *DDR1* expression across multiple tissues (**Table S11**). DDR1 is a collagen receptor[66] shown to bind type IV collagen[67], and is highly expressed in kidneys, particularly upon renal injury[68]. Upon renal injury, *Ddr1*-deficient mice display lower levels of collagen[69], decreased proteinuria, and an increased survival rate compared to wild-type controls[70], with *Ddr1/Col4a3* double knockout mice displaying protection from progressive renal fibrosis and prolonged lifespan compared to *Col4a3*

knockout mice alone[69]. Thus, through its role in collagen binding DDR1 has been suggested as a possible therapeutic target for kidney disease[69].

The association of rs12615970, an intronic variant on chromosome 2 near the *COLEC11* gene, met genome-wide significance for the CKD phenotype, as well as nominal significance for multiple albuminuria-based traits. The rs12615970 containing fragment was found to interact with *COLEC11*, *ALLC*, and *ADI1* transcription start sites in chromatin conformation data on GM12878 cell line (**Table S11**)[65,71]. Collectin-11 is an innate immune factor synthesized by multiple cell types, including renal epithelial cells with a role in pattern recognition and host defense against invasive pathogens, through binding to fructose and mannose sugar moieties[72,73]. Mice with kidney-specific deficiency of *COLEC11* are protected against ischemia-induced tubule injury due to reduced complement deposition[74], and mutations in *COLEC11* have been identified in families with 3MC syndrome, a series of rare autosomal recessive disorders resulting in birth defects and abnormal development, including kidney abnormalities[75].

The intronic variant rs144434404, associated at study-wide significance with the microalbuminuria phenotype, resides within the bone morphogenetic protein 7 (*BMP7*) gene. *BMP7* encodes a secreted ligand of the transforming growth factor-beta superfamily of proteins. Developmental processes are regulated by the BMP family of glycosylated extracellular matrix molecules, via serine/threonine kinase receptors and canonical Smad pathway signaling. Coordinated regulation of both BMP and BMP-antagonist expression is essential for developing tissues, and changes in the levels of either BMP or BMP-antagonists can contribute to disease progression such as fibrosis and cancer[76]. BMP7 is required for renal morphogenesis, and *Bmp7* knockout mice die soon after birth due to reduced ureteric bud branching[77-79]. Maintenance of *Bmp7* expression in glomerular podocytes and proximal tubules of diabetic mice prevents podocyte loss and reduces overall diabetic renal injury[38]. More recently, we have identified a mechanism through which BMP7 orchestrates renal protection through Akt inhibition and

highlights Akt inhibitors as potential anti-fibrotic therapeutics[80]. It is also noteworthy that the BMP7 antagonist grem-1 is implicated in DKD[81-83] and gremlin has been implicated as a biomarker of kidney disease[84].

Strengths of this analysis include the large sample size, triple that of the previous largest GWAS; the uniform genotyping and quality control procedures; standardized imputation for all studies (1,000 Genomes reference panel); the inclusion of exome array content; the exploration of multiple standardized phenotype definitions of DKD; and supportive data from various sources of human kidney samples. Several of the loci identified have known correlations with kidney biology, suggesting that these are likely true associations with DKD. However, we acknowledge a number of limitations. First, nine variants have low MAF and were driven by only two cohorts, indicating that further validation will be required to increase confidence in these associations. Second, seven variants were significantly associated with microalbuminuria only, a trait shown to be less heritable in previous studies. We included these loci to maximize comprehensiveness in reporting novel DKD associations. Replication in independent samples and functional confirmation is required to validate all of these loci. Even though the gene-level, gene set and pathway analyses had limited power, these analyses identified several additional potential DKD loci and pathways, some with relevance to kidney biology, that require further follow-up. Finally, while we included only controls with a minimum diabetes duration of 15 years, we cannot fully rule out that some of the controls would progress to DKD in the future, as the improvements in diabetes treatment in the last decades have postponed the onset of complications. We also excluded cases with short diabetes duration to avoid renal complications that might be due to other causes. These phenotypic definitions were meant to overcome the limitation that in clinical practice kidney biopsies are rarely taken from individuals with diabetes to verify the diagnosis. As for any late onset disease, these challenges in phenotypic definition may have reduced our power to detect additional associations. We note, however, that this relatively small degree of contamination would lead to loss of power

and increased type II error rather than false positive findings; therefore, it does not undermine the robustness of the associations reported here.

Diabetic complications are unquestionably driven by hyperglycemia and partially prevented by improved glycemic control in both T1D and T2D, but there has been doubt over what contribution, if any, inherited factors contribute to disease risk. In line with previous genetic studies, this study with a markedly expanded sample size identified several loci strongly associated with DKD risk. These findings suggest that larger studies, aided by novel analyses and including T2D, will continue to enhance our understanding of the complex pathogenesis of DKD, paving the way for molecularly targeted preventive or therapeutic interventions.

AUTHOR CONTRIBUTIONS

DATA AND SOFTWARE AVAILABILITY

GWAS summary statistics for all ten DKD phenotypes and two adjustment models are available for download at the AMP-Type 2 Diabetes Knowledge Portal (http://www.type2diabetesgenetics.org/informational/data), under "JDRF Diabetic Nephropathy Collaborative Research Initiative GWAS" datasets. Individual level genotype data cannot be shared for all cohorts due to restrictions set by study consents and by EU and national regulations, individual genotype data.

Table of Contents (Supplemental Results)

REFERENCES

1.   Centers for Disease Control and Prevention. National Diabetes Statistics Report, 2017 Estimates of Diabetes and Its Burden in the United States Background. in *Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services; 2017.*
2.   Tuttle, K.R. *et al.* Diabetic kidney disease: a report from an ADA Consensus Conference. *Am J Kidney Dis* **64**, 510-33 (2014).
3.   Krolewski, M., Eggers, P.W. & Warram, J.H. Magnitude of end-stage renal disease in IDDM: a 35 year follow-up study. *Kidney Int* **50**, 2041-2046 (1996).
4.   Harjutsalo, V., Katoh, S., Sarti, C., Tajima, N. & Tuomilehto, J. Population-based assessment of familial clustering of diabetic nephropathy in type 1 diabetes. *Diabetes* **53**, 2449-54 (2004).
5.   Quinn, M., Angelico, M.C., Warram, J.H. & Krolewski, A.S. Familial factors determine the development of diabetic nephropathy in patients with IDDM. *Diabetologia* **39**, 940-5 (1996).
6.   Seaquist, E., Goetz, F., Rich, S. & Barbosa, J. Familial clustering of diabetic kidney disease. Evidence for genetic susceptibility to diabetic nephropathy. *N Engl J Med* **320**, 1161-1165 (1989).
7.   Sandholm, N. *et al.* The Genetic Landscape of Renal Complications in Type 1 Diabetes. *J Am Soc Nephrol* **28**, 557-574 (2017).
8.   Iyengar, S.K. *et al.* Genome-Wide Association and Trans-ethnic Meta-Analysis for Advanced Diabetic Kidney Disease: Family Investigation of Nephropathy and Diabetes (FIND). *PLoS Genet* **11**, e1005352 (2015).
9.   Pattaro, C. *et al.* Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat Commun* **7**, 10023 (2016).
10.   Sandholm, N. *et al.* Genome-wide association study of urinary albumin excretion rate in patients with type 1 diabetes. *Diabetologia* **57**, 1143-53 (2014).
11.   Sandholm, N. *et al.* Chromosome 2q31.1 Associates with ESRD in Women with Type 1 Diabetes. *Journal of the American Society of Nephrology* **24**, 1537-1543 (2013).
12.   Sandholm, N. *et al.* New susceptibility loci associated with kidney disease in type 1 diabetes. *PLoS Genet* **8**, e1002921 (2012).
13.   van Zuydam, N.R. *et al.* A Genome-Wide Association Study of Diabetic Kidney Disease in Subjects With Type 2 Diabetes. *Diabetes* **67**, 1414-1427 (2018).
14.   Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955-9 (2012).
15.   Fuchsberger, C., Abecasis, G.R. & Hinds, D.A. minimac2: faster genotype imputation. *Bioinformatics* **31**, 782-4 (2015).
16.   Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)* **95**, 221-7 (2005).
17.   McKelvey, R.D. & Zavoina, W. A statistical model for the analysis of ordinal level dependent variables. *The Journal of Mathematical Sociology* **4**, 103-120 (1975).
18.   Mauer, M. *et al.* Renal and retinal effects of enalapril and losartan in type 1 diabetes. *N Engl J Med* **361**, 40-51 (2009).
19.   Cohen, C.D., Frach, K., Schlondorff, D. & Kretzler, M. Quantitative gene expression analysis in renal biopsies: a novel protocol for a high-throughput multicenter application. *Kidney Int* **61**, 133-40 (2002).
20.   Berthier, C.C. *et al.* Enhanced expression of Janus kinase-signal transducer and activator of transcription pathway members in human diabetic nephropathy. *Diabetes* **58**, 469-77 (2009).

21. Schmid, H. *et al.* Modular activation of nuclear factor-kappaB transcriptional programs in human diabetic nephropathy. *Diabetes* **55**, 2993-3003 (2006).
22. Ko, Y.A. *et al.* Genetic-Variation-Driven Gene-Expression Changes Highlight Genes with Important Functions for Kidney Disease. *Am J Hum Genet* **100**, 940-953 (2017).
23. Qiu, C. *et al.* Renal compartment-specific genetic variation analyses identify new pathways in chronic kidney disease. *Nat Med* **24**, 1721-1731 (2018).
24. Park, J. *et al.* Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758-763 (2018).
25. Bernstein, B.E. *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* **28**, 1045-8 (2010).
26. Barbeira, A.N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun* **9**, 1825 (2018).
27. Yagame, M. *et al.* Differential distribution of type IV collagen chains in patients with diabetic nephropathy in non-insulin-dependent diabetes mellitus. *Nephron* **70**, 42-8 (1995).
28. Caramori, M.L., Parks, A. & Mauer, M. Renal lesions predict progression of diabetic nephropathy in type 1 diabetes. *J Am Soc Nephrol* **24**, 1175-81 (2013).
29. Fufaa, G.D. *et al.* Structural Predictors of Loss of Renal Function in American Indians with Type 2 Diabetes. *Clin J Am Soc Nephrol* **11**, 254-61 (2016).
30. Mauer, M., Caramori, M.L., Fioretto, P. & Najafian, B. Glomerular structural-functional relationship models of diabetic nephropathy are robust in type 1 diabetic patients. *Nephrol Dial Transplant* **30**, 918-23 (2015).
31. Gorski, M. *et al.* 1000 Genomes-based meta-analysis identifies 10 novel loci for kidney function. *Sci Rep* **7**, 45040 (2017).
32. Thorn, L.M. *et al.* Metabolic syndrome in type 1 diabetes: Association with diabetic nephropathy and glycemic control (the FinnDiane study). *Diabetes Care* **28**, 2019-2024 (2005).
33. Morris, A.D. *et al.* The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. *BMJ* **315**, 524-8 (1997).
34. Nathan, D.M. & Group, D.E.R. The diabetes control and complications trial/epidemiology of diabetes interventions and complications study at 30 years: overview. *Diabetes Care* **37**, 9-16 (2014).
35. The Diabetes Control and Complications Trial Research Group. Implementation of treatment protocols in the Diabetes Control and Complications Trial. *Diabetes Care* **18**, 361-76 (1995).
36. Blunsom, N.J., Gomez-Espinosa, E., Ashlin, T.G. & Cockcroft, S. Mitochondrial CDP-diacylglycerol synthase activity is due to the peripheral protein, TAMM41 and not due to the integral membrane protein, CDP-diacylglycerol synthase 1. *Biochim Biophys Acta Mol Cell Biol Lipids* **1863**, 284-298 (2018).
37. Tamura, Y. *et al.* Tam41 is a CDP-diacylglycerol synthase required for cardiolipin biosynthesis in mitochondria. *Cell Metab* **17**, 709-18 (2013).
38. Wang, S. *et al.* Renal bone morphogenetic protein-7 protects against diabetic nephropathy. *J Am Soc Nephrol* **17**, 2504-12 (2006).
39. Sandholm, N. *et al.* Confirmation of GLRA3 as a susceptibility locus for albuminuria in Finnish patients with type 1 diabetes. *Sci Rep* **8**, 12408 (2018).
40. Lamparter, D., Marbach, D., Rueedi, R., Kutalik, Z. & Bergmann, S. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. *PLoS Comput Biol* **12**, e1004714 (2016).

41. de Leeuw, C.A., Mooij, J.M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).
42. Bell, C.G. *et al.* Genome-wide DNA methylation analysis for diabetic nephropathy in type 1 diabetes mellitus. *BMC Medical Genomics* **3**(2010).
43. Sapienza, C. *et al.* DNA methylation profiling identifies epigenetic differences between diabetes patients with ESRD and diabetes patients without nephropathy. *Epigenetics* **6**, 20-8 (2011).
44. Loo, Y.M. & Gale, M., Jr. Immune signaling by RIG-I-like receptors. *Immunity* **34**, 680-92 (2011).
45. Mollen, K.P. *et al.* Emerging paradigm: toll-like receptor 4-sentinel for the detection of tissue damage. *Shock* **26**, 430-7 (2006).
46. Mora, C. & Navarro, J.F. Inflammation and diabetic nephropathy. *Curr Diab Rep* **6**, 463-8 (2006).
47. Wada, J. & Makino, H. Inflammation and the pathogenesis of diabetic nephropathy. *Clin Sci (Lond)* **124**, 139-52 (2013).
48. Lin, M. *et al.* Toll-like receptor 4 promotes tubular inflammation in diabetic nephropathy. *J Am Soc Nephrol* **23**, 86-102 (2012).
49. Ma, J. *et al.* TLR4 activation promotes podocyte injury and interstitial fibrosis in diabetic nephropathy. *PLoS One* **9**, e97985 (2014).
50. Colhoun, H.M. & Marcovecchio, M.L. Biomarkers of diabetic kidney disease. *Diabetologia* **61**, 996-1011 (2018).
51. McKnight, A.J. *et al.* Genetic polymorphisms in nitric oxide synthase 3 gene and implications for kidney disease: a meta-analysis. *Am J Nephrol* **32**, 476-81 (2010).
52. Prabhakar, S.S. Role of nitric oxide in diabetic nephropathy. *Semin Nephrol* **24**, 333-44 (2004).
53. Chasman, D.I. *et al.* Integration of genome-wide association studies with biological knowledge identifies six novel genes related to kidney function. *Hum Mol Genet* **21**, 5329-43 (2012).
54. Swan, E.J., Maxwell, A.P. & McKnight, A.J. Distinct methylation patterns in genes that affect mitochondrial function are associated with kidney disease in blood-derived DNA from individuals with Type 1 diabetes. *Diabet Med* **32**, 1110-5 (2015).
55. Muller-Deile, J. *et al.* Podocytes regulate the glomerular basement membrane protein nephronectin by means of miR-378a-3p in glomerular diseases. *Kidney Int* **92**, 836-849 (2017).
56. Hodgin, J.B. *et al.* Identification of cross-species shared transcriptional networks of diabetic nephropathy in human and mouse glomeruli. *Diabetes* **62**, 299-308 (2013).
57. Zanetti, D. *et al.* Identification of 22 novel loci associated with urinary biomarkers of albumin, sodium, and potassium excretion. *Kidney Int* **95**, 1197-1208 (2019).
58. Holme, A. *et al.* Glomerular involvement in the arthrogryposis, renal dysfunction and cholestasis syndrome. *Clin Kidney J* **6**, 183-8 (2013).
59. Kleppel, M.M., Fan, W., Cheong, H.I. & Michael, A.F. Evidence for separate networks of classical and novel basement membrane collagen. Characterization of alpha 3(IV)-alport antigen heterodimer. *J Biol Chem* **267**, 4137-42 (1992).
60. Khoshnoodi, J., Pedchenko, V. & Hudson, B.G. Mammalian collagen IV. *Microsc Res Tech* **71**, 357-70 (2008).
61. Kashtan, C.E. *et al.* Alport syndrome: a unified classification of genetic disorders of collagen IV alpha345: a position paper of the Alport Syndrome Classification Working Group. *Kidney Int* **93**, 1045-1051 (2018).
62. Xie, J. *et al.* COL4A3 mutations cause focal segmental glomerulosclerosis. *J Mol Cell Biol* **6**, 498-505 (2014).

63. Parkin, J.D. *et al.* Mapping structural landmarks, ligand binding sites, and missense mutations to the collagen IV heterotrimers predicts major functional domains, novel interactions, and variation in phenotypes in inherited diseases affecting basement membranes. *Hum Mutat* **32**, 127-43 (2011).

64. Guan, M. *et al.* Association of kidney structure-related gene variants with type 2 diabetes-attributed end-stage kidney disease in African Americans. *Hum Genet* **135**, 1251-1262 (2016).

65. Schofield, E.C. *et al.* CHiCP: a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets. *Bioinformatics* **32**, 2511-3 (2016).

66. Alves, F. *et al.* Distinct structural characteristics of discoidin I subfamily receptor tyrosine kinases and complementary expression in human cancer. *Oncogene* **10**, 609-18 (1995).

67. Vogel, W., Gish, G.D., Alves, F. & Pawson, T. The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol Cell* **1**, 13-23 (1997).

68. Dorison, A. & Chantziantoniou, C. DDR1: A major player in renal diseases. *Cell Adh Migr* **12**, 299-304 (2018).

69. Gross, O. *et al.* Loss of collagen-receptor DDR1 delays renal fibrosis in hereditary type IV collagen disease. *Matrix Biol* **29**, 346-56 (2010).

70. Kerroch, M. *et al.* Genetic inhibition of discoidin domain receptor 1 protects mice against crescentic glomerulonephritis. *FASEB J* **26**, 4079-91 (2012).

71. Mifsud, B. *et al.* Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet* **47**, 598-606 (2015).

72. Selman, L. & Hansen, S. Structure and function of collectin liver 1 (CL-L1) and collectin 11 (CL-11, CL-K1). *Immunobiology* **217**, 851-63 (2012).

73. Hansen, S. *et al.* Collectin 11 (CL-11, CL-K1) is a MASP-1/3-associated plasma collectin with microbial-binding activity. *J Immunol* **185**, 6096-104 (2010).

74. Farrar, C.A. *et al.* Collectin-11 detects stress-induced L-fucose pattern to trigger renal epithelial injury. *J Clin Invest* **126**, 1911-25 (2016).

75. Rooryck, C. *et al.* Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome. *Nat Genet* **43**, 197-203 (2011).

76. Walsh, D.W., Godson, C., Brazil, D.P. & Martin, F. Extracellular BMP-antagonist regulation in development and disease: tied up in knots. *Trends Cell Biol* **20**, 244-56 (2010).

77. Zeisberg, M., Shah, A.A. & Kalluri, R. Bone morphogenic protein-7 induces mesenchymal to epithelial transition in adult renal fibroblasts and facilitates regeneration of injured kidney. *J Biol Chem* **280**, 8094-100 (2005).

78. Vukicevic, S., Kopp, J.B., Luyten, F.P. & Sampath, T.K. Induction of nephrogenic mesenchyme by osteogenic protein 1 (bone morphogenetic protein 7). *Proc Natl Acad Sci U S A* **93**, 9021-6 (1996).

79. Luo, G. *et al.* BMP-7 is an inducer of nephrogenesis, and is also required for eye development and skeletal patterning. *Genes Dev* **9**, 2808-20 (1995).

80. Higgins, D.F. *et al.* BMP7-induced-Pten inhibits Akt and prevents renal fibrosis. *Biochim Biophys Acta Mol Basis Dis* **1863**, 3095-3104 (2017).

81. Roxburgh, S.A. *et al.* Allelic depletion of grem1 attenuates diabetic kidney disease. *Diabetes* **58**, 1641-50 (2009).

82. Dolan, V. *et al.* Expression of gremlin, a bone morphogenetic protein antagonist, in human diabetic nephropathy. *Am J Kidney Dis* **45**, 1034-9 (2005).

83. McMahon, R. *et al.* IHG-2, a mesangial cell gene induced by high glucose, is human gremlin. Regulation by extracellular glucose concentration, cyclic mechanical strain, and transforming growth factor-beta1. *J Biol Chem* **275**, 9901-4 (2000).

84. Afkarian, M. *et al.* Urinary excretion of RAS, BMP, and WNT pathway components in diabetic kidney disease. *Physiol Rep* **2**, e12010 (2014).

TABLES

Table 1. Loci associated with DKD at study-wide ($P<6.76\times10^{-9}$, bold) and genome-wide ($P<5\times10^{-8}$) significance.
Common variants and/or genes with relevant kidney biology are reported in the top half of the table. Uncommon variants (MAF < 2%) with no known relevant kidney biology are reported in the bottom half of the table. Genes are annotated as follows: missense variant in the indicated gene (M); intronic, synonymous, or noncoding variant in the indicated gene (G); gene nearest to lead variant (N); gene has relevant kidney (B). Chr, chromosome; pos, position; EAF, effect allele frequency; OR, odds ratio; min, minimally adjusted covariate model; full, fully adjusted covariate model.

SNP	Chr:pos	Effect allele	Other allele	EAF	Notable gene(s)	Phenotype	OR_{min}	$P\text{-value}_{min}$	OR_{full}	$P\text{-value}_{full}$
rs55703767	2:228121101	T	G	0.206	*COL4A3* (M, B, N)	DN	0.79	**5.34×10^{-12}**	0.78	**8.19×10^{-11}**
						All vs. ctrl	0.83	**3.88×10^{-10}**	0.84	9.68×10^{-9}
						CKD+DN	0.77	**5.30×10^{-9}**	0.76	3.77×10^{-8}
						Macro	0.78	9.28×10^{-9}	0.77	9.38×10^{-9}
rs12615970	2:3745215	G	A	0.133	*COLEC11 (B) ALLC (N, G)*	CKD	0.76	9.43×10^{-9}	0.77	1.60×10^{-7}
rs142823282	3:11910635	G	A	0.011	*TAMM41* (N, B)	Micro	6.73	**8.32×10^{-10}**	9.18	**1.13×10^{-11}**
rs145681168	4:174500806	G	A	0.014	*HAND2-AS1* (N, G, B)	Micro	5.53	2.06×10^{-7}	7.47	**5.40×10^{-9}**
rs118124843	6:30887465	T	C	0.011	*DDR1 (B) VARS2 (G)*	Micro	3.79	4.42×10^{-8}	3.99	3.37×10^{-8}
rs77273076	7:99728546	T	C	0.008	*MBLAC1* (N, B)	Micro	9.16	1.04×10^{-8}	7.10	2.28×10^{-7}
rs551191707	8:128100029	CA	C	0.122	*PRNCR1* (N)	ESRD vs. macro	1.70	4.39×10^{-8}	1.71	3.15×10^{-6}
rs61983410	14:26004712	T	C	0.213	*STXBP6* (N)	Micro	0.79	9.84×10^{-8}	0.78	3.06×10^{-8}
rs144434404	20:55837263	T	C	0.011	*BMP7* (N, G, B)	Micro	6.78	**2.67×10^{-9}**	6.66	**4.65×10^{-9}**

rs115061173	3:926345	A	T	0.014	*LINC01266* (N)	ESRD vs. ctrl	9.40	$4.07×10^{-8}$	8.34	$4.08×10^{-5}$
rs116216059	3:36566312	A	C	0.016	*STAC* (N, G)	ESRD vs. non-ESRD	8.73	$1.37×10^{-8}$	11.78	$1.41×10^{-4}$
rs191449639	4:71358776	A	T	0.005	*MUC7* (N)	DN	32.42	$1.32×10^{-8}$	32.47	$2.09×10^{-8}$
rs149641852	5:121774582	T	G	0.012	*SNCAIP* (N, G)	CKD extreme	9.01	$1.37×10^{-8}$	---	---
rs183937294	11:16937846	G	T	0.007	*PLEKHA7* (N, G)	Micro	17.22	$1.65×10^{-8}$	23.62	$2.10×10^{-6}$
rs113554206	14:73740250	A	G	0.012	*PAPLN* (N, G)	Macro	4.60	$5.39×10^{-7}$	10.42	$8.46×10^{-9}$
rs185299109	18:1811108	T	C	0.007	*intergenic*	CKD	20.75	$1.28×10^{-8}$	44.75	$4.99×10^{-7}$

Figure 1
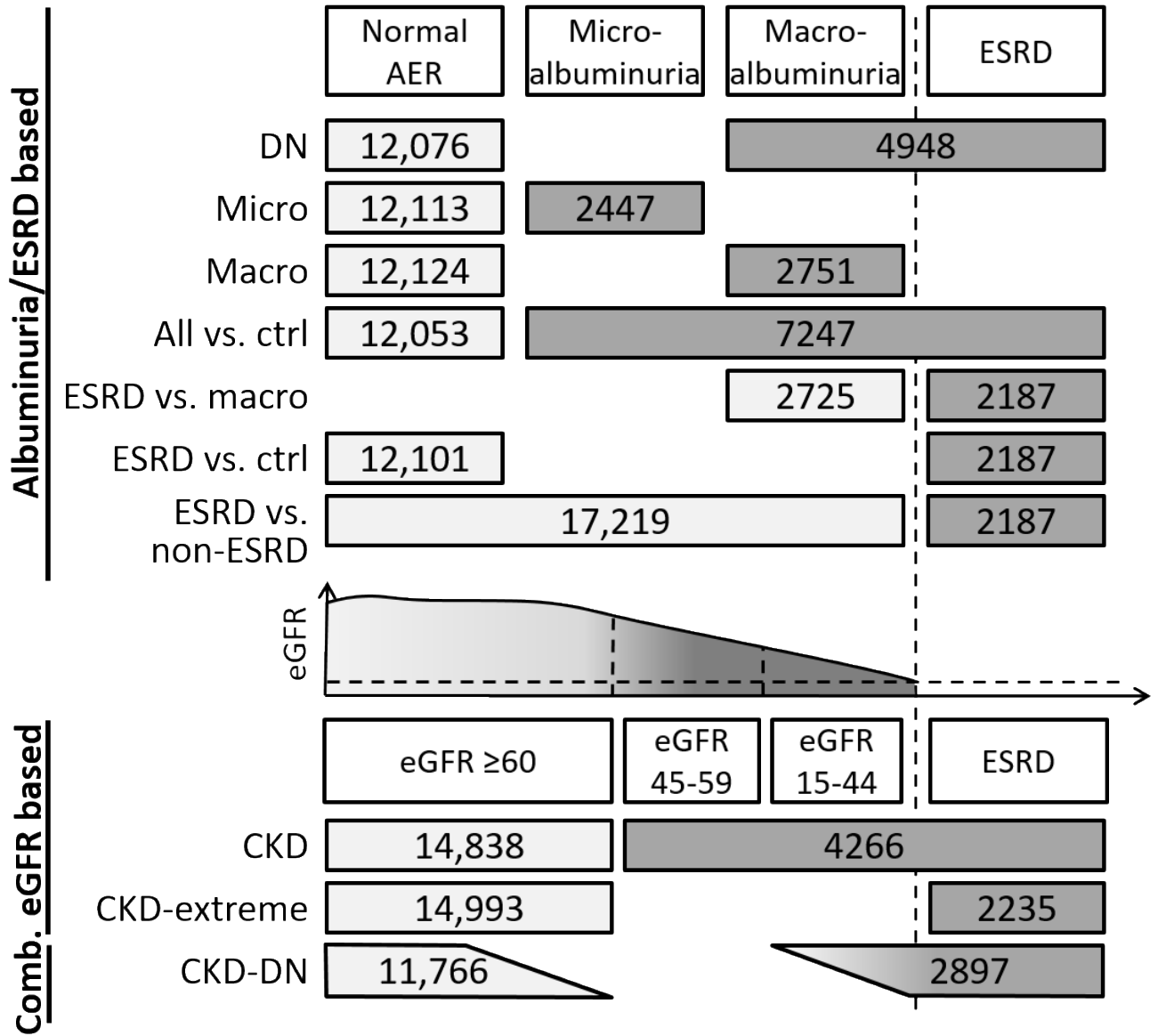
Figure 1. Phenotypic analysis of DKD. Schematic diagram of outcomes analyzed in this study. Numbers indicate the total number of cases (darker gray) and controls (lighter gray) included in the meta-analyses for each phenotype. Microalb.: microalbuminuria; macroalb.: macroalbuminuria; eGFR: estimated glomerular filtration rate; ESRD: End-stage renal disease, defined as eGFR <15 mL/min/1.73m^2 or undergoing dialysis or having renal transplant; CKD: chronic kidney disease.

Figure 2
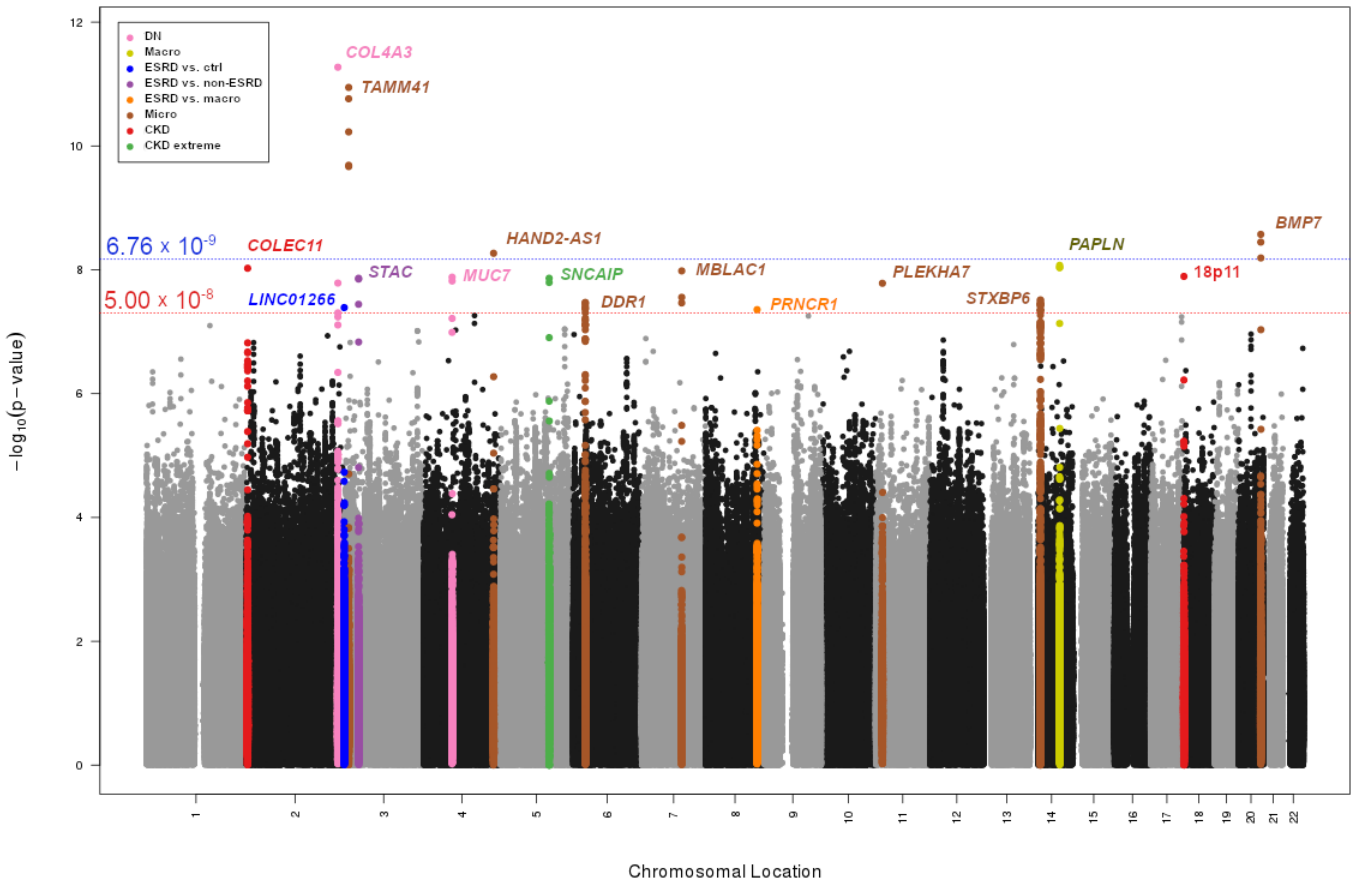
Figure 2 Genome-wide association testing of all 10 phenotypic comparisons. Multiphenotype Manhattan plot shows lowest *P*-value at each marker for each of the 10 phenotypic comparisons, under the standard and fully-adjusted model. Significance of SNPs (-log$_{10}$[*P*-value], y axis) is plotted against genomic location (x axis). Loci surpassing genome-wide significance (red line) and/or study-wide significance (blue line) are colored by phenotype.

Figure 3
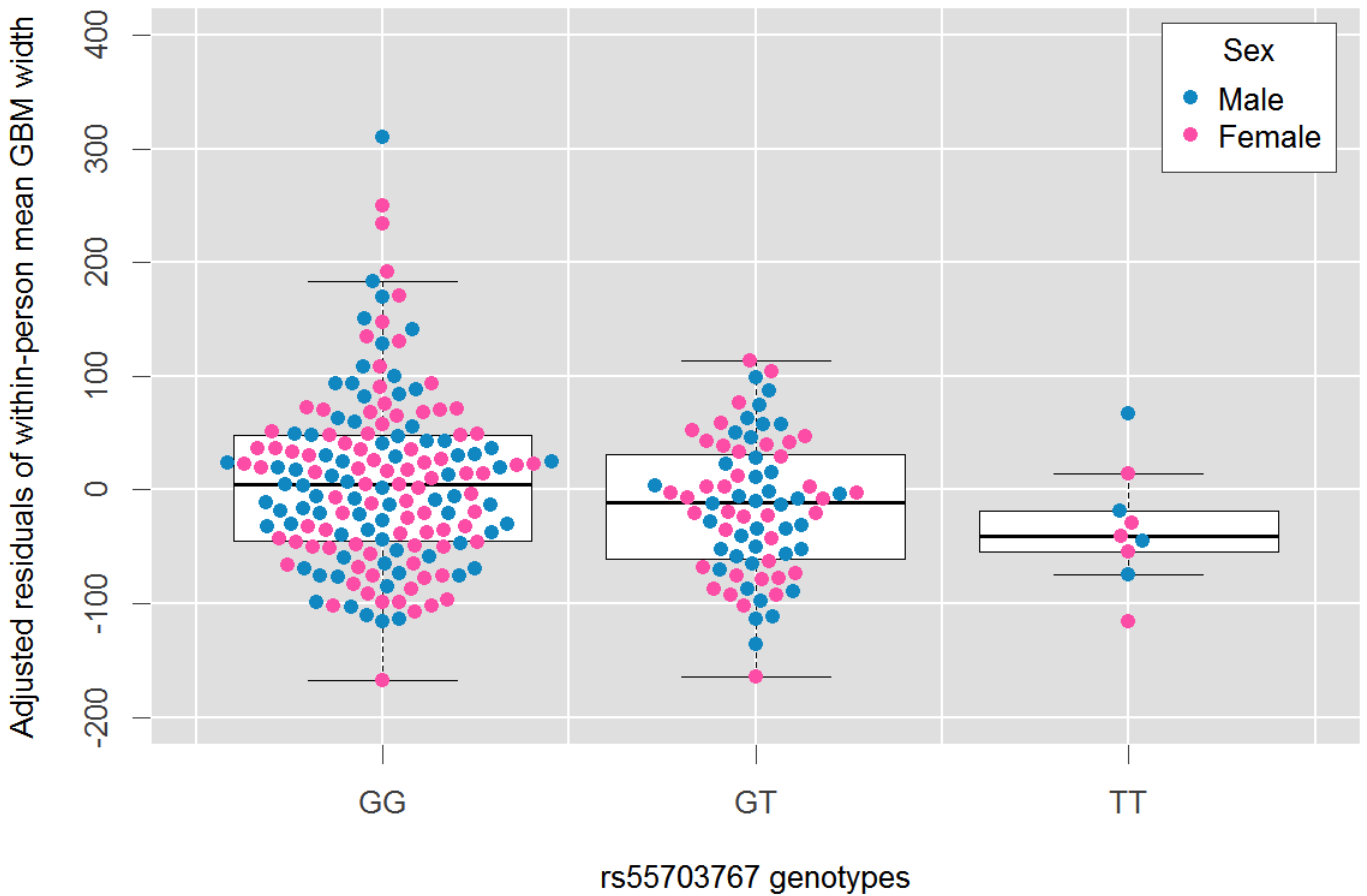
Figure 3. Adjusted residuals of GBM width by rs55703767 genotype and sex. Box and whisker plot of residuals of mean GBM width after adjusting for age, sex, and diabetes duration, stratified by GG, GT, or TT genotype at rs55703767, with overlay of individual data points for both females (pink) and males (blue).

Figure 4

Pheno	P		P HbA1c<7.5	P HbA1>7.5
All vs. ctrl	0.007		0.87	0.0018
DN	0.0022		0.659	0.001
Macro	0.06		0.663	0.0066
CKD-DN	0.0117		0.973	0.0079
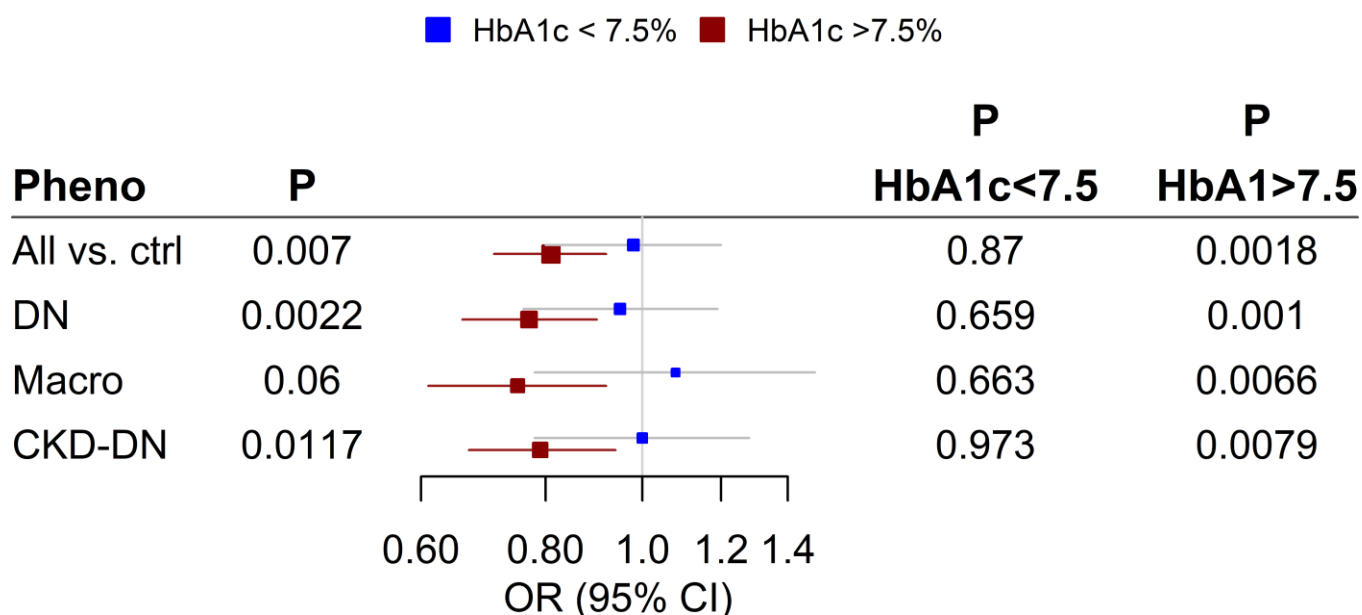
Figure 4. Association at rs55703767 (*COL4A3*) stratified by HbA1c below or above 7.5%, for the phenotypes reaching genome-wide significance in the combined meta-analysis. Analysis included 1344 individuals with time-weighted mean HbA1c <7.5% (58 mmol/mol), and 2977 with mean HbA1c ≥7.5% from the FinnDiane study; the individuals had median 19 HbA1c measurements (range 1 – 129).

Figure 5
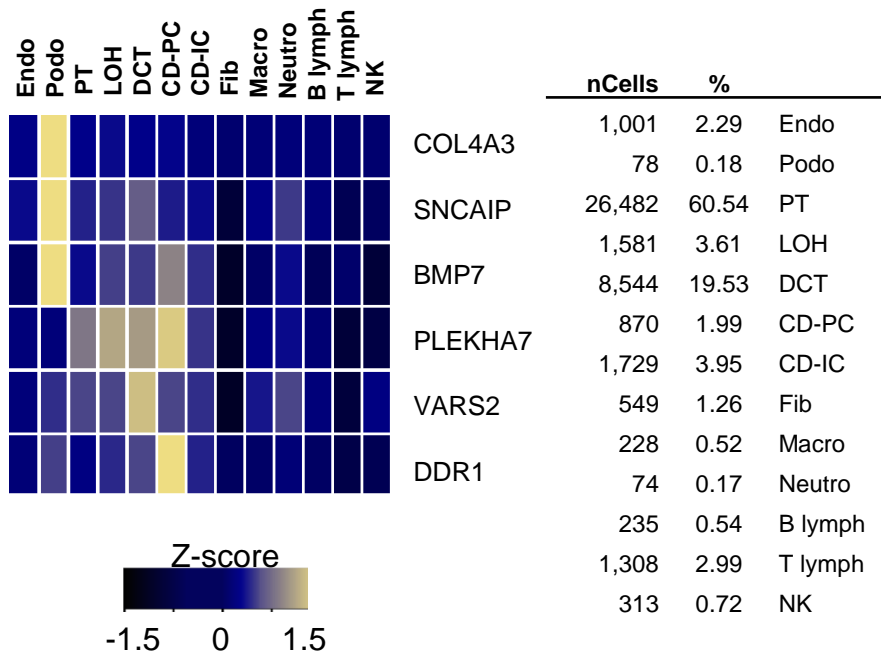
Figure 5. Single cell RNA-sequencing in mouse kidney shows *COL4A3*, *SNCAIP*, and *BMP7* are specifically expressed in podocytes. Mean expression values of the genes were calculated in each cluster. The color scheme is based on z-score distribution; the map shows genes with z-score>2. In the heatmap, each row represents one gene and each column is single cell type. Percentages of assigned cell types are summarized in the right panel. Endo, containing endothelial, vascular, and descending loop of Henle; Podo, podocyte; PT, proximal tubule; LOH, ascending loop of Henle; DCT, distal convoluted tubule; CD-PC, collecting duct principal cell; CD-IC, collecting duct intercalated cell; CD-Trans, collecting duct transitional cell; Fib, fibroblast; Macro, macrophage; Neutro, neutrophil; lymph, lymphocyte; NK, natural killer cell.

SUPPLEMENTAL INFORMATION

Contents

SUPPLEMENTAL METHODS.

Cohorts in GWAS. The GWAS meta-analysis included up to 19,406 patients with type 1 diabetes and of European origin from 17 cohorts: The Austrian Diabetic Nephropathy Study (AusDiane); The Coronary Artery Calcification in Type 1 Diabetes (CACTI)[1]; the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC)[2, 3]; Pittsburgh Epidemiology of Diabetes Complications Study (EDC)[4]; The Finnish Diabetic Nephropathy (FinnDiane) Study[5, 6]; French and Belgian subjects from the Genetics of Diabetic Nephropathy (GENEDIAB)[7] and Genesis[8] studies; Genetics of Kidneys in Diabetes US Study (GoKinD) from George Washington University (GWU-GoKinD)[9]; patients from the Joslin Kidney Study[9, 10]; individuals with T1D from Italy[5]; The Latvian Diabetic Nephropathy Study (LatDiane)[11]; The Lithuanian Diabetic Nephropathy Study (LitDiane) [Reference pending, submitted]; The Romanian Diabetic Nephropathy Study (RomDiane)[12]; The Scottish Diabetes Research Network Type 1 Bioresource (SDRNT1BIO)[13, 14]; individuals with T1D from Steno Diabetes Center[15]; individuals with T1D from Uppsala, Sweden[16, 17]; UK GoKinD, Warren 3 and All Ireland (UK-ROI) study[18]; and The Wisconsin Epidemiologic Study of Diabetic Retinopathy (WESDR)[19]. All participants gave informed consent and all studies were approved by ethics committees from all participating institutions.

GWAS Genotyping. Samples were genotyped on the HumanCore BeadChip (Illumina, San Diego, CA, USA), which contains 250,000 genome-wide tag SNPs (and other variants) and over 200,000 exome-focused variants. All samples were passed through a stringent quality control protocol. Following initial genotype calling with Illumina software, all samples were re-called with zCall, a calling algorithm specifically designed for rare SNPs from arrays. Once calling was completed for all cohorts, our pipeline updated variant orientation and position aligned to hg19

(Genome Reference Consortium Human Build 37, GRCh37). Variant names were updated using 1000 Genomes as a reference. The data were then filtered for low quality variants (e.g. call rates <95% or excessive deviation from Hardy-Weinberg equilibrium) or samples (e.g. call rates <98%, gender mismatch, extreme heterozygosity). Principal Component Analysis (PCA) was performed separately for each cohort in order to empirically detect and exclude outliers with evidence of non-European ancestry. Genotypes were expanded to a total of approximately 49 million by imputation, using 1,000 Genomes Project (phase 3 version 5) as a reference.

GWAS Phenotype definitions. Participant renal status was evaluated on the basis of both albuminuria and eGFR. We defined a total of 10 different case-control outcomes to cover the different aspects of renal complications (**Figure 1**). Five comparisons ("All vs. ctrl", "Micro", "DN", "Macro", and "ESRD vs. macro") were based on albuminuria, measured by albumin excretion rate (AER) from overnight or 24-h urine collection, or by albumin creatinine ratio (ACR). Two out of three consecutive collections were required (when available) to classify the renal status of subjects as either normoalbuminuria, microalbuminuria, macroalbuminuria, or ESRD; for detailed thresholds, see **Table S9**. Controls with normal AER were required to have a minimum diabetes duration of 15 years; subjects with microalbuminuria/ macroalbuminuria/ ESRD were required to have minimum diabetes duration of 5/ 10/ 10 years, respectively, in order to exclude renal complications of non-diabetic origins. Two comparisons ("ESRD vs. ctrl" and "ESRD vs. non-ESRD") were based on presence of end-stage renal disease as defined by eGFR< 15 mL/min or dialysis or renal transplant. Two phenotypes ("CKD" and "CKD extreme") were defined based on estimated glomerular filtration rate (eGFR; evaluated with the CKD-EPI formula): Controls had eGFR ≥ 60ml/min/1.73m^2 for both phenotypes, and minimum of 15 years of diabetes duration; cases had eGFR <60ml/min/1.73m^2 for the "CKD" phenotype, and eGFR <15 ml/min/1.73m^2 or dialysis or renal transplant for the "CKD extreme" phenotype, and

minimum of 10 years of diabetes duration. For the "CKD-DN" phenotype that combined both albuminuria and eGFR data, controls were required to have both eGFR ≥60ml/min/1.73m^2 and normoalbuminuria; cases had both eGFR <45ml/min/1.73m^2 and micro- or macroalbuminuria, or ESRD.

GWAS Statistical Analysis. A genome-wide association analysis of each of the case-control definitions was performed using logistic regression under an additive genetic model, adjusting for age, sex, diabetes duration, study site (where applicable) and principal components. As disease onset and progression is also closely related to BMI and HbA1c levels,[20] we conducted a second set of analyses adjusting for BMI and HbA1c which we refer to as our fully adjusted covariate model. Allele dosages were used to account for imputation uncertainty. Inverse-variance fixed effects meta-analysis was performed using METAL and the following filters: INFO score >0.3, minor allele count >10, and presence of variant in at least two cohorts. The X chromosome was similarly analyzed for males and females both separately and in a combined analysis, with the exception of using hard call genotypes in place of allele dosages. The study-wide significance threshold (P<6.76×10^{-9}) was calculated by applying a Bonferroni correction to the traditional GWAS threshold (P<5.00×10^{-8}), based on the number of effectively independent tests, using methods previously described on the eigenvalues of the GWAS summary statistics correlation matrix[21].

Glomerular basement membrane measurement in Renin-Angiotensin System Study (RASS). RASS was a double-blind placebo-controlled randomized trial of the angiotensin converting enzyme inhibitor (ACEi) enalapril and the angiotensin II receptor blocker (ARB) losartan on renal pathology among 285 normoalbuminuric, normotensive subjects with T1D and had normal or increased measured glomerular filtration rate (>90 ml/min/1.73m^2)[22]. Beginning in

2005, participants were recruited from three centers: University of Minnesota (Minneapolis, Minnesota), McGill University (Montreal, Canada) and University of Toronto (Toronto, Canada) and included those with 2 to 20 years of diabetes and excluded those on any antihypertensive medications. Written informed consent was obtained from each participant and the study was approved by the relevant institutional review boards. RASS study participants were followed for 5 years with percutaneous kidney biopsy completed prior to randomization and at 5 years. Structural parameters measured by electron microscopy on biopsy included GBM width, measured by the electron microscopic orthogonal intercept method[22].

RASS study participants were followed for 5 years with percutaneous kidney biopsy completed prior to randomization and at 5 years. Structural parameters measured by electron microscopy on biopsy included GBM width, measured by the electron microscopic orthogonal intercept method22.

RASS genotyping: All RASS participants contributed DNA for genotyping on the Illumina HumanOmni1-Quad and HumanCoreExome beadchip arrays. Genotypes were called using BeadStudio/Genomestudio software (Illumina®). Quality control (QC) measures included removing duplicate samples, samples with evidence of contamination (heterozygosity range 0.25-0.32) and those with cryptic relatedness identity-by-state (IBS) (n=24). Principal component analyses were completed and 7 non-European outliers were removed. Of those genotyped, 1 participant was missing kidney biopsy data.

RASS GBM width analysis: We completed linear regression of the COL4A3 variant (rs55703767) and within person mean GBM width (nm) from both baseline and 5 year measures, in additive and genotypic genetic models. Both univariate and multivariate analyses were run including sex, baseline age and diabetes duration, within person mean HbA1c over 5 years, indicators for treatment group assignment and treatment center. A two-sided significance threshold of alpha <0.05 was applied.

In silico replication in SUMMIT consortium. The SUMMIT consortium included up to 5193 subjects with type 2 diabetes, with and without kidney disease, of European ancestry. All studies were approved by ethics committees from relevant institutions and all participants gave informed consent[23]. Complete list of SUMMIT Consortium members provided in Table S13.

SUMMIT genotyping and statistical analysis: SUMMIT Cohorts were genotyped on the Affymetrix SNP 6.0, the Illumina Omni express and the Illumina 610Quad arrays. QC measures included filtering out low frequency (<1% MAF) variants, filtering out low quality variants or samples, removal of duplicate samples, and removal of non-European samples based on principal component analysis.[23] Genome-wide association analyses were performed for DKD trait definitions harmonized with seven of our primary T1D analyses: "DN", "Micro", "Macro", "ESRD", "ESRD vs. non-ESRD", "CKD", and "CKD-DN" under an additive model, adjusting for age, gender and duration of diabetes.

RNA-sequencing and cis-eQTL analysis in human kidney samples from University of Pennsylvania cohort. Human kidney tissue collection was approved by the University of Pennsylvania Institutional Review Board. Kidney samples were obtained from surgical nephrectomies. Nephrectomies were de-identified, and the corresponding clinical information was collected through an honest broker; therefore, no consent was obtained from the subjects. Tubular and glomerular eQTL data sets were generated by 121 samples of tubules and 119 samples of glomeruli, respectively. The cis window was defined as 1 megabase up- and down-stream of the transcriptional start site (±1Mb). Whole kidney cis-eQTL (further just referred to as eQTL) data set was generated from 96 human samples were obtained from The Cancer Genome Atlas (TCGA) through the TCGA Data portal[24].

7

RNA-sequencing of human kidney samples in the University of Pennsylvania cohort: Human kidney tissue was manually microdissected under a microscope in RNAlater for glomerular and tubular compartments. The local renal pathologist performed an unbiased review of the tissue section by scoring multiple parameters, and RNA were prepared using RNAeasy mini columns (Qiagen, Valencia, CA) according to manufacturer's instructions. RNA quality was assessed with the Agilent Bioanalyzer 2100 and RNA integrity number scores above 7 were used for cDNA production. The library was prepared in the DNA Sequencing Core at University of Texas Southwestern Medical Center. One microgram total RNA was used to isolate poly(A) purified mRNA using the Illumina TruSeq RNA Preparation Kit. We sequenced samples for single-end 100bp, and the annotated RNA counts (fastq) were calculated by Illumina's CASAVA 1.8.2. Illumina sequence quality was surveyed with FastQC. Adaptor and lower-quality bases were trimmed with Trim-galore. Trimmed reads were aligned to the Gencode human genome(GRCh37) with STAR-2.4.1d. The readcount of each sample was obtained using HTSeq-0.6.1 (htseq-count) and then normalized fragments per kilobase million values were used to perform association analysis with fibrosis and sclerosis using linear regression.

Human kidney cis-eQTL analysis. Nominal p-values were calculated for each SNP-gene pair with FastQTL using linear regression with an additive effects model, and adjusted by six genotype PCs.

RNA-sequencing of human kidney samples. Normalized fragment per kilobase million values were used to perform association analysis with fibrosis and sclerosis using linear regression.

RNAseq and microarray profiling of human kidney samples from the Pima cohort. Kidney biopsy samples from the Pima Indian cohort were manually micro-dissected into 119 glomerular and 100 tubule-interstitial tissues to generate gene expression profiles[25]. Expression profiling in the Pima Indian cohort kidney biopsies was carried out using Affymetrix GeneChip Human

Genome U133 Array and U133Plus2 Array, as reported previously, and Affymetrix Human Gene ST Genechip 2.1[26, 27], and on RNA-seq (Illumina). The libraries were prepared using the ClonTech SMARTSeq v4 Ultra Low Input polyA selection kit. Samples were sequenced on a HiSeq 4000, single end, 75bp. Mapping to human reference genome GRCh38.7 was performed with STAR 2.5.2b (https://github.com/alexdobin/STAR). For annotation and quantification of mapping results we used cufflinks, cuffquant and cuffnorm in version 2.2.1 (https://cole-trapnell-lab.github.io/cufflinks/). After mapping and quantification, PCA and Hierarchical Clustering was used to identify outliers and reiterated until no more outliers could be identified.

eQTL analysis. Analysis was performed with Robust Multi-array Average quantile normalization[28] after removing probes overlapping with variants identified by WGS. Batch effects between platforms were corrected using ComBat[29] and unknown batch effects were also adjusted using singular value decomposition with first four eigenvectors. eQTL mapping was performed using EPACTS (https://genome.sph.umich.edu/wiki/EPACTS) software tool using linear mixed model accounting for hidden familial relatedness, after inverse Gaussian transformation of expression levels, adjusting for age and sex.

Mouse kidney single cell RNA-sequencing. Animal studies were approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania. We mated Cdh16[Cre] mice (Jackson Lab, 012237), Nphs2[Cre] mice (Jackson Lab, 008205) and SclCre mice (MGI number is 3579158) with Tomato-GFP (mT/mG) mice (Jackson Lab, 007576) to generate Cdh16[Cre]mT/mG, Scl[Cre]mT/mG and Nphs2cremT/mG mice[30].

Mouse kidney single cell RNA-sequencing: Kidneys were harvested from 4 to 8-week-old male mice with C57BL/6 background and dissociated into single cell suspension as described in our previous study31. The single cell sequencing libraries were sequenced on an Illumina HiSeq with 2x150 paired-end kit. The sequencing reads were demultiplexed, aligned to the mouse

genome (mm10) and processed to generate gene-cell data matrix using Cell Ranger 1.3 (http://10xgenomics.com)[31].

To calculate the average expression level for each cluster, a z-score of normalized expression value was first obtained for every single cell. Then, we calculated the mean z-scores for individual cells in the same cluster, resulting in 16 values for each gene.

Genomic features of human kidney. Human kidney-specific chromatin immunoprecipitation followed by sequencing (ChIP-seq) data can be found at GEO: GSM621634, GSM670025, GSM621648, GSM772811, GSM621651, GSM1112806, GSM621638. Different histone markers were combined into chromatin states using ChromHMM[32].

Gene and gene set analysis. PASCAL gene and pathway scores were conducted on all 20 sets of GWAS summary statistics (10 outcomes and 2 covariate models). Gene scores were derived using the sum option, averaging association signal across each gene using the default 50kb window size. Pathway scores were then computed from pathway member gene scores where membership was defined using default pathway libraries from BioCarta, REACTOME, and KEGG. Using a similar approach, MAGMA (v1.06) gene and pathway scores were conducted on all GWAS summary statistics using both the default gene region defined by the transcription start and stop sites and a 5kb window definition. MAGMA pathway analysis included all 1077 of the PASCAL reported libraries plus an additional 252 pathways included in MSigDB canonical pathway set. MAGENTA (vs2, July 2011) pathway analysis included 4725 pathways with a minimum of five genes within the gene set. Gene sets were obtained with the MAGENTA distribution and included Gene ontology terms, PANTHER sets (biological processes, molecular functions, metabolic and signaling pathways), KEGG pathways, and

Ingenuity pathways. DEPICT gene set enrichment uses a more comprehensive collection of gene sets that allows genes to have a continuous probability for gene set membership. We conducted DEPICT individually on all 20 sets of GWAS summary statistics with $P < 1.0 \times 10^{-5}$. We conducted two additional pooled analyses using genome-wide minimum P-values from: 1) All 20 analyses (10 phenotypes and 2 covariate models) and 2) Sixteen analyses of the 8 most related phenotypes (8 phenotypes and 2 covariate models) which excluded ESRD vs Macro and Micro.

Data and Software Availability

All cohorts can share genome-wide meta-analysis summary statistics. Individual level genotype data: due to restrictions set by study consents and by EU and national regulations, individual genotype data cannot be shared for all cohorts

Table S1. Cohorts contributing to analyses.

This table can be found in a separate excel sheet, Supplemental_table_S1.xlsx

Table S2. Characteristics of RASS participants. Categorical variables display counts and percentage. Continuous values are mean ± standard deviation.

Variables (Total N = 253)	Freq(%)/Mean±SD
Sex - Female	134 (53%)
Age (years)	30 ± 10
T1D duration (years)	11 ± 5
Within-person mean HbA1c (%) (mmol/mol)	8.6 ± 1.4 70 ± 15
Mean GBMW (nm)	480 ± 88
rs55703767 – GG GT TT	163 (64%) 80 (32%) 10 (4%)

Table S3. Multivariate analysis of association between rs55703767 and GBM width

Variables		Adjusted model			Fully adjusted model*	
		Effect (SE)	P		Effect (SE)	P
rs55703767 (T allele)[¶]		-22.8 (8.2)	0.006		-19.7 (8.2)[¶]	0.0172
Females (vs males)		-48.4 (9.3)	<.0001		-50.4 (9.3)	<.0001
Age at baseline (yrs)		-2.4 (0.5)	<.0001		-2.4 (0.5)	<.0001
Diabetes duration (yrs)		3.8 (1.0)	0.0002		3.8 (1.0)	0.0002
Mean HbA1c (%)		27.2 (3.3)	<.0001		27.4 (3.3)	<.0001
Treatments	Placebo	-	-		Reference	
	Enalapril	-	-		-6.9 (11.2)	0.538
	Losartan	-	-		1.4 (10.9)	0.896
Centres	Montreal	-	-		Reference	
	Toronto	-	-		0.8 (12.8)	0.952
	Minnesota	-	-		18.9 (13.7)	0.169
* Fully adjusted model also included 3 principal components for population structure within Europeans. [¶]SNP genotypes modelled as additive genetic effects.						

Table S4: Look-up of the lead loci in GWAS on eGFR in the general population (Gorski et al., 2017).[33]

Nearest Gene	SNP	EA	NEA	EAF	SE	OR	P_{min}	P_{Full}	EAF	β	SE	P	N
				Meta-analysis results					GWAS on eGFR (Gorski 2017)				
COL4A3	rs55703767	T	G	0.206	0.03	0.79	**$5.34{\times}10^{-12}$**	**$8.19{\times}10^{-11}$**	0.142	0.002	0.0013	0.132	110517
COL4A3	rs55703767	T	G	0.209	0.04	0.79	**$9.28{\times}10^{-9}$**	**$9.38{\times}10^{-9}$**	0.142	0.002	0.0013	0.132	110517
COL4A3	rs55703767	T	G	0.205	0.03	0.84	**$3.88{\times}10^{-10}$**	**$9.68{\times}10^{-9}$**	0.142	0.002	0.0013	0.132	110517
COL4A3	rs55703767	T	G	0.208	0.04	0.77	**$5.30{\times}10^{-9}$**	**$3.77{\times}10^{-8}$**	0.142	0.002	0.0013	0.132	110517
PRNCR1	rs551191707	CA	C	0.122	0.1	1.7	**$4.39{\times}10^{-8}$**	$3.15{\times}10^{-6}$					
STXBP6	rs61983410	T	C	0.787	0.04	1.26	$9.84{\times}10^{-8}$	**$3.06{\times}10^{-8}$**	0.841	-0.001	0.0012	0.336	110516
COLEC11	rs12615970	A	G	0.867	0.05	1.31	**$9.43{\times}10^{-9}$**	$1.60{\times}10^{-7}$					
LINC01266	rs115061173	A	T	0.014	0.41	9.39	**$4.07{\times}10^{-8}$**	$4.08{\times}10^{-5}$					
SNCAIP	rs149641852	T	G	0.012	0.39	9.03	**$1.37{\times}10^{-8}$**	---	0.009	0.002	0.0042	0.643	109257
PAPLN	rs113554206	A	G	0.012	0.3	4.62	$5.39{\times}10^{-7}$	**$8.46{\times}10^{-9}$**	0.007	-0.005	0.0064	0.408	95870
STAC	rs116216059	A	C	0.016	0.38	8.76	**$1.37{\times}10^{-8}$**	$1.41{\times}10^{-4}$	0.006	$3.00{\times}10^{-4}$	0.0043	0.953	108165
HAND2-AS1	rs145681168	A	G	0.986	0.33	0.18	$2.06{\times}10^{-7}$	**$5.40{\times}10^{-9}$**	0.993	0.003	0.0067	0.612	64752
TAMM41	rs142823282	A	G	0.983	0.31	0.15	**$8.32{\times}10^{-10}$**	**$1.13{\times}10^{-11}$**					
VARS2	rs118124843	T	C	0.011	0.24	3.78	**$4.42{\times}10^{-8}$**	**$3.37{\times}10^{-8}$**	0.031	0.011	0.0055	*0.040*	58794
MUC7	rs191449639	A	T	0.005	0.61	32.46	**$1.32{\times}10^{-8}$**	**$2.09{\times}10^{-8}$**					
MBLAC1	rs77273076	T	C	0.008	0.39	9.12	**$1.04{\times}10^{-8}$**	$2.28{\times}10^{-7}$	0.007	0.006	0.0051	0.236	108694
BMP7	rs144434404	T	C	0.011	0.32	6.75	**$2.67{\times}10^{-9}$**	**$4.65{\times}10^{-9}$**	0.004	$1.00{\times}10^{-4}$	0.0072	0.993	91428
PLEKHA7	rs183937294	T	G	0.993	0.5	0.06	**$1.65{\times}10^{-8}$**	$2.10{\times}10^{-6}$					
	rs185299109	T	C	0.007	0.53	20.7	**$1.28{\times}10^{-8}$**	$4.99{\times}10^{-7}$					

EA: Effect allele. Positive odds ratio indicates that EA is associated with higher risk; positive beta indicates that EA is associated with higher eGFR, i.e. lower renal risk.

Table S5: Look-up of the lead loci in GWAS in the SUMMIT consortium (van Zuydam et al., 2018).[23]

SNP	Chr:pos	EA	NEA	EAF	Notable gene(s)	Phenotype	N	OR	P-value
rs55703767	2:228121101	T	G	0.211	*COL4A3*	**DN**	**5190**	**0.911**	**0.08**
				0.213		**CKD+DN**	**2243**	**0.867**	**0.09**
rs145681168	4:174500806	G	A	0.017	*HAND2-AS1*	**Micro**	**3477**	**1.034**	**0.97**
rs149641852	5:121774582	T	G	0.018	*SNCAIP*	**CKD**	**4676**	**1.032**	**0.30**
rs118124843	6:30887465	T	C	0.018	*DDR1, VARS2*	**Micro**	**2439**	**1.137**	**0.63**
rs77273076	7:99728546	T	C	0.014	*MBLAC1*	Micro	3252	0.866	0.48
rs61983410	14:26004712	T	C	0.184	*STXBP6*	**Micro**	**3760**	**0.990**	**0.58**
rs144434404	20:55837263	T	C	0.011	*BMP7*	**Micro**	**2439**	**1.100**	**0.78**

Chr, chromosome; pos, position; EA: Effect allele; EAF, effect allele frequency; OR, odds ratio.

Table S6: Association at lead loci stratified by HbA1c <7.5%.

Locus	SNP	Pheno	EA	NEA	ALL				HbA1c < 7.5%			HbA1c >= 7.5%			P_HET
					N	MAF	P	INFO	N (case/ctrl)	P	OR (95% CI)	N (case/ctrl)	P	OR (95% CI)	
COL4A3	rs55703767	MACROESRD	G	T	3611	0.19	**2.16E-03**	1.00	1165 (499/666)	0.659	0.95 (0.76;1.19)	2495 (884/1611)	**9.55E-04**	0.77 (0.66;0.9)	0.132
COL4A3	rs55703767	MACRO	G	T	2803	0.19	0.06	1.00	837 (164/673)	0.663	1.08 (0.78;1.49)	2006 (373/1633)	**6.63E-03**	0.75 (0.61;0.92)	0.068
COL4A3	rs55703767	ALLvCTRL	G	T	4271	0.19	**7.04E-03**	1.00	1344 (692/652)	0.870	0.98 (0.8;1.2)	2977 (1391/1586)	**1.76E-03**	0.81 (0.71;0.92)	0.114
COL4A3	rs55703767	CKDDN	G	T	3059	0.19	**1.17E-02**	1.00	984 (379/605)	0.973	1 (0.78;1.28)	2102 (624/1478)	**7.90E-03**	0.79 (0.67;0.94)	0.136
PRNCR1	rs551191707	ESRDvMACRO	C	CA	1371	0.14	**2.50E-03**	0.81	498 (340/158)	**1.92E-02**	1.71 (1.09;2.67)	885 (524/361)	**4.79E-02**	1.38 (1;1.91)	0.453
STXBP6	rs61983410	MICRO	T	C	2976	0.23	**3.75E-03**	0.93	863 (195/668)	**1.34E-02**	0.69 (0.52;0.93)	2155 (526/1629)	0.067	0.85 (0.71;1.01)	0.248
COLEC11	rs12615970	CKD	A	G	4264	0.14	**3.13E-03**	0.82	1432 (531/901)	0.086	0.81 (0.63;1.03)	3014 (833/2181)	**1.62E-02**	0.8 (0.66;0.96)	0.949
LINC01266	rs115061173	ESRD	T	A	3119	0.00	**1.89E-02**	0.36	1012 (340/672)	0.284	2.63 (0.45;15.36)	2156 (524/1632)	0.085	5.98 (0.78;45.9)	0.550
SNCAIP	rs149641852	CKDEXTREMES	G	T	3907	0.01	**3.04E-03**	0.33	1323 (415/908)	0.559	1.53 (0.37;6.35)	2765 (559/2206)	**6.27E-04**	10.78 (2.76;42.09)	0.052
PAPLN	rs113554206	MACRO	G	A	2803	0.00	0.32	0.34	837 (164/673)	0.793	0.39 (0;417.59)	2006 (373/1633)	0.114	21.87 (0.48;999.19)	0.322
STAC	rs116216059	ESRDvALL	C	A	4272	0.01	0.48	0.67	1340 (340/1000)	0.867	0.88 (0.2;3.89)	2984 (524/2460)	0.453	0.67 (0.23;1.93)	0.764
HAND2-AS1	rs145681168	MICRO	A	G	2976	0.01	0.50	0.48	863 (195/668)	0.509	3.48 (0.09;141.38)	2155 (526/1629)	0.395	0.61 (0.19;1.91)	0.378
TAMM41	rs142823282	MICRO	A	G	2976	0.00	0.93	0.15				2155 (526/1629)	0.886	1.19 (0.11;13.33)	NA
VARS2	rs118124843	MICRO	C	T	2976	0.01	0.93	1.00	863 (195/668)	0.533	0.59 (0.11;3.16)	2155 (526/1629)	0.769	1.15 (0.46;2.85)	0.492
MUC7	rs191449639	MACROESRD	T	A	3611	0.00	0.09	0.28	1165 (499/666)	0.487	2.17 (0.24;19.36)	2495 (884/1611)	0.246	3.58 (0.42;30.94)	0.749
MBLAC1	rs77273076	MICRO	C	T	2976	0.01	**1.36E-04**	0.37	863 (195/668)	**3.98E-03**	168.26 (5.14;5507.78)	2155 (526/1629)	**1.68E-03**	11.25 (2.48;50.97)	0.163
BMP7	rs144434404	MICRO	C	T	2976	0.01	0.57	0.67	863 (195/668)	0.407	0.49 (0.09;2.61)	2155 (526/1629)	0.911	0.94 (0.3;2.91)	0.534
PLEKHA7	rs183937294	MICRO	T	G	2976	0.00	0.22	0.26				2155 (526/1629)	0.288	3.79 (0.32;44.52)	NA
18p11.32	rs185299109	CKD	C	T	4264	0.00	0.68	0.32	1432 (531/901)	0.147	0.12 (0.01;2.09)	3014 (833/2181)	0.956	0.96 (0.21;4.45)	0.212

Association stratified by HbA1c in the FinnDiane study. *P*-values <0.05 are given with scientific notation and bold. Lines with gray text had minor allele count (MAC)<10 in cases and/or controls and did not contribute to the meta-analysis.

Table S7. Association of rs55703767 with DN in DCCT/EDIC subgroups.

Cohort	Treatment Group	DN %	MAF	Last measure		Time to Event	
				OR (95%CI)	P value	HR (95%CI)	P value
Primary Prevention (diabetes dur 1-5 yrs)	Intensive	3%	0.22	2.86 (0.4-22)	0.32	0.91 (0.2-4.0)	0.90
	Conventional	10%	0.21	0.67 (0.3-1.4)	0.31	0.66 (0.32-1.33)	0.24
Secondary Intervention (diabetes dur 1-15 yrs)	Intensive	5%	0.20	0.86 (0.3-2.6)	0.79	0.65 (0.22-1.9)	0.43
	Conventional	13%	0.22	0.18 (0.1-0.5)	0.003	0.30 (0.13-0.68)	0.004

OR=Odds Ratio for last measure, HR=Hazard Ratio for time to event phenotype

Table S8. Significant (*P*<0.05/18,222 genes tested = 2.74 × 10^{-6}) gene level associations with diabetic kidney disease in MAGMA.

Gene	Phenotype	Model	Window	Number of SNPs	Total Sample Size	MAGMA P-value	PASCAL P-value
SLC46A2	All vs. ctrl	Min	nowindow	66	17817	6.74 × 10^{-7}	1.57 × 10^{-5}
			5kbwindow	93	17832	7.38 × 10^{-7}	
		Full	nowindow	64	16821	8.13 × 10^{-7}	6.93 × 10^{-5}
			5kbwindow	90	16855	1.03 × 10^{-6}	
SFXN4	Macro	Full	nowindow	69	11953	3.98 × 10^{-7}	1.45 × 10^{-4}
			5kbwindow	86	11857	1.65 × 10^{-7}	
COL20A1	Ckdextreme	Min	nowindow	111	11165	2.47 × 10^{-6}	7.88 × 10^{-5}
			5kbwindow	137	11603	2.01 × 10^{-6}	
		Full	nowindow	110	8533	6.65 × 10^{-7}	4.47 × 10^{-5}
			5kbwindow	136	9044	5.77 × 10^{-7}	
	ESRD vs. All	Min	nowindow	111	12063	1.34 × 10^{-6}	3.76 × 10^{-5}
			5kbwindow	137	12362	1.04 × 10^{-6}	
		Full	nowindow	110	8638	1.12 × 10^{-6}	5.81 × 10^{-5}
			5kbwindow	136	9045	9.53 × 10^{-7}	
GLT6D1	ESRD vs. Macro	min	5kbwindow	96	4248	1.49 × 10^{-6}	2.15 × 10^{-5}
SNX30	All vs. ctrl	min	5kbwindow	434	18249	2.49 × 10^{-6}	1.05 × 10^{-5}

Table S9. Top nominally significant gene level associations (P < 1.0 x 10^{-5}) with diabetic kidney disease in PASCAL.

Gene	Phenotype	Model	Number of SNPs	PASCAL P-value
INIP	All vs. ctrl	Min	248	1.99 × 10^{-6}
		Full	248	5.54 × 10^{-6}
LCN9	ESRD vs. macro	Min	301	5.25 × 10^{-6}
CBX8	DN	Min	119	8.47 × 10^{-6}

Table S10: Significant gene set and pathway analysis results. Significantly enriched gene sets identified from at least one of the following methods: MAGENTA (FDR < 0.05, MAGMA (P<0.05 empirical permutation multiple testing correction), PASCAL (P<0.05/1,078 gene sets tested = 4.64 × 10-5), and DEPICT (FDR < 0.01).

Gene set	Gene set database	Phenotype	Model	Method
negative regulators of RIG I MDA5 signaling	REACTOME	ESRD vs. Macro	Full	MAGMA
Platelet aggregation plug formation	REACTOME	Micro	Min	MAGMA
negative regulators of RIG I MDA5 signaling	REACTOME	ESRD vs. Macro	Full	PASCAL
RIG I MDA5 mediated induction of IFN alpha beta pathways	REACTOME	ESRD vs. Macro	Full	PASCAL
TRAF3 dependent IRF activation pathway	REACTOME	ESRD vs. Macro	Full	PASCAL
TRAF6 mediated IRF activation	REACTOME	ESRD vs. Macro	Full	PASCAL
Nitric Oxide Signaling in the Cardiovascular System	Ingenuity	ESRD vs. ctrl	Min	MAGENTA
Nicotinic acetylcholine receptor signaling pathway	Panther	ESRD vs. non-ESRD	Min	MAGENTA
ACTIVATED TLR4 SIGNALLING	REACTOME	All vs. ctrl	Min	MAGENTA
Other lipid, fatty acid and steroid metabolism	PANTHER BIOLOGICAL PROCESS	CKD	Min	MAGENTA
DNA degradation	PANTHER BIOLOGICAL PROCESS	CKD	Min	MAGENTA
Tumor necrosis factor family member	PANTHER MOLECULAR FUNCTION	CKD-extreme	Min	MAGENTA
TUFM (Tu Translation Elongation Factor, Mitochondrial) PPI subnetwork	InWeb protein-protein interaction database	DN	Min	DEPICT

Table S11. eQTL associations and chromatin conformation interactions for the lead SNPs.

SNP	Chr:pos	EA	NEA	EAF	Notable gene(s)	eQTL				PC-HiC	
						GENE	P	HIGH A	Tissue	Gene	Score (Tissue)
rs12615970	2:3745215	G	A	0.133	COLEC11 (B); ALLC (G)					**ALLC**	10.42 (GM12878);
										COLEC11, AC010907.2	9.67 (GM12878);
										ADI1, AC142528.1	8.75 (GM12878);
										RP13-512J5.1	8.58 (GM12878);
										RPS7	8.13 (GM12878);
rs55703767	2:228121101	T	G	0.206	COL4A3 (M, B, N)	**MFF**	5.63×10⁻³⁸	T	blood	**COL4A3, COL4A4**	8.89 (GM12878);
						MFF	9.0×10⁻⁸	T	Cells - Transformed fibroblasts	**IRS**, RP11-395N3.2	9.36 (GM12878);
						TM4SF20	2.2×10⁻⁷	T	Cells - Transformed fibroblasts		
rs115061173	3:926345	A	T	0.014	LINC01266 (N)						
rs142823282	3:11910635	G	A	0.011	TAMM41 (N, B)	**PPARG**	4.6×10⁻⁷	G	Colon - Sigmoid	**TAMM41**	10.65 (GM12878);
rs116216059	3:36566312	A	C	0.016	STAC (G)					**DCLK3**	8.8 (GM12878);
										STAC	10.87 (GM12878);
rs191449639	4:71358776	A	T	0.005	MUC7 (N)						
rs145681168	4:174500806	G	A	0.014	HAND2-AS1 (G, B)					**HAND2**, HAND2-AS1	10.49 (GM12878);
rs149641852	**5:121774582**	**T**	**G**	**0.012**	**SNCAIP (G)**					**SNX24**	9.2 (GM12878);
										snoU13	8.93 (GM12878);
										SNCAIP, CTD-2544H17.2	9.69 (GM12878);
										CTD-2280E9.1	10.81 (GM12878);
rs118124843	6:30887465	T	C	0.011	DDR1 (B);	**HLA-C**	1.00×10⁻¹⁸	C	eQTLgen blood	**PSORS1C1**	12.3 (Endothelial Precursors); 12.3 (Endothelial Precursors); 7.84 (Megacaryocytes); 5.63 (Pancreatic islets);
					VARS2 (G)					**DPCR1, HCG21**	11.49 (GM12878);
						HLA-U	3.56×10⁻¹⁰	T	eQTLgen blood	DDR1-AS1, **DDR1**	10.97 (GM12878);
						PSORS1C3	4.13×10⁻⁹	T	eQTLgen blood	RNU6-1133P	7.27 (Macrophages M2);

SNP	Chr:pos	EA	NEA	EAF	Notable gene(s)	eQTL				PC-HiC	
						GENE	P	HIGH A	Tissue	Gene	Score (Tissue)
						NCR3	9.35×10⁻⁶	C	eQTLgen blood	*RN7SL175P, **DDR1, GTF2H4, VARS2***	7.07 (Endothelial Precursors); 7.07 (Endothelial Precursors); 6.95 (Cardiomyocytes); 6.89 (Pancreatic islets); 5.43 (Megacaryocytes); 5.09 (Macrophages M1);
						HCG22	1.5×10⁻⁵	T	eQTLgen blood	*C6orf15*	6.95 (Macrophages M1); 6.95 (Macrophages M1); 6.18 (Macrophages M0);
						VARS2	1.71×10⁻⁵	C	eQTLgen blood		
						GTF2H4	9.70×10⁻⁷	T	Esophagus - Gastroesophageal Junction		
						POU5F1	3.3×10⁻⁵	T	Esophagus - Gastroesophageal Junction		
						PSORS1C3	5.6×10⁻⁵	T	Esophagus - Gastroesophageal Junction		
						C6orf48	1×10⁻⁴	C	Nerve - Tibial		
rs77273076	7:99728546	T	C	0.008	*MBLAC1* (N, B)	*CNPY4*	1.17×10⁻⁷	C	eQTLgen blood	***MBLAC1,** AC073842.19, RP11-506M12.1*	**NA (withn the same fragment);**
						AP4M1	1.04×10⁻⁵	C	eQTLgen blood	***LAMTOR4, GAL3ST4, GPC2,** C7orf43, MIR4658*	14.82 (CD34); 14.82 (CD34); 14.56 (GM12878);
						ZSCAN21	1.29×10⁻⁵	C	eQTLgen blood	***LAMTOR4***	14.14 (CD34); 14.14 (CD34); 13.1 (GM12878);
										***GATS, STAG3, PVRIG,** AC005071.1*	14.11 (CD34); 14.11 (CD34); 13.68 (GM12878);
										MCM7, AP4M1	14.11 (CD34); 14.11 (CD34); 13.68 (GM12878);
										STAG3, GPC2	13.75 (CD34); 13.75 (CD34); 13.36 (GM12878);
										***MCM7, COPS6,** MIR93, MIR106B, MIR25*	13.69 (CD34); 13.69 (CD34); 13.67 (GM12878);
										CNPY4, TAF6	13.37 (GM12878); 13.37 (GM12878); 13.14 (CD34);
										ZKSCAN1	13.23 (GM12878); 13.23 (GM12878); 11.87 (CD34);
										ZSCAN21	13.14 (GM12878);

SNP	Chr:pos	EA	NEA	EAF	Notable gene(s)	eQTL			PC-HiC	
						GENE	*P*	*HIGH A Tissue*	*Gene*	*Score (Tissue)*
									ZCWPW1, MEPCE	12.72 (GM12878); 12.72 (GM12878); 11.48 (CD34);
									PILRA	12.65 (GM12878);
									TRIM4	12.64 (GM12878);
									ZCWPW1	12.61 (GM12878);
									SAP25, FBXO24, LRCH4, RP11-44M6.3	12.59 (GM12878);
									PILRB, PVRIG2P, STAG3L5P-PVRIG2P-PILRB,	12.38 (GM12878); 12.38 (GM12878); 5.94 (Naive B); 5.02 (Total B);
									TSC22D4, NYAP1, AC092849.1, RN7SL161P, C7orf61	12.21 (GM12878);
									RP11-758P17.2, PPP1R35, RP11-758P17.3	12.02 (GM12878); 12.02 (GM12878); 11.55 (CD34);
									AZGP1, AZGP1P1	11.3 (Neutrophils); 11.3 (Neutrophils); 6.07 (Macrophages M2); 5.65 (Total CD4 MF); 5.65 (Total CD4 MF); 5.1 (Total CD4 Activated);
									BUD31, snoU13	11.03 (GM12878);
									ZNF3	10.83 (GM12878);
									ZCWPW1	8.39 (Naive B); 8.39 (Naive B); 5.52 (Total CD4 Activated);
									PMS2P1	5.16 (Foetal Thymus);
rs551191707	8:128100029	CA	C	0.122	*PRNCR1* (N)	*-NONE-*				
rs183937294	11:16937846	G	T	0.007	*PLEKHA7* (G)	*-NONE-*			RNU6-585, PRP11-466H18.1	10.77 (cd34); 10.77 (cd34); 9.2 (GM12878);
									AC116533.1, SNORD14B, SNORD14A, rps13	10.19 (GM12878);
									PLEKHA7, OR7E14P	9.49 (GM12878);

SNP	Chr:pos	EA	NEA	EAF	Notable gene(s)	eQTL			PC-HiC	
						GENE	P	HIGH A Tissue	*Gene*	*Score (Tissue)*
									SOX6, *C11orf58*	9.15 (GM12878);
									SERGEF, *RP1-59M18.2*	7.56 (GM12878);
									OTOG	5.89 (Total CD8);
									USH1C	5.07 (Naive CD8);
rs61983410	14:26004712	T	C	0.213	*STXBP6* (N)	*-NONE-*			*SNORD37*	10.97 (GM12878);
rs113554206	14:73740250	A	G	0.012	*PAPLN* (G)	*-NONE-*			*RP4-647C14.3*	***NA (within the same fragment);***
									NUMB	21.42 (Endotheial precursors); 21.42 (Endotheial precursors); 17.38 (Pancreatic islets); 12.32 (Megacaryocytes); 11.56 (CD34); 9.65 (Neutrophils); 9.61 (Naive B); 9.06 (Total B); 7.55 (cardiomyocytes); 5.33 (Naive CD4); 5.11 (Naive CD8);
									PAPLN, *RNU6-419P, RP4-647C14.2*	13.67 (CD34); 13.67 (CD34); 13.24 (GM12878);
									PAPLN	13.24 (CD34); 13.24 (CD34); 12.47 (GM12878);
									PSEN1	12.08 (GM12878); 12.08 (GM12878); 5.4 (cardiomyocytes);
									HEATR4	12.05 (Endotheial precursors); 12.05 (Endotheial precursors); 10.98 (Megacaryocytes); 10.45 (GM12878); 10.37 (CD34); 9.92 (Neutrophils); 8.84 (Pancreatic islets); 7.98 (Monocytes); 7.3 (Total B); 7.02 (Naive B); 6.6 (cardiomyocytes); 5.81 (Erythroblasts); 5.53 (Total CD4 Activated);
									RP1-240K6.3	11.84 (GM12878); 11.84 (GM12878); 11.15 (CD34); 7.49 (Endotheial precursors); 5.62 (Pancreatic islets); 5.34

SNP	Chr:pos	EA	NEA	EAF	Notable gene(s)	eQTL			PC-HiC	
						GENE	*P*	*HIGH A Tissue*	*Gene*	*Score (Tissue)*
										(Total B); 5.02 (Megacaryocytes);
									DNAL1, RNU6-240P	11.07 (GM12878);
									PSEN1	10.96 (Endotheial precursors); 10.96 (Endotheial precursors); 10.56 (CD34);
									PNMA1	10.9 (GM12878);
									RP3-414A15.2	10.64 (GM12878); 10.64 (GM12878); 7.24 (Monocytes); 5.99 (Neutrophils);
									ZFYVE1	10.53 (GM12878);
									PTGR2, Y_RNA, RP5-1021I20.4	10.42 (CD34); 10.42 (CD34); 10.16 (GM12878);
									RP4-693M11.3	10.33 (GM12878);
									RP4-687K1.2	9.96 (GM12878);
									HEATR4, C14orf169, AC005280.1	9.52 (GM12878); 9.52 (GM12878); 5.35 (Pancreatic islets);
									RBM25	9.34 (GM12878);
									RP3-414A15.10	9.32 (GM12878);
									ELMSAN1	9.12 (GM12878);
									CCDC176	8.85 (GM12878);
									RBM25, RP11-109N23.5	8.74 (GM12878);
									ACOT6	8.74 (GM12878);
									DNAL1	8.7 (GM12878);
									FAM161B, RP5-1021I20.5	5.58 (Total CD8);
rs185299109	18:1811108	T	C	0.007		*-NONE-*			*-NONE-*	
rs144434404	**20:55837263**	**T**	**C**	**0.011**	*BMP7* **(G, B)**	*-NONE-*			*-NONE-*	

SNP	Chr:pos	EA	NEA	EAF	Notable gene(s)	eQTL				PC-HiC	
						GENE	P	HIGH A	Tissue	*Gene*	*Score (Tissue)*

Notable Genes: based on genetic findings (G), (B), (N), (M); eQTL associations were searched from GTEX and eQTLgen (cis-eQTL) data sets. HIGH A: allele associated with higher gene expression levels. Promoter Capture Hi-C (PCHi-C) data: searched from www.chicp.org (date accessed: 1.12.2018; Schofield EC, Carver T, Achuthan P, Freire-Pritchett P, Spivakov M, Todd JA, Burren OS. CHiCP: a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets. Bioinformatics. (2016) 15:32(16):2511-3), including 16 primary blood cell types and foetal thymocytes (Javierre et al.), CD34 and GM12878 cell line (Mifsud et al.), pancreatic isles (Miguel-Escalada et al,), and hESC derived cardiomyocytes (Choy et al.). Score: CHiCAGO score, values >5 were considered significant and listed. Protein coding genes are highleghted with bold typing.

Table S12: Transcriptome-wide association analysis (TWAS) results with p<1×10^{-4}

tissue	GWAS phenotype	GENE	Z Score	Effect	P-value	Var_G	r2	P-value	Q-value	used	in cov	in model
								Prediction performance			**N SNPs**	
tub	ESRD vs macro_min	ACOT8	4.02	0.76	5.80E-05	0.06	0.04	2.46E-02	2.02E-02	30	30	31
tub	DN_min	AKIRIN2	4.17	0.32	3.03E-05	0.09	0.05	1.30E-02	1.19E-02	36	39	42
tub	Macro_min	AKIRIN2	4.36	0.42	1.32E-05	0.09	0.05	1.30E-02	1.19E-02	36	39	42
glom	Macro_min	ARL17B	-4.02	-0.21	5.91E-05	0.32	0.51	1.28E-19	1.77E-18	53	54	65
tub	All vs ctrl_min	CALCOCO2	3.93	0.29	8.55E-05	0.06	0.10	3.23E-04	4.93E-04	37	37	40
glom	All vs ctrl_max	EXOC2	4.10	0.20	4.06E-05	0.15	0.21	1.74E-07	3.79E-07	60	60	62
glom	ESRD vs non-ESRD_max	FAM132B	-3.95	-0.52	7.73E-05	0.06	0.07	4.84E-03	3.91E-03	33	33	35
tub	CKD extreme_min	FES	-3.89	-0.40	9.97E-05	0.08	0.09	7.57E-04	1.05E-03	34	34	34
tub	ESRD_min	FES	-3.97	-0.42	7.07E-05	0.08	0.09	7.57E-04	1.05E-03	34	34	34
tub	ESRD vs non-ESRD_min	FES	-4.24	-0.42	2.27E-05	0.08	0.09	7.57E-04	1.05E-03	34	34	34
glom	ESRD vs macro_min	GSDMB	-3.96	-0.46	7.37E-05	0.11	0.07	4.54E-03	3.71E-03	35	35	35
glom	Macro_max	HOXD1	4.02	0.88	5.71E-05	0.01	0.03	5.51E-02	3.10E-02	12	12	13
glom	Macro_min	HOXD1	4.13	0.83	3.70E-05	0.01	0.03	5.51E-02	3.10E-02	12	12	13
glom	DN_max	ITPR3	4.30	0.21	1.74E-05	0.19	0.38	1.09E-13	6.21E-13	33	33	35
glom	DN_min	ITPR3	4.00	0.17	6.26E-05	0.19	0.38	1.09E-13	6.21E-13	33	33	35
glom	Macro_max	ITPR3	4.31	0.25	1.62E-05	0.19	0.38	1.09E-13	6.21E-13	33	33	35
glom	Macro_min	ITPR3	4.03	0.21	5.69E-05	0.19	0.38	1.09E-13	6.21E-13	33	33	35
glom	ESRD_min	MORC1	3.94	0.44	8.03E-05	0.06	0.08	1.58E-03	1.47E-03	83	87	88
glom	Macro_max	NLN	4.07	0.26	4.61E-05	0.21	0.15	1.09E-05	1.71E-05	85	86	110
glom	Macro_min	NLN	4.53	0.27	5.99E-06	0.21	0.15	1.09E-05	1.71E-05	85	86	110
glom	All vs ctrl_max	NPNT	3.90	0.37	9.51E-05	0.08	0.16	5.26E-06	8.75E-06	2	2	4
glom	ESRD_min	PRC1	4.03	0.41	5.54E-05	0.09	0.11	3.23E-04	3.59E-04	33	33	34
tub	CKD-DN_min	PRRC2C	3.92	0.81	8.74E-05	0.02	0.03	4.40E-02	3.27E-02	10	10	10
tub	ESRD_min	PRRC2C	3.92	0.91	8.76E-05	0.02	0.03	4.40E-02	3.27E-02	10	10	10
tub	Macro_min	TENM2	3.98	0.83	6.92E-05	0.02	0.11	1.59E-04	2.64E-04	23	59	60
glom	DN_min	VPS33B	4.09	0.24	4.40E-05	0.15	0.37	2.23E-13	1.19E-12	24	24	26
glom	ESRD_min	VPS33B	4.20	0.35	2.69E-05	0.15	0.37	2.23E-13	1.19E-12	24	24	26
tub	CKD extreme_max	VPS9D1	-4.17	-0.54	3.01E-05	0.06	0.11	2.78E-04	4.30E-04	20	20	21

Tissue: glomeruli (glom) or tubuli (tub); Z-core, Effect and P-value: MetaXcan's association results for the gene. Var_g: variance of the gene expression, calculated as W' * G * W (where W is the vector of SNP weights in a gene's model, W' is its transpose, and G is the covariance matrix). Prediction performance r2, P-value and Q-value: r2, p-value and q-value of tissue model's correlation to gene's measured transcriptome (prediction performance). N SNPs ... used: number of snps from GWAS that got used in MetaXcan analysis; ... in cov: number of snps in the covariance matrix; ... in model: number of snps in the model

Table S13: Pseudo-R2 of all SNPs across all GWAS as calculated by the McKelvey and Zavoina method.[34] Total variance explained is the sum of pseudo-R2 across all SNPs with minor allele frequency (MAF) greater than 5% or 1%, noting that effect size and therefore variance explained tend to be overestimated with rare variants. Missing values indicate SNPs that did not pass our GWAS filters for those disease definitions as described in the methods section.

Minimally Adjusted Model

SNP	Minor allele frequency	DN	All vs. ctrl	CKD	CKD-DN	CKD extreme	ESRD vs. ctrl	ESRD vs. non-ESRD	ESRD vs. macro	Macro	Micro
rs61983410	0.213	0.00%	0.09%	0.01%	0.00%	0.04%	0.01%	0.04%	0.05%	0.01%	0.54%
rs55703767	0.206	0.57%	0.33%	0.23%	0.65%	0.34%	0.55%	0.27%	0.03%	0.60%	0.11%
rs12615970	0.133	0.16%	0.06%	0.52%	0.48%	0.41%	0.26%	0.21%	0.08%	0.05%	0.00%
rs551191707	0.122	0.02%	0.00%	0.14%	0.33%	0.70%	0.69%	0.75%	1.76%	0.06%	0.01%
rs142823282	0.017	0.04%	0.58%	0.01%	NA	NA	NA	NA	NA	0.11%	3.50%
rs116216059	0.016	0.00%	0.00%	0.13%	0.23%	2.96%	1.95%	4.40%	NA	0.01%	0.00%
rs145681168	0.014	0.01%	0.15%	0.00%	0.08%	0.17%	NA	NA	NA	0.03%	2.41%
rs115061173	0.014	0.47%	0.24%	0.34%	1.44%	2.24%	3.96%	2.57%	NA	0.12%	0.01%
rs113554206	0.012	0.97%	0.27%	0.42%	NA	NA	NA	NA	NA	1.64%	NA
rs149641852	0.012	0.12%	0.02%	0.21%	2.14%	3.39%	1.94%	1.30%	NA	0.07%	0.03%
rs144434404	0.011	0.05%	0.38%	0.12%	NA	NA	NA	NA	NA	NA	2.43%
rs118124843	0.011	0.06%	0.22%	0.05%	0.09%	NA	NA	NA	NA	NA	1.17%
rs77273076	0.008	0.08%	0.30%	0.09%	0.12%	NA	NA	NA	NA	NA	2.28%
rs183937294	0.007	0.09%	0.47%	NA	NA	NA	NA	NA	NA	NA	3.49%
rs185299109	0.007	0.08%	0.05%	3.84%	NA	NA	NA	NA	NA	NA	NA
rs191449639	0.005	3.46%	0.17%	NA	NA	NA	NA	NA	NA	NA	NA
variance explained (MAF>5%)		0.75%	0.48%	0.89%	1.47%	1.49%	1.50%	1.26%	1.92%	0.73%	0.65%
variance explained (MAF>1%)		2.46%	2.34%	2.17%	5.44%	10.26%	9.36%	9.53%	1.92%	2.68%	10.21%

Fully Adjusted Model

SNP	Minor allele frequency	DN	All vs. ctrl	CKD	CKD-DN	CKD extreme	ESRD vs. ctrl	ESRD vs. non-ESRD	ESRD vs. macro	Macro	Micro
rs61983410	0.213	**0.01%**	**0.14%**	**0.02%**	**0.01%**	**0.04%**	**0.00%**	**0.02%**	**0.01%**	**0.02%**	**0.63%**
rs55703767	0.206	**0.60%**	0.31%	0.25%	0.72%	0.33%	0.51%	0.27%	0.02%	0.70%	0.11%
rs12615970	0.133	**0.17%**	0.07%	0.50%	0.46%	0.34%	0.22%	0.22%	0.29%	0.04%	0.00%
rs551191707	0.122	**0.01%**	0.00%	0.08%	0.23%	0.41%	0.49%	0.49%	1.88%	0.07%	0.00%
rs142823282	0.017	**0.04%**	0.86%	0.01%	NA	NA	NA	NA	NA	0.12%	4.65%
rs116216059	0.016	**0.00%**	0.00%	0.03%	NA	NA	NA	NA	NA	0.00%	0.00%
rs145681168	0.014	**0.01%**	0.18%	0.00%	NA	NA	NA	NA	NA	NA	3.33%
rs115061173	0.014	**0.61%**	0.26%	0.40%	2.53%	2.11%	3.01%	1.63%	NA	NA	0.00%
rs113554206	0.012	**1.12%**	0.26%	NA	NA	NA	NA	NA	NA	3.99%	NA
rs149641852	0.012	**0.27%**	0.04%	0.74%	2.69%	NA	NA	NA	NA	0.01%	0.05%
rs144434404	0.011	**0.03%**	0.40%	0.11%	NA	NA	NA	NA	NA	NA	2.38%
rs118124843	0.011	**0.12%**	0.34%	0.17%	NA	NA	NA	NA	NA	NA	1.23%
rs77273076	0.008	**NA**	0.37%	0.04%	NA	NA	NA	NA	NA	NA	1.80%
rs183937294	0.007	**NA**	0.90%	NA	NA	NA	NA	NA	NA	NA	NA
rs185299109	0.007	**NA**	0.00%	NA	NA	NA	NA	NA	NA	NA	NA
rs191449639	0.005	**3.60%**	0.26%	NA	NA	NA	NA	NA	NA	NA	NA
variance explained (MAF>5%)		**0.78%**	0.52%	0.84%	1.41%	1.12%	1.22%	1.00%	2.20%	0.83%	0.74%
variance explained (MAF>1%)		**2.98%**	2.86%	2.30%	6.62%	3.23%	4.23%	2.63%	2.20%	4.95%	12.38%

Table S14. Physicians and nurses at health care centers participating in the collection of FinnDiane patients.

FinnDiane Study Centers	Physicians and nurses
Anjalankoski Health Centre	S. Koivula, T. Uggeldahl
Central Finland Central Hospital, Jyväskylä	T. Forslund, A. Halonen, A. Koistinen, P. Koskiaho, M. Laukkanen, J. Saltevo, M. Tiihonen

FinnDiane Study Centers	Physicians and nurses
Central Hospital of Åland Islands, Mariehamn	M. Forsen, H. Granlund, A-C. Jonsson, B. Nyroos
Central Hospital of Kanta-Häme, Hämeenlinna	P. Kinnunen, A. Orvola, T. Salonen, A. Vähänen
Central Hospital of Länsi-Pohja, Kemi	H. Laukkanen, P. Nyländen, A. Sademies
Central Ostrabothnian Hospital District, Kokkola	S. Anderson, B. Asplund, U. Byskata, P. Liedes, M. Kuusela, T. Virkkala
City of Espoo Health Centre	
Espoonlahti	A. Nikkola, E. Ritola
Tapiola	M. Niska, H. Saarinen
Samaria	E. Oukko-Ruponen, T. Virtanen
Viherlaakso	A. Lyytinen
City of Helsinki Health Centre	
Puistola	H. Kari, T. Simonen
Suutarila	A. Kaprio, J. Kärkkäinen, B. Rantaeskola
Töölö	P. Kääriäinen, J. Haaga, A-L. Pietiläinen
City of Hyvinkää Health Centre	S. Klemetti, T. Nyandoto, E. Rontu, S. Satuli-Autere
City of Vantaa Health Centre	
Korso	R. Toivonen, H. Virtanen
Länsimäki	R. Ahonen, M. Ivaska-Suomela, A. Jauhiainen
Martinlaakso	M. Laine, T. Pellonpää, R. Puranen
Myyrmäki	A. Airas, J. Laakso, K. Rautavaara
Rekola	M. Erola, E. Jatkola
Tikkurila	R. Lönnblad, A. Malm, J. Mäkelä, E. Rautamo
Heinola Health Centre	P. Hentunen, J. Lagerstam
Helsinki University Central Hospital, Department of Medicine, Division of Nephrology	A. Ahola, J. Fagerudd, M. Feodoroff, D. Gordin, O. Heikkilä, K Hietala, L. Kyllönen, J. Kytö, S. Lindh, K. Pettersson-Fernholm, M. Rosengård-Bärlund, M. Rönnback, A. Sandelin, A-R Salonen, L. Salovaara, L. Thorn, J. Tuomikangas, T. Vesisenaho, J. Wadén
Herttoniemi Hospital, Helsinki	V. Sipilä
Hospital of Lounais-Häme, Forssa	T. Kalliomäki, J. Koskelainen, R. Nikkanen, N. Savolainen, H. Sulonen, E. Valtonen
Iisalmi Hospital	E. Toivanen
Jokilaakso Hospital, Jämsä	A. Parta, I. Pirttiniemi

FinnDiane Study Centers	Physicians and nurses
Jorvi Hospital, Helsinki University Central Hospital	S. Aranko, S. Ervasti, R. Kauppinen-Mäkelin, A. Kuusisto, T. Leppälä, K. Nikkilä, L. Pekkonen
Jyväskylä Health Centre, Kyllö	K. Nuorva, M. Tiihonen
Kainuu Central Hospital, Kajaani	S. Jokelainen, P. Kemppainen, A-M. Mankinen, M. Sankari
Kerava Health Centre	H. Stuckey, P. Suominen
Kirkkonummi Health Centre	A. Lappalainen, M. Liimatainen, J. Santaholma
Kivelä Hospital, Helsinki	A. Aimolahti, E. Huovinen
Koskela Hospital, Helsinki	V. Ilkka, M. Lehtimäki
Kotka Heath Centre	E. Pälikkö-Kontinen, A. Vanhanen
Kouvola Health Centre	E. Koskinen, T. Siitonen
Kuopio University Hospital	E. Huttunen, R. Ikäheimo, P. Karhapää, P. Kekäläinen, M. Laakso, T. Lakka, E. Lampainen, L. Moilanen, L. Niskanen, U. Tuovinen, I. Vauhkonen, E. Voutilainen
Kuusamo Health Centre	T. Kääriäinen, E. Isopoussu
Kuusankoski Hospital	E. Kilkki, I. Koskinen, L. Riihelä
Laakso Hospital, Helsinki	T. Meriläinen, P. Poukka, R. Savolainen, N. Uhlenius
Lahti City Hospital	A. Mäkelä, M. Tanner
Lapland Central Hospital, Rovaniemi	L. Hyvärinen, S. Severinkangas, T. Tulokas
Lappeenranta Health Centre	P. Linkola, I. Pulli
Lohja Hospital	T. Granlund, M. Saari, T. Salonen
Loimaa Health Centre	A. Mäkelä, P. Eloranta
Länsi-Uusimaa Hospital, Tammisaari	I-M. Jousmaa, J. Rinne
Malmi Hospital, Helsinki	H. Lanki, S. Moilanen, M. Tilly-Kiesi
Mikkeli Central Hospital	A. Gynther, R. Manninen, P. Nironen, M. Salminen, T. Vänttinen
Mänttä Regional Hospital	I. Pirttiniemi, A-M. Hänninen
North Karelian Hospital, Joensuu	U-M. Henttula, P. Kekäläinen, M. Pietarinen, A. Rissanen, M. Voutilainen
Nurmijärvi Health Centre	A. Burgos, K. Urtamo
Oulankangas Hospital, Oulainen	E. Jokelainen, P-L. Jylkkä, E. Kaarlela, J. Vuolaspuro
Oulu Health Centre	L. Hiltunen, R. Häkkinen, S. Keinänen-Kiukaanniemi
Oulu University Hospital	R. Ikäheimo
Päijät-Häme Central Hospital	H. Haapamäki, A. Helanterä, S. Hämäläinen, V. Ilvesmäki, H. Miettinen
Palokka Health Centre	P. Sopanen, L. Welling
Pieksämäki Hospital	V. Javtsenko, M. Tamminen
Pietarsaari Hospital	M-L. Holmbäck, B. Isomaa, L. Sarelin

FinnDiane Study Centers	Physicians and nurses
Pori City Hospital	P. Ahonen, P. Merensalo, K. Sävelä
Porvoo Hospital	M. Kallio, B. Rask, S. Rämö
Raahe Hospital	A. Holma, M. Honkala, A. Tuomivaara, R. Vainionpää
Rauma Hospital	K. Laine, K. Saarinen, T. Salminen
Riihimäki Hospital	P. Aalto, E. Immonen, L. Juurinen
Salo Hospital	A. Alanko, J. Lapinleimu, P. Rautio, M. Virtanen
Satakunta Central Hospital, Pori	M. Asola, M. Juhola, P. Kunelius, M-L. Lahdenmäki, P. Pääkkönen, M. Rautavirta
Savonlinna Central Hospital	E. Korpi-Hyövälti, T. Latvala, E. Leijala
South Karelia Central Hospital, Lappeenranta	T. Ensala, E. Hussi, R. Härkönen, U. Nyholm, J. Toivanen
Tampere Health Centre	A. Vaden, P. Alarotu, E. Kujansuu, H. Kirkkopelto-Jokinen, M. Helin, S. Gummerus, L. Calonius, T. Niskanen, T. Kaitala, T. Vatanen
Tampere University Hospital	I. Ala-Houhala, T. Kuningas, P. Lampinen, M. Määttä, H. Oksala, T. Oksanen, K. Salonen, H. Tauriainen, S. Tulokas
Tiirismaa Health Centre, Hollola	T. Kivelä, L, Petlin, L. Savolainen
Turku Health Centre	I. Hämäläinen, H. Virtamo, M. Vähätalo
Turku University Central Hospital	K. Breitholz, R. Eskola, K. Metsärinne, U. Pietilä, P. Saarinen, R. Tuominen, S. Äyräpää
Vaajakoski Health Centre	K. Mäkinen, P. Sopanen
Valkeakoski Regional Hospital	S. Ojanen, E. Valtonen, H. Ylönen, M. Rautiainen, T. Immonen
Vammala Regional Hospital	I. Isomäki, R. Kroneld, M. Tapiolinna-Mäkelä
Vaasa Central Hospital	S. Bergkulla, U. Hautamäki, V-A. Myllyniemi, I. Rusk

Table S14: Members of the SUMMIT consortium.

Partner	Name	Position
1	**Michael Mark**	**Coordinator, WP6 leader**
Boehringer-Ingelheim	Markus Albertini	Project manager
Ingelheim, Germany	Carine Boustany	Chronic Kidney Disease, Head of Lab
	Alexander Ehlgen	Transmed
	Martin Gerl	Biomarker & Bioanlysis, Groupleader
	Jochen Huber	In vivo Scientist CMDR, Head of Lab
	Corinna Schölch	Biomarker & Bioanlysis, Head of Lab
	Heike Zimdahl-Gelling	Pharmacogenomics, Head of Lab
2	**Leif Groop**	Prof. Endocrinology; Coordinator Managing entity IMI-JU; PI; **WP1 and WP6 leader**
Lund University	Elisabet Agardh	Prof. Ophthalmology
Clinical Research Centre	Emma Ahlqvist	Postdoc
Malmö, Sweden	Tord Ajanki	Communication strategist
	Nibal Al Maghrabi	Research nurse
	Peter Almgren	Biostatistician
	Jan Apelqvist	Diabetologist
	Eva Bengtsson	Assis. Prof. Cardiovascular research
	Lisa Berglund	Postdoc
	Harry Björckbacka	Assis. Prof. Cardiovascular research
	Ulrika Blom-Nilsson	LUDC administrator
	Mattias Borell	Website, server management
	Agneta Burström	Research nurse
	Corrado Cilio	Assoc. Prof. Cellular autoimmunity
	Magnus Cinthio	Assist. Prof. Electrical Measurements, Lund Technical University
	Karl Dreja	Nephrologist
	Pontus Dunér	Postdoc Exp. Cardiovasc. Research
	Daniel Engelbertsen	PhD student Exp. Cardiovasc. Research
	Joao Fadista	Postdoc
	Maria Gomez	Assoc. Prof. Cardiovascular disease, **WP4 co-leader**
	Isabel Goncalves	Assis. Prof. Cardiovascular research

	Bo Hedblad	Prof. Cardiovascular epidemiology
	Anna Hultgårdh	Prof. Vessel Wall Biology
	Martin E. Johansson	Pathologist
	Cecilia Kennbäck	Laboratory Engineer
	Jasmina Kravic	Database manager
	Claes Ladenvall	Genetic statistician
	Åke Lernmark	Prof. Type 1 diabetes and celiac disease
	Eero Lindholm	Physician, Researcher Diabetic Complications
	Charlotte Ling	Assist. Prof. Epigenetics
	Holger Luthman	Prof. Medical genetics
	Olle Melander	Assoc. Prof. Hypertension and cardiovascular disease
	Malin Neptin	Biomedical analyst
	Jan Nilsson	Prof. Experimental Cardiovascular research, **WP3 leader**
	Peter Nilsson	Prof. Internal medicine
	Tobias Nilsson	PhD student Electrical Measurements, Lund Technical University
	Gunilla Nordin Fredriksson	Prof. Cardiovascular research
	Marju Orho-Melander	Prof. Genetic epidemiology
	Emilia Ottoson-Laakso	PhD student
	Annie Persson	Research nurse
	Margaretha Persson	Laboratory Engineer
	Mats-Åke Persson	Database manager
	Jacqueline Postma	Project manager
	Elisabeth Pranter	Research nurse
	Sara Rattik	PhD student Exp. Cardiovasc. Research
	Gunnar Sterner	Chief physician Internal Medicine Research Unit
	Lilian Tindberg	Research nurse
	Maria Wigren	Postdoc Exp. Cardiovasc. Research
	Anna Zetterqvist	PhD student
	Mikael Åkerlund	Postdoc
	Gerd Östling	Laboratory Engineer
3	**Timo Kanninen**	Technical director; PI
Biocomputing Platforms	Anni Ahonen-Bishopp	Software development manager

(BC Platforms)	Anita Eliasson	Financial and administrative director
Espoo, Finland	Timo Herrala	System (server) specialist
	Päivi Tikka-Kleemola	Service manager
4	**Anders Hamsten**	Prof. Cardiovascular disease; Atherosclerosis Research Unit; PI
Karolinska Institute	Christer Betsholtz	Prof. Vascular biology
Stockholm, Sweden	Ami Björkholm	Administrator
	Ulf de Faire	Professor emeritus Cardiovascular epidemiology
	Fariba Foroogh	Research engineer
	Guillem Genové	Scientist
	Karl Gertow	Research Assist. Prof. Cardiovascular genetics
	Bruna Gigante	Assoc. Professor Cardiovascular epidemiology
	Bing He	Postdoc
	Karin Leander	Assoc. Professor Cardiovascular epidemiology
	Olga McLeod	Postdoc
	Maria Nastase-Mannila	Postdoc
	Jaako Patrakka	Postdoc
	Angela Silveira	Assoc. Prof. Cardiovascular genetics
	Rona Strawbridge	Postdoc
	Karl Tryggvason	Prof. Medical Chemistry
	Max Vikström	Statistician
	John Öhrvik	Professor
	Anne-May Österholm	Postdoc
5	**Barbara Thorand**	Nutritional scientist, epidemiologist
Helmholtz Centre	Christian Gieger	Statistician
Munich, Germany	Harald Grallert	Biologist
	Tonia Ludwig	Statistician
	Barbara Nitz	Scientist
	Andrea Schneider	Data manager
	Rui Wang-Sattler	Scientist
	Astrid Zierer	Statistician

6	**Giuseppe Remuzzi**	Institute director; PI
Mario Negri Institute for	Ariela Benigni	Head of department Molecular Medicine
Pharmacological Research	Roberta Donadelli	Scientist
	Maria Domenica Lesti	Researcher
Bergamo, Italy	Marina Noris	Head Laboratory Immunology and genetics of transplantation and rare diseases
	Norberto Perico	Senior scientist
	Annalisa Perna	Biostatistician
	Rossella Piras	Postdoc
	Piero Ruggenenti	Head of department Renal medicine, Assist. Prof. Nephrology and dialysis
	Erica Rurali	Postdoc
7	**David Dunger (att: Jane Horsford)**	Prof. Paediatrics; PI
University of Cambridge	Ludo Chassin	Senior Data Manager
UK	Neil Dalton, London	Clinical biochemistry
	John Deanfield, London	Paediatric cardiology
	Jane Horsford	PA to Prof. Dunger
	Clare Rice	Operations manager/financial contact
	James Rudd	Cardiovascular imaging
	Neil Walker	Head Data services
	Karen Whitehead	Technician
	Max Wong	Postdoc
8	**Helen Colhoun**	Prof. Public health and epidemiology; PI; Vice coordinator Managing entity; **WP2 leader**
	Fiona Adams	
University of Dundee	Tahira Akbar	PA to Helen Colhoun
Scotland	Jill Belch	Prof. Vasucular disease
	Harshal Deshmukh	PhD student
	Fiona Dove	
	Angela Ellingford	NHS Tayside Diabetic Retinopathy Screening Programme manager
	Bassam Farran	Statistician
	Mike Ferguson	Dean of research Biological chemistry and drug discovery
	Gary Henderson	

	Graeme Houston	Consultant radiologist/senior lecturer
	Faisel Khan	Reader, Vascular & Inflammatory Diseases Research Unit
	Graham Leese	Consultant diabetologist/reader
	Yiyuan Liu	PhD student
	Shona Livingstone	Senior statistician
	Helen Looker	Epidemiologist
	Margaret McCann	Project assistant
	Stuart McGurnaghan	Lead data programmer
	Andrew Morris	Prof. Diabetic medicine
	David Newton	
	Colin Palmer	Prof. Pharmacogenomics
	Ewan Pearson	Consultant diabetologist/senior lecturer
	Gillian Reekie	Research Nurse
	Natalie Smith	Research Nurse
9	**Angela Shore**	Prof. Cardiovascular Science, PI
Peninsula Medical School	Kuni Aizawa	Postdoc
Exeter, UK	Claire Ball	Research nurse
	Nick Bellenger	Cardiologist
	Francesco Casanova	Associate Research Fellow Vascular medicine
	Tim Frayling	Prof. Genetics
	Phil Gates	Senior lecturer Cardiovascular science
	Kim Gooding	Postdoc Vascular medicine
	Andrew Hatttersley	Prof. Molecular medicine
	Roland Ling	Consultant opthalmologist
	David Mawson	Research technician
	Robin Shandas	Prof. Bioengineering (Colorado)
	David Strain	Stroke physician, clinical lecturer
	Clare Thorn	Postdoc Vascular medicine
10	**Ulf Smith**	Prof. ; PI
University of Gothenburg	Ann Hammarstedt	Researcher Molecular and clinical medicine
Sweden	Hans Häring	Prof. University of Tübingen

	Oluf Pedersen	Prof. Steno Centre, Copenhagen
	Georgio Sesti	Prof. Universtiy of Catanzaro
11	**Per-Henrik Groop**	Prof. Diabetes genetics; PI
	Emma Fagerholm	MSc; PhD student, genetics
Folkhälsan	Carol Forsblom	Clinical coordinator
Helsinki, Finland	Valma Harjutsalo	PhD; FinnDiane Co-PI
	Maikki Parkkonen	Laboratory manager
	Niina Sandholm	DSc(PhD); GWAS and bioinformatics, FinnDiane Co-PI
	Nina Tolonen	MD PhD
	Iiro Toppila	BSc, MSc; bioinformatician
	Erkka Valo	MSc; PhD student, bioinformatician
12	**Veikko Salomaa**	Prof. Epidemiology; PI; **deputy leader WP2**
The National Institute for Health and Welfare	Aki Havulinna	DSc. (tech), statistician
Helsinki, Finland	Kati Kristiansson	PhD
	Pia Okamo	THL press officer
	Tomi Peltola	PhD
	Markus Perola	Professor
	Arto Pietilä	Statistician
	Samuli Ripatti	Professor, Statistics
	Marketta Taimi	Research assistant
13	**Seppo Ylä-Herttuala**	Prof.; PI; **WP4 leader**
University of Eastern Finland	Mohan Babu	PhD student
Kuopio, Finland	Marike Dijkstra	PhD student
	Erika Gurzeler	PhD student
	Jenni Huusko	PhD student
	Ivana Kholová	Postdoc
	Markku Laakso	Prof.
	Mari Merentie	PhD student

	Marja Poikolainen	PA Prof Ylä-Herttuala
14	**Mark McCarthy**	Prof. Human type 2 diabetes; Oxford Centre for Diabetes, Endocrinology and Metabolism; Wellcome Trust Centre for Human Genetics; PI; **deputy leader WP1**
University of Oxford	Will Rayner	Database manager
UK	Neil Robertson	Informatics
	Natalie van Zuydam	Postdoc
15	**Claudio Cobelli**	Prof. ; PI; **WP5 leader**
University of Padova	Barbara Di Camillo	Assist. Prof.
Italy	Francesca Finotello	PhD student
	Francesco Sambo	Postdoctoral fellow
	Gianna Toffolo	Prof.
	Emanuele Trifoglio	PhD student
16	**Riccardo Bellazzi**	Prof. Bioengineering; PI; **deputy leader WP5**
	Nicola Barbarini	Postdoctoral fellow
University of Pavia	Mauro Bucalo	Software engineer
Italy	Christiana Larizza	Assist. Prof.
	Paolo Magni	Assoc. Prof.
	Alberto Malovini	Postdoctoral fellow
	Simone Marini	Postdoctoral fellow
	Francesca Mulas	Postdoctoral fellow
	Silvana Quaglini	Prof.
	Lucia Sacchi	Assist. Prof.
	Francesca Vitali	
17	**Ele Ferrannini**	Prof. Medicine; PI
	Beatrice Boldrini	Postdoctoral fellow
University of Pisa	Michaela Kozakova	Senior investigator Medical Pathophysiology
Italy	Andrea Mari	Senior researcher Biomedical engineering (ISIB-CNR, Padova)
	Carmela Morizzo	Biologist, Sonographer Cardiovascular ultrasound
	Lucrecia Mota	EGIR administrative office

	Andrea Natali	Assoc. Prof. Medicine
	Carlo Palombo	Assoc. Prof. Medicine; **deputy leader WP3**
	Elena Venturi	Researcher
	Mark Walker	Prof. Molecular diabetic medicine (Univ Newcastle-upon-Tyne)
18	**Carlo Patrono**	Prof.Pharmacology; PI
Catholic University of Rome	Francesca Pagliaccia	PhD student
Italy	Bianca Rocca	Assist. Prof. Pharmacology
19	**Pirjo Nuutila**	Prof. ; PI
University of Turku	Johanna Haukkala	PhD student
Finland	Juhani Knuuti	Prof. ; Director Turku PET Centre
	Anne Roivainen	Prof.
	Antti Saraste	Adj. Prof.
20	**Paul McKeague**	Prof. Genetic Epidemiology; PI
University of Edinburgh	Norma Brown	Research administrator, Public Health Services
Scotland	Marco Colombo	Bioinformaticist
21	**Birgit Steckel-Hamann**	Deputy coordinator; PI, Manager IMI, LRL
Eli Lilly	Krister Bokvist	Biostatistician
	Sudha Shankar	Diabetologist
	Melissa Thomas	Translational Science
22	**Li-ming Gan**	Prof.; Translational Science Director Cardiovascular Disease; PI, **WP3 leader**
AstraZeneca	Suvi Heinonen	PhD, Internal AZ postdoc, Bioscience
	Ann-Cathrine Jönsson-Rylander	PhD, Assoc. Prof., Team Leader Bioscience, **WP4 leader**
	Remi Momo	Postdoctoral fellow
	Volker Schnecke	Informatician Translational Science, **WP5 leader**
	Robert Unwin	Translational Science Director Diabetic Nephropathy
	Anna Walentinsson	Geneticist Translational Science
	Carl Whatling	Bioscientist

23	**Everson Nogoceke**	Pre-clinical and clinical aspects of metabolic and vascular disease; PI; **WP2 leader**
Roche	Gonzalo Durán Pacheco	Senior Research Statistician
	Ivan Formentini	Biomarker & Experimental Medicine Leader
	Thomas Schindler	Pre-clinical and clinical and clinical biomarkers
24	**Piero Tortoli**	Professor of Electronics
University of Florence	Luca Bassi	Postdoctoral fellow
	Enrico Boni	Postdoctoral fellow
	Alessandro Dallai	Postdoctoral fellow
	Francesco Guidi	Technician
	Matteo Lenge	PhD student
	Riccardo Matera	PhD student
	Alessandro Ramalli	PhD student
	Stefano Ricci	Assist. Prof.
	Jacopo Viti	PhD student
25	**Bernd Jablonka**	SAD internal IMI coordinator
Sanofi-aventis	Dan Crowther	Biomarker researcher
	Johan Gassenhuber	Biostatistician
	Sibylle Hess	Biomarker researcher
	Thomas Hübschle	Pharmacologist Diabetes
	Hans-Paul Juretschke	Imaging
	Hartmut Rütten	Head Translational Medicine
	Thorsten Sadowski	Pharmacologist Diabetes
	Paulus Wohlfart	Pharmacologist Diabetes
26	**Julia Brosnan**	Biochemist, (pre)clinical research CVD, Pfizer US; **WP2 leader**
Pfizer	Valerie Clerin	Cardio-renal biologist, WP2
	Eric Fauman	Computational biologist
	Craig Hyde	Statistician
	Anders Malarstig	Human genetics, Pfizer Europé; **WP1 leader**
	Nick Pullen	Renal Disease Research Director

	Mera Tilley	
	Theresa Tuthill	Imaging specialist
	Ciara Vangjeli	Cardiovascular genetic epidemiologist, Pfizer Europe
	Daniel Ziemek	Computational biologist

Figure S1. Manhattan and QQ Plots for each case-control definition and covariate model (minimal and full)

DN - minimal

$\lambda = 1.011$

DN - full

λ = 1.009

macro - min

QQ Plot: JDRF_DNCRI_macro_min_mac10_study2_meta_1KG_Dec20171

λ = 1.010

Observed -log10(p-value)

Expected -log10(p-value)
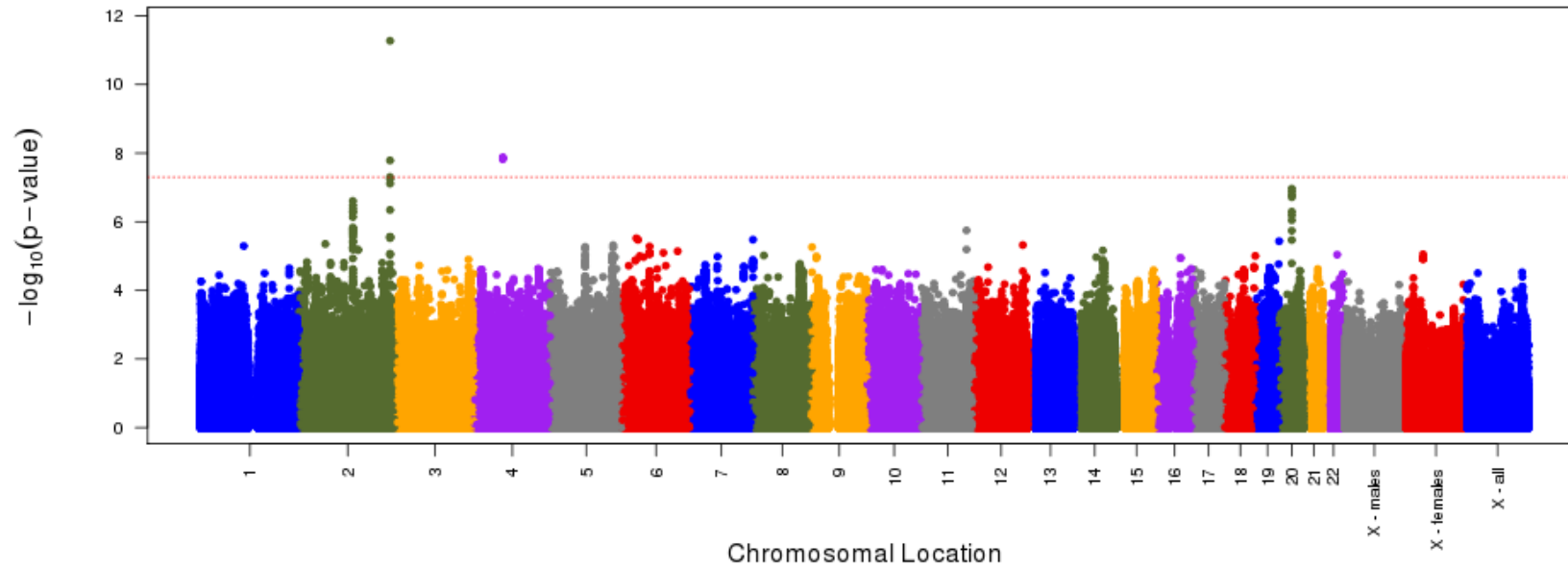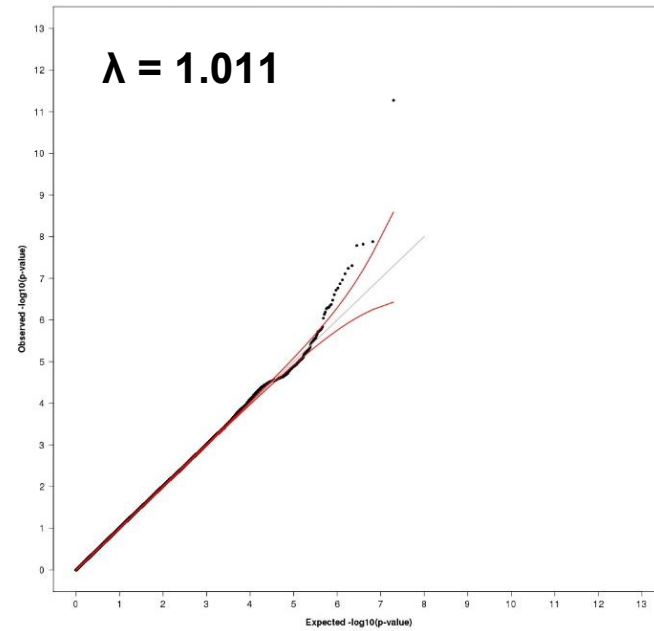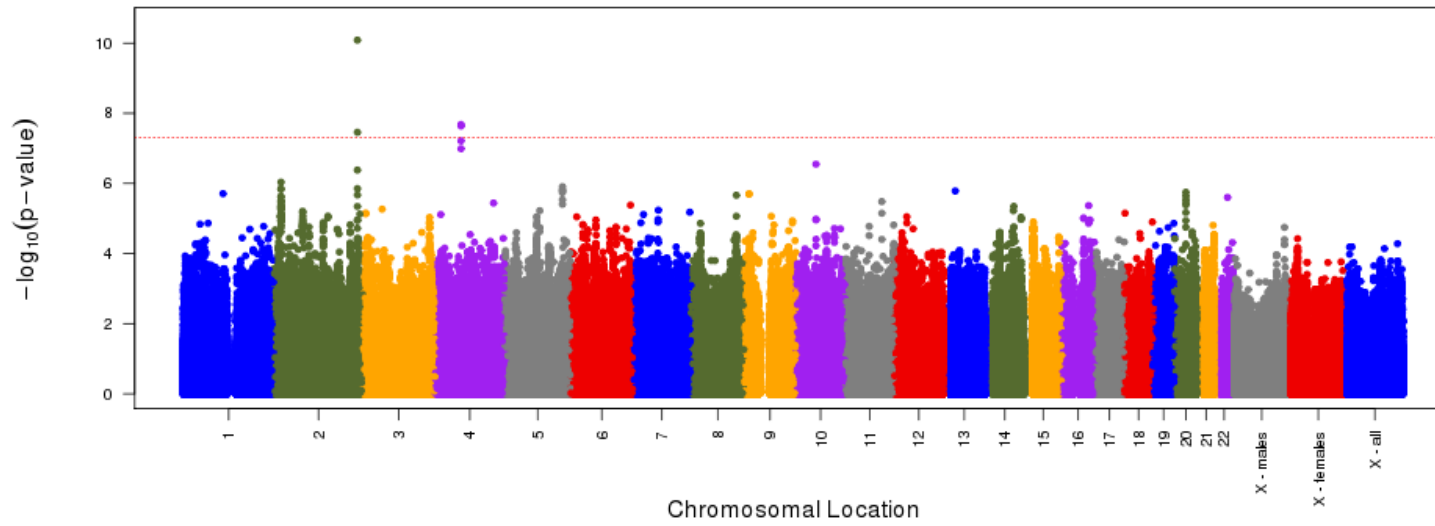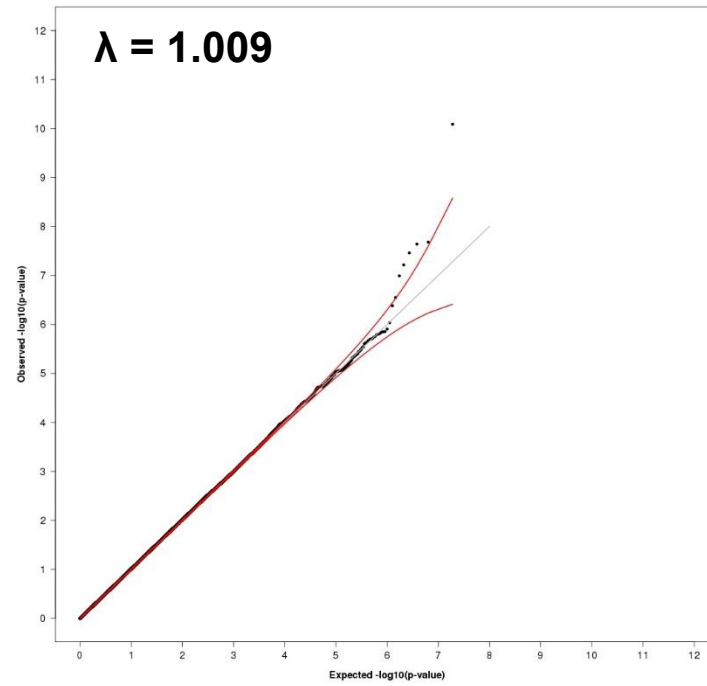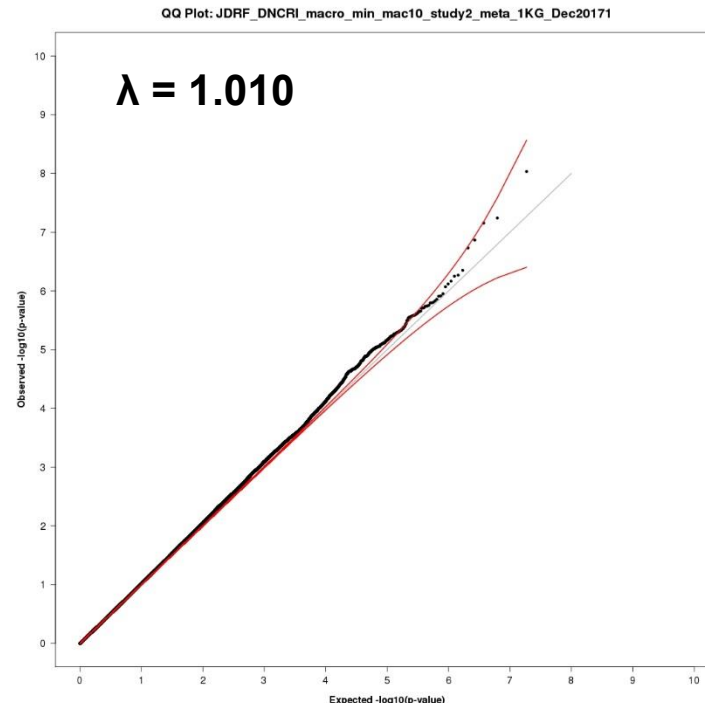
Manhattan Plot - PHENO2_macro
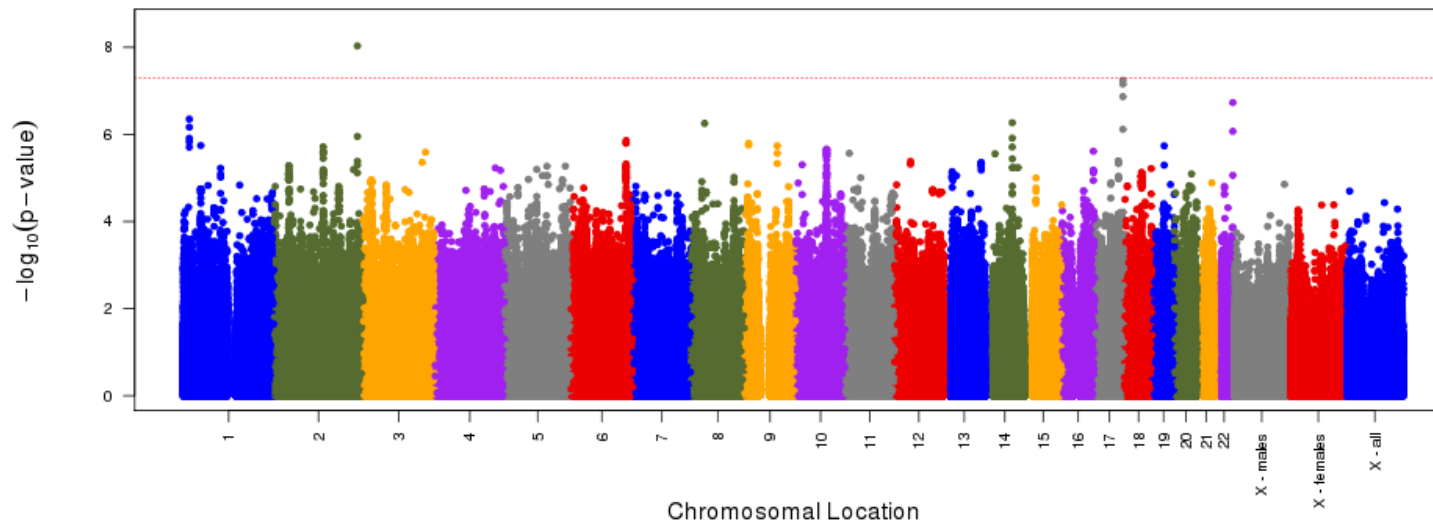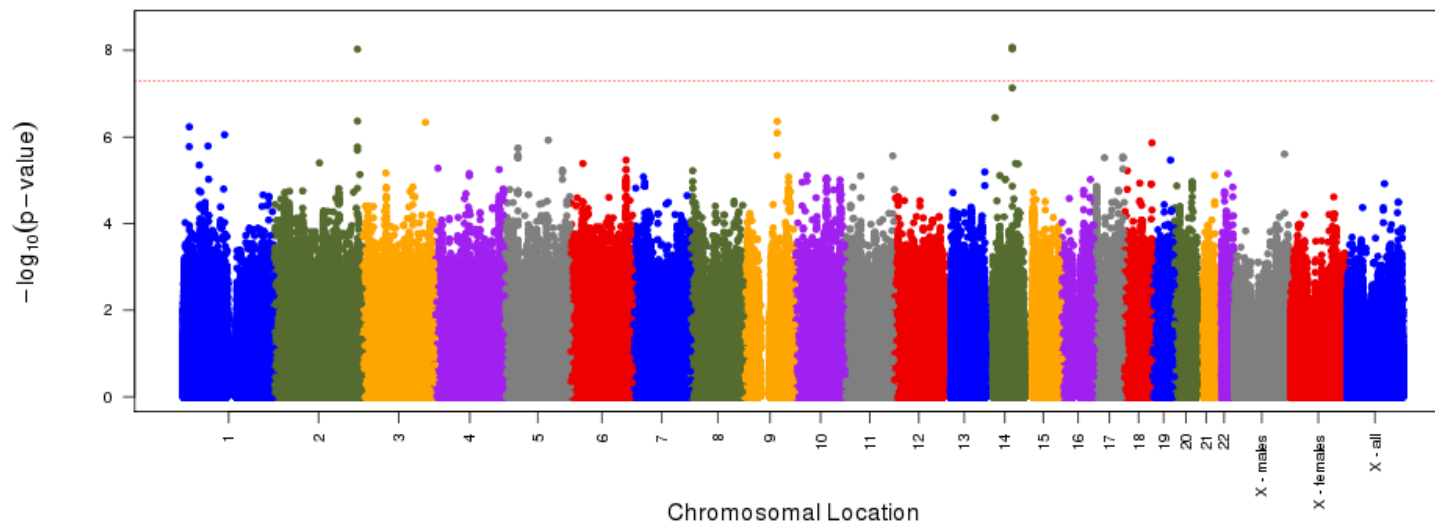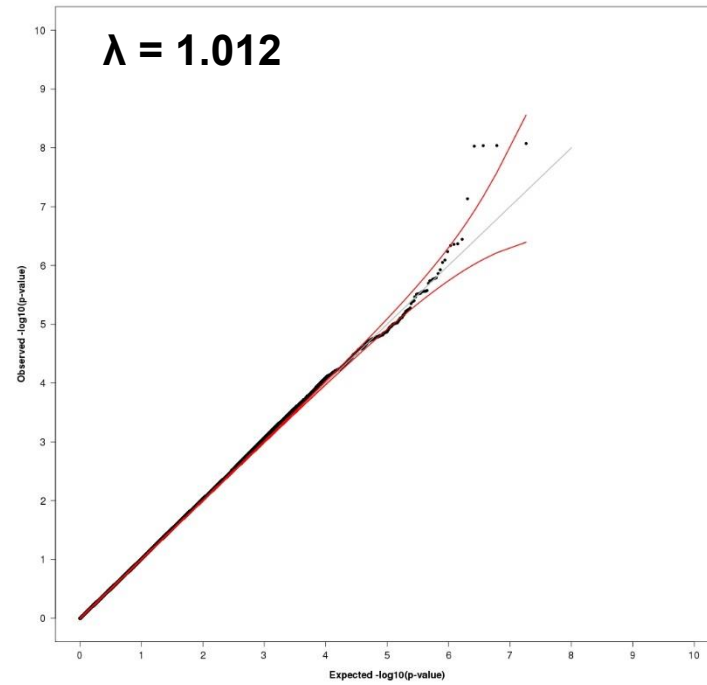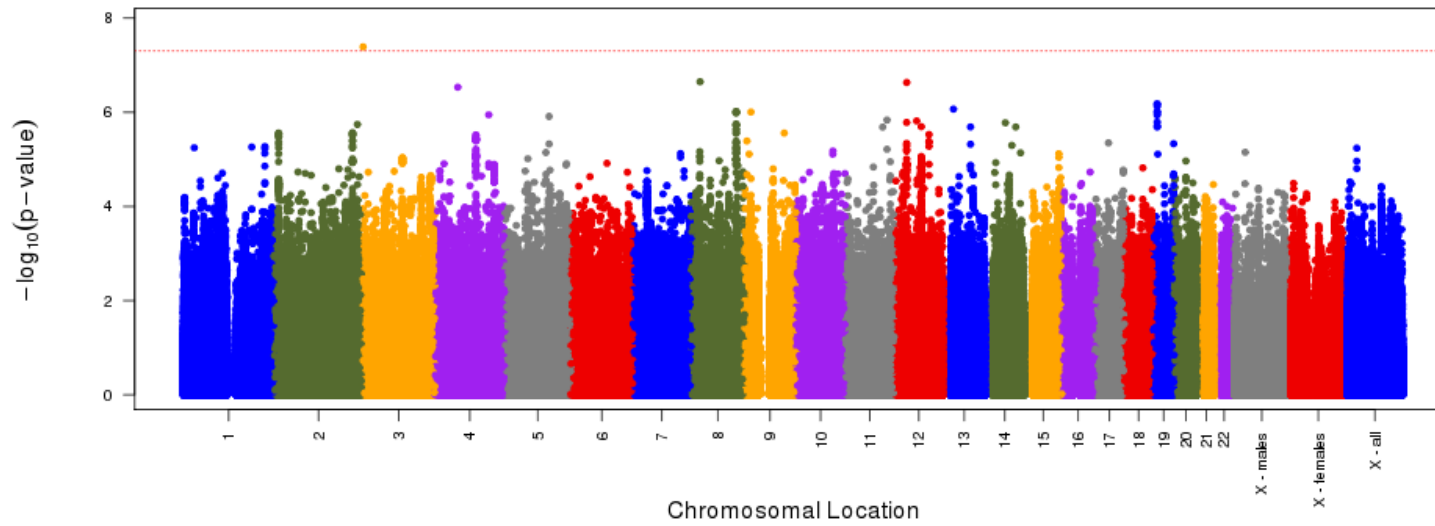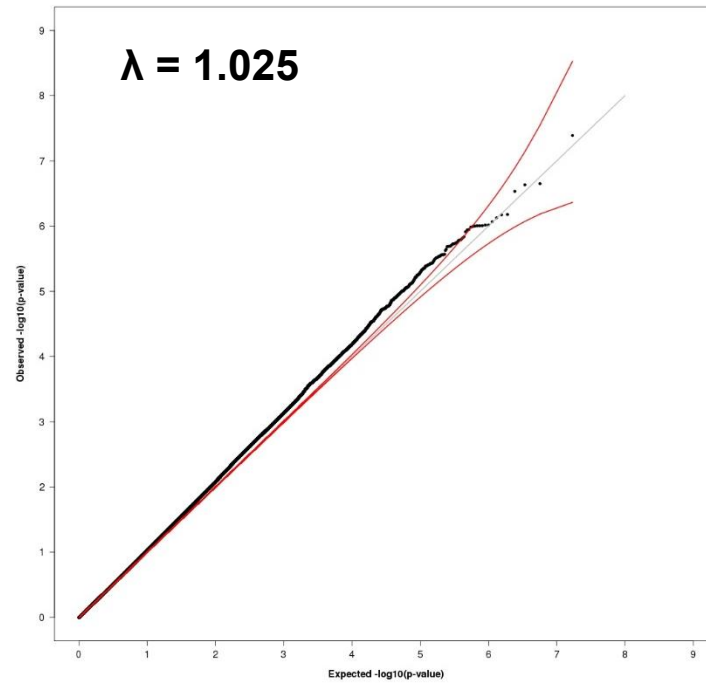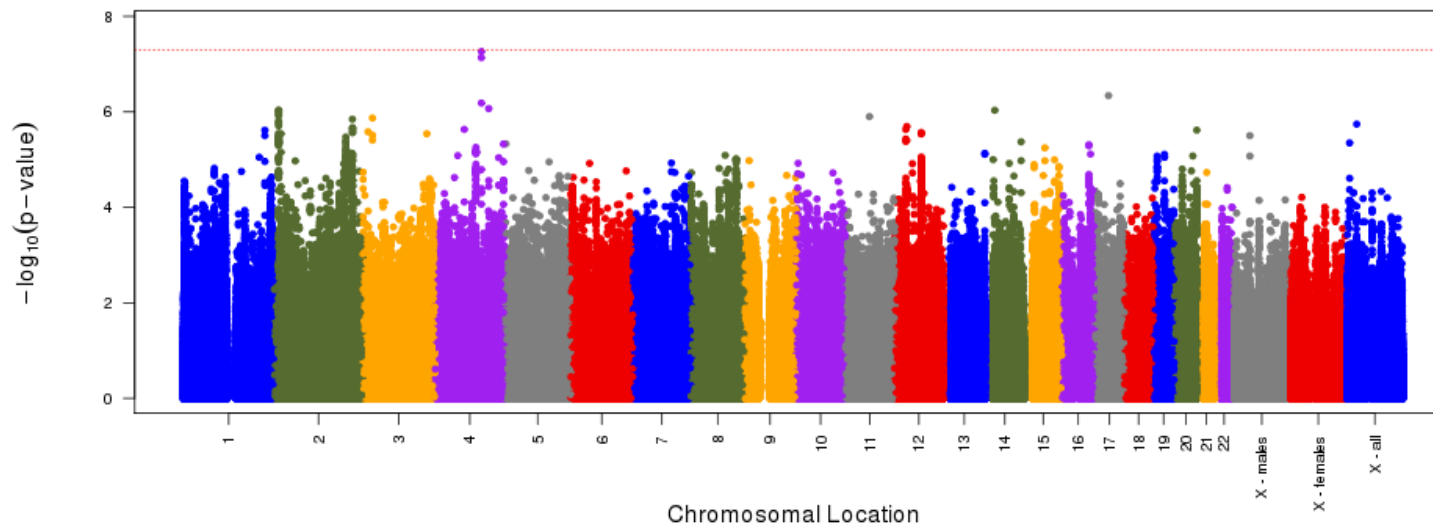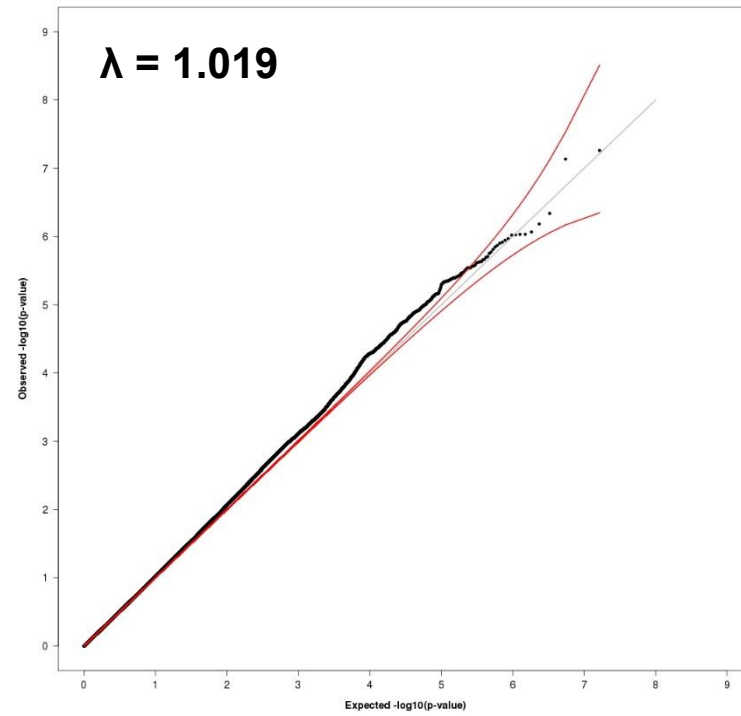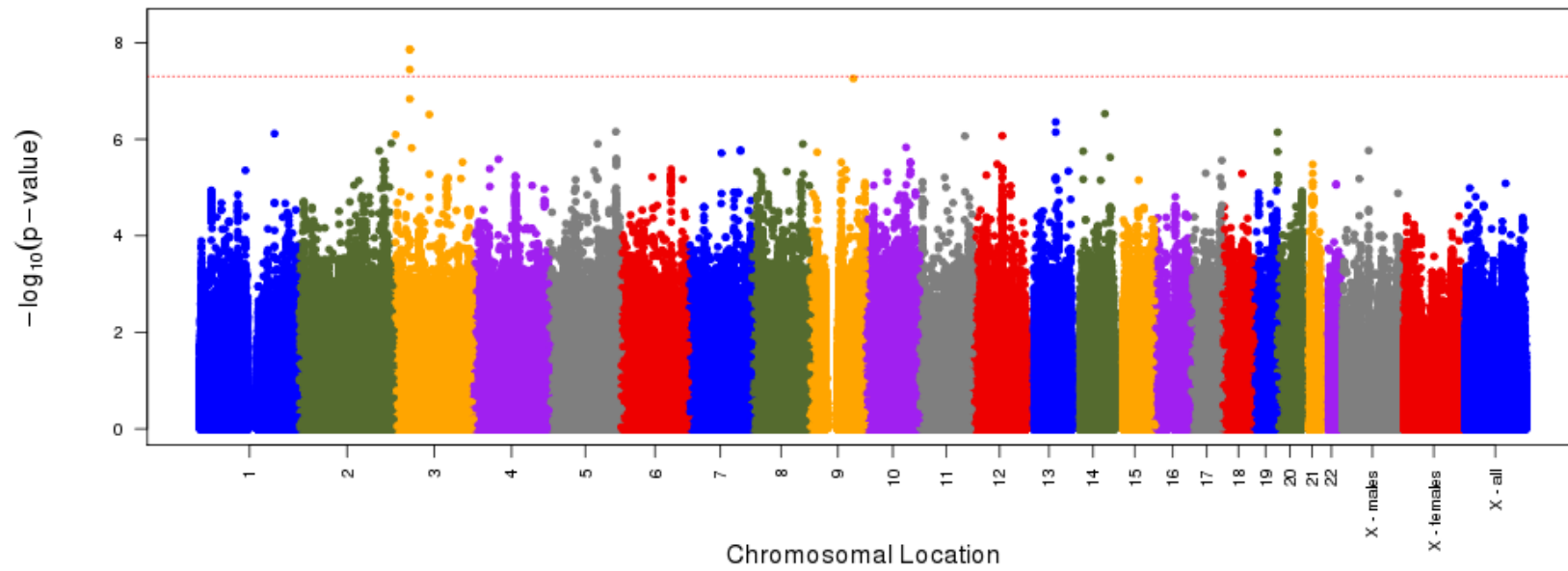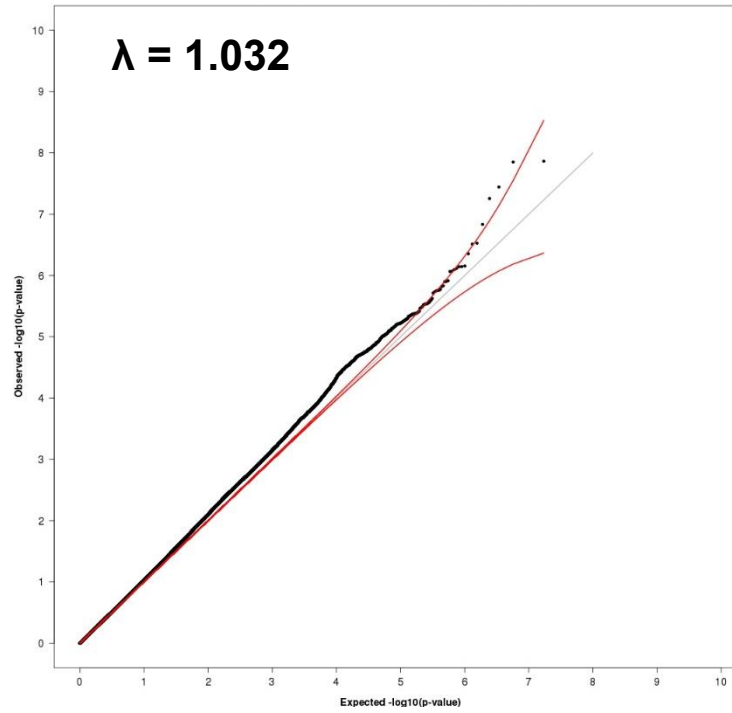
$-\log_{10}(p-value)$

Chromosomal Location

macro - full   **λ = 1.012**

ESRD vs. ctrl- min

λ = 1.025

ESRD vs. ctrl - full   λ = 1.019

ESRD vs. non-ESRD -   λ = 1.032

ESRD vs. non-ESRD -

λ = 1.021

ESRD vs. macro - min   λ = 1.001

ESRD vs. macro - full   λ = 1.014

All vs. ctrl - min

λ = 1.019

All vs. ctrl - full   λ = 1.013

Micro - min

λ = 1.017

Micro - full

λ = 1.010

CKD - min   λ = 1.019

CKD - full

λ = 1.017

CKD extreme - min

λ = 1.027

CKD extreme - full λ = 1.023

CKD-DN - min

λ = 1.018

CKD-DN - full **λ = 1.017**
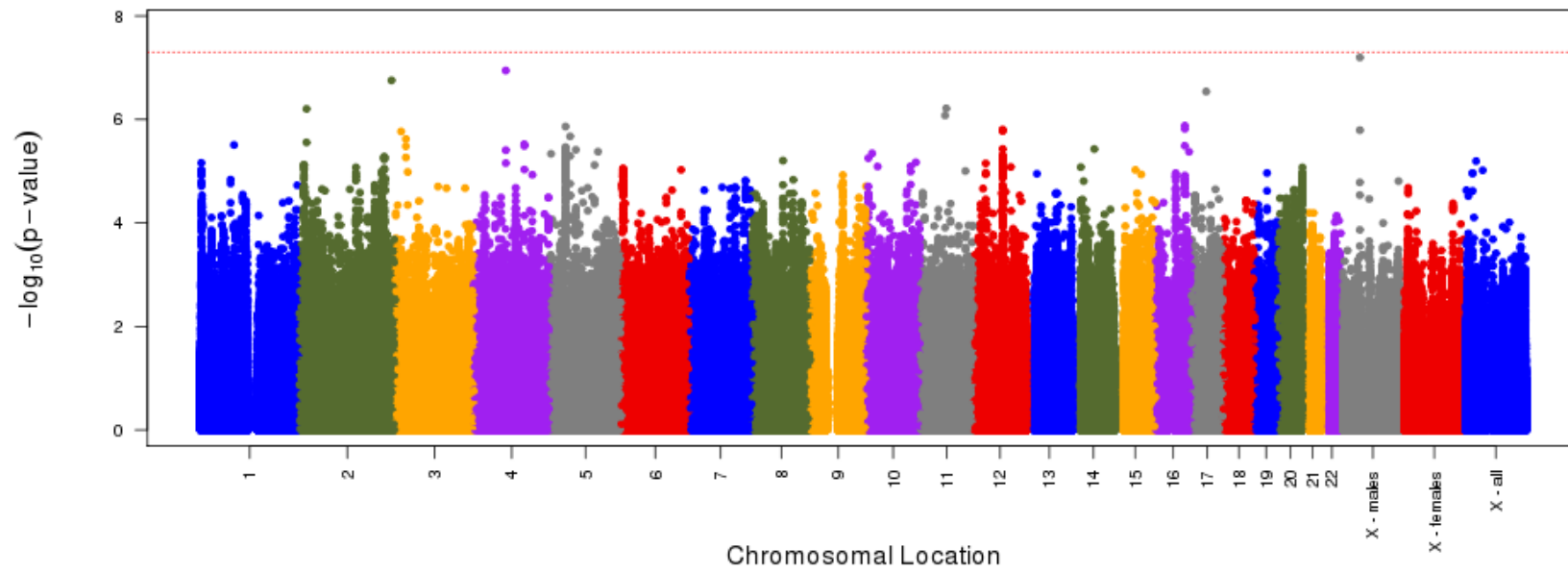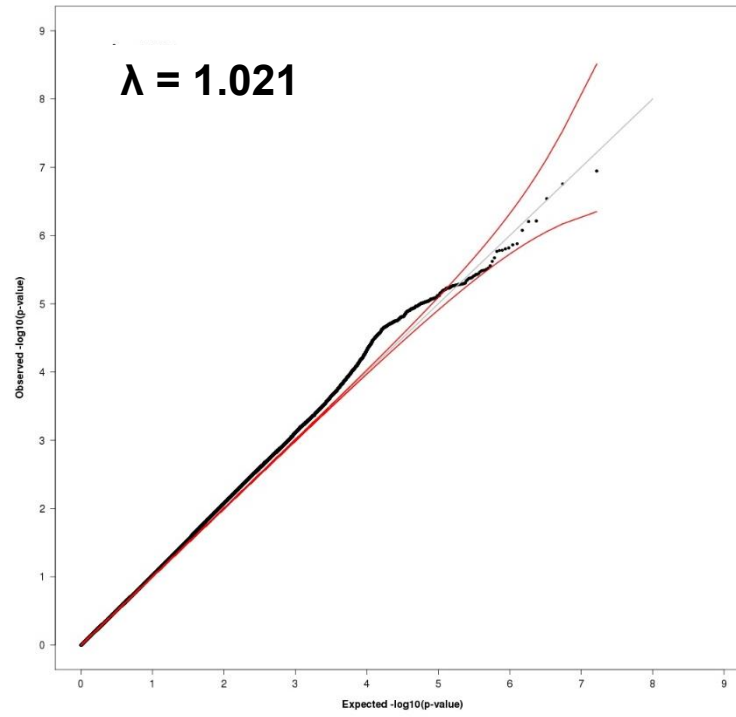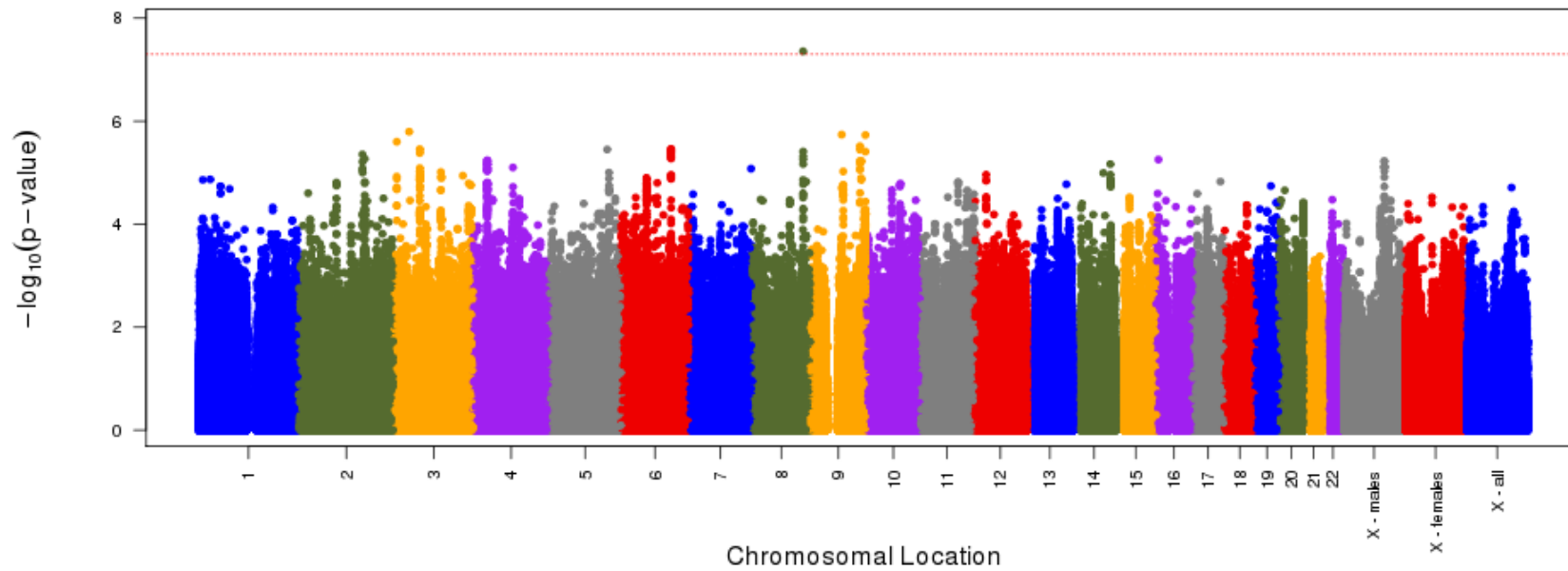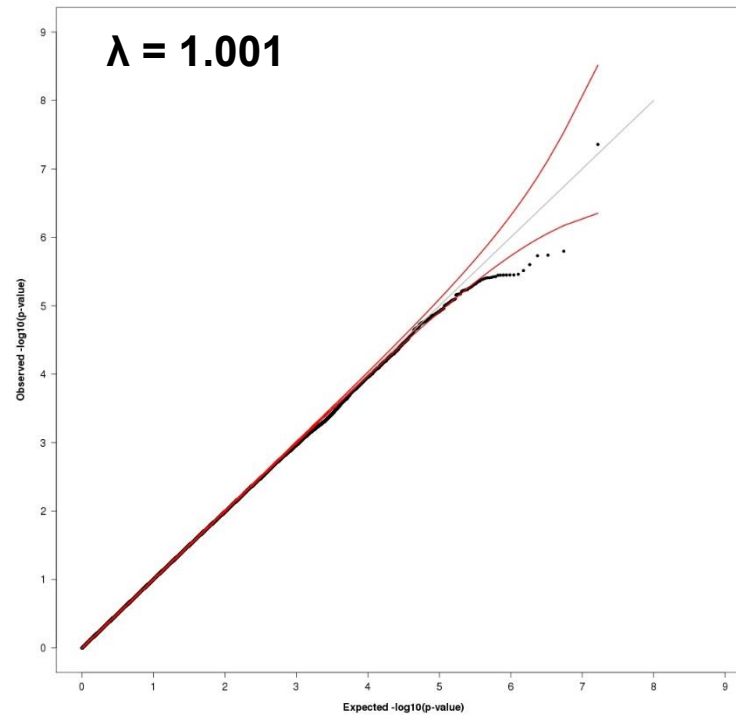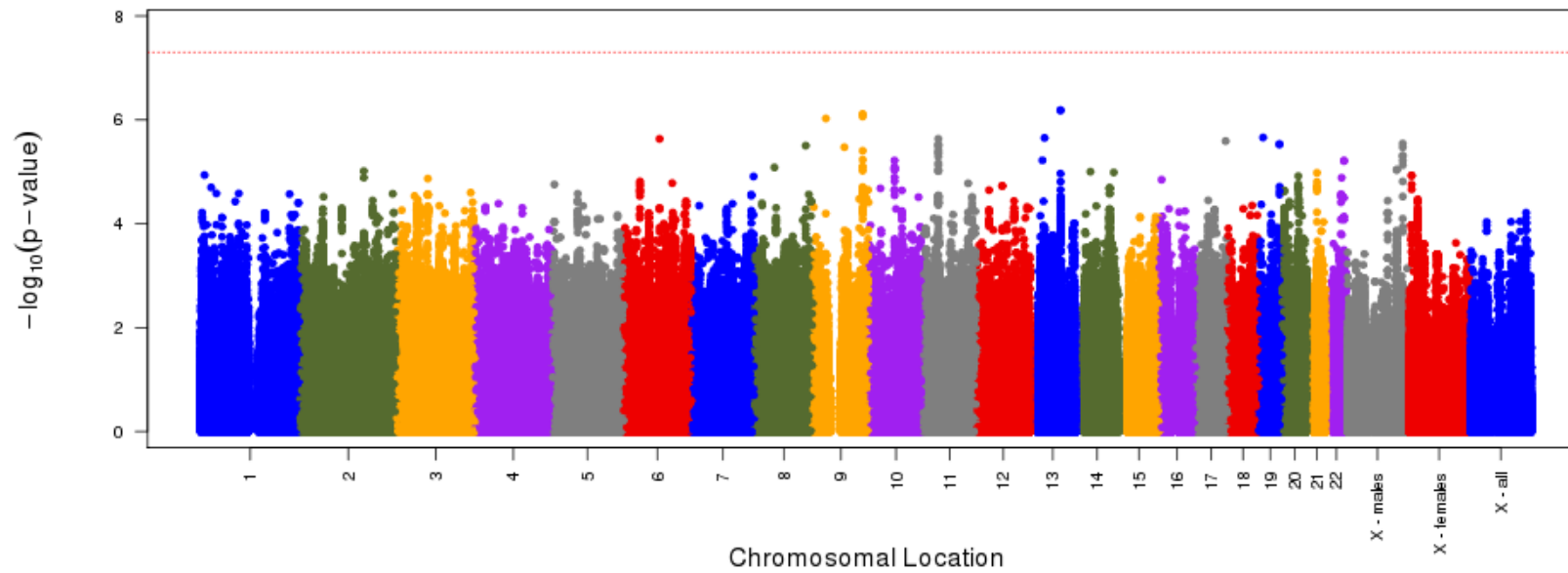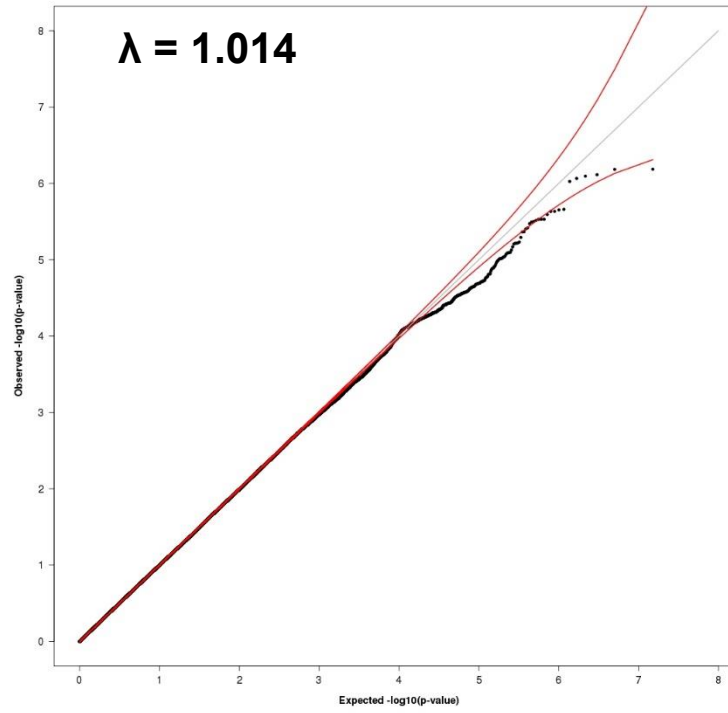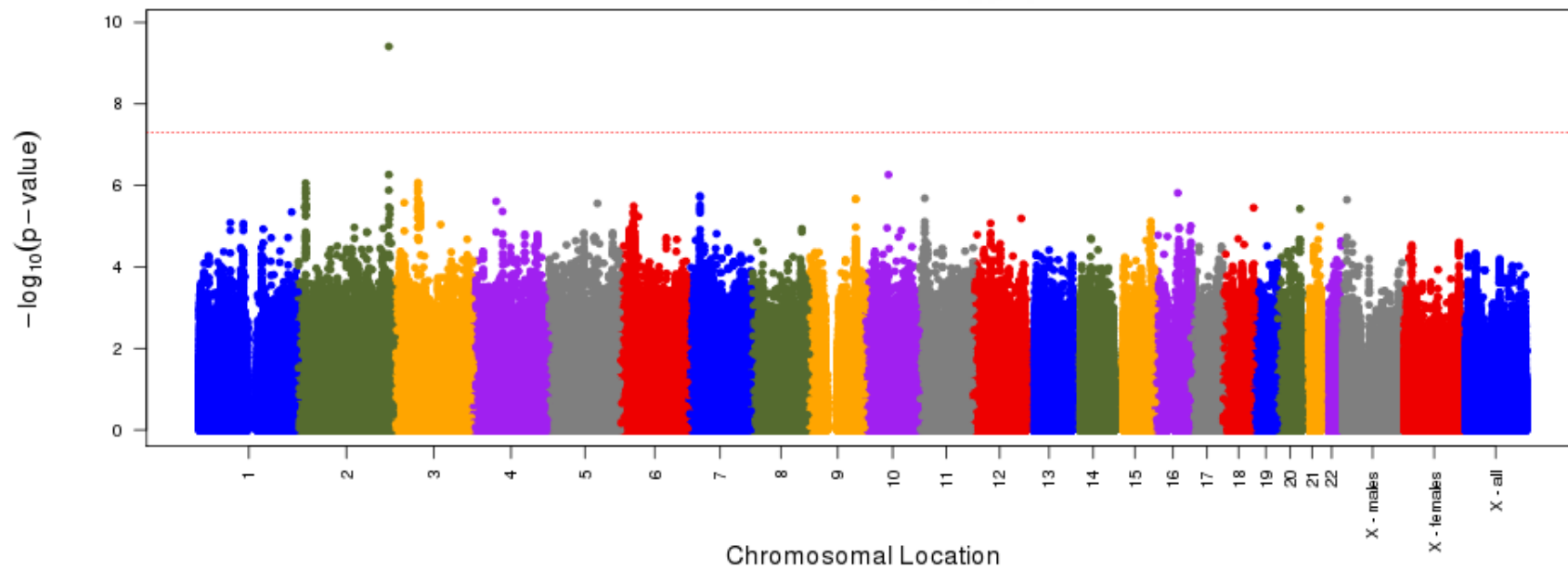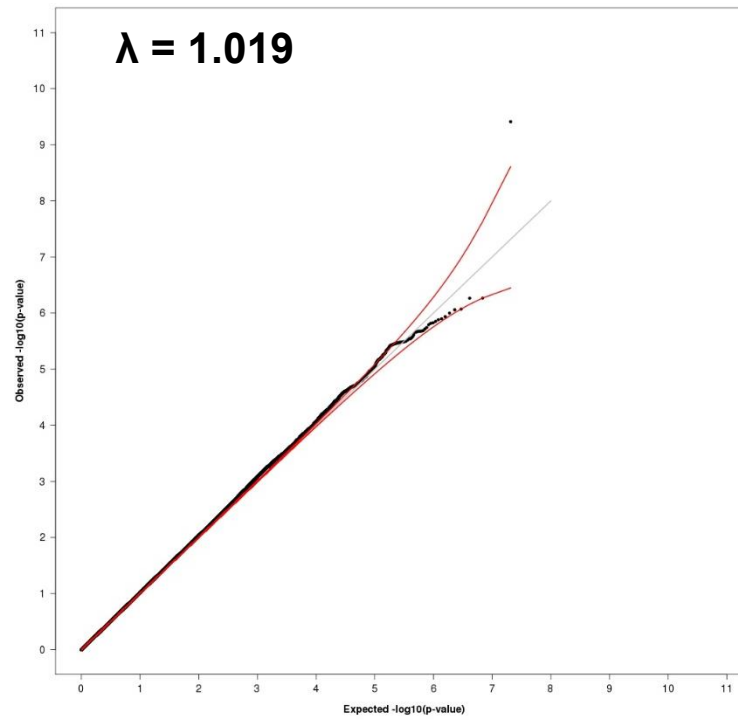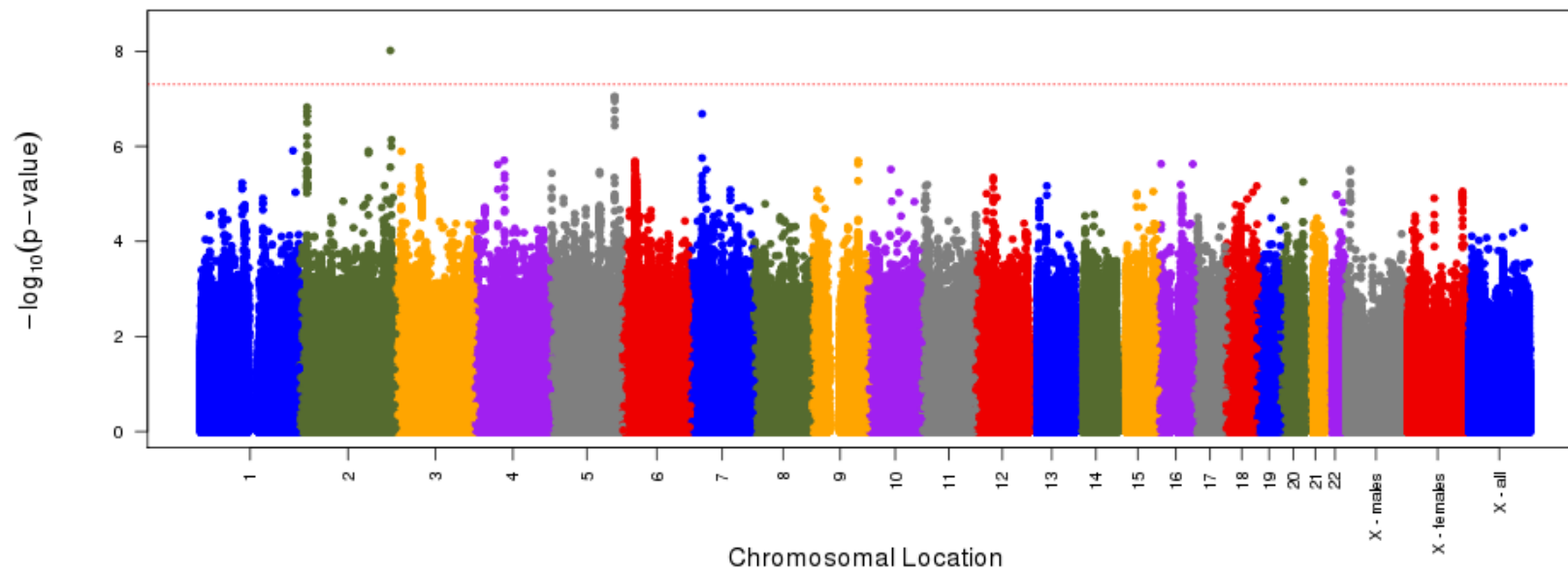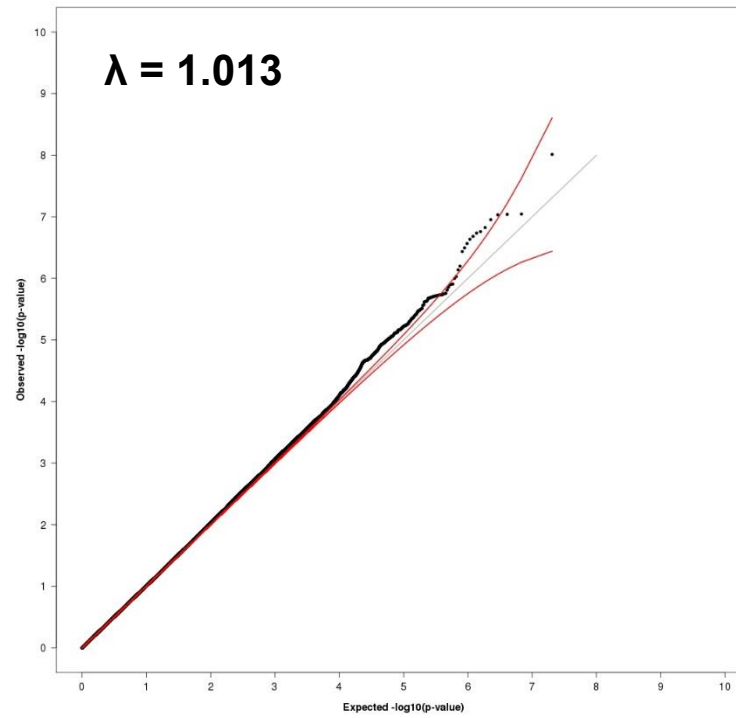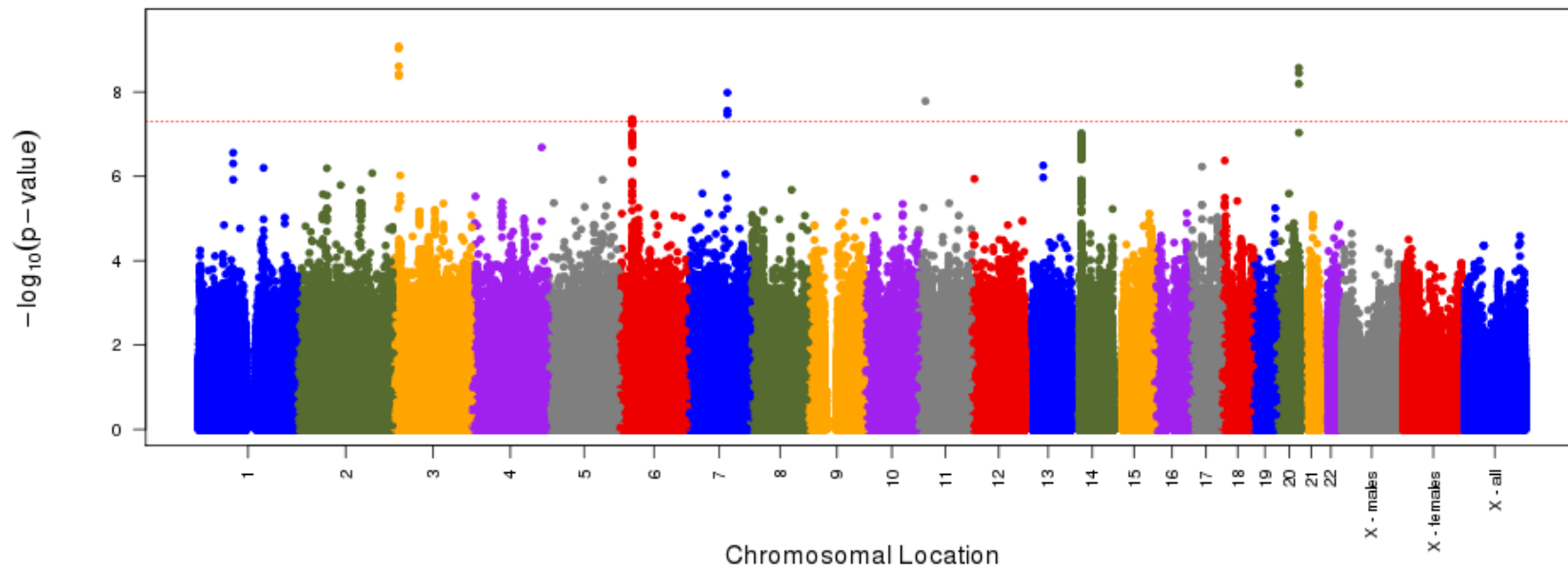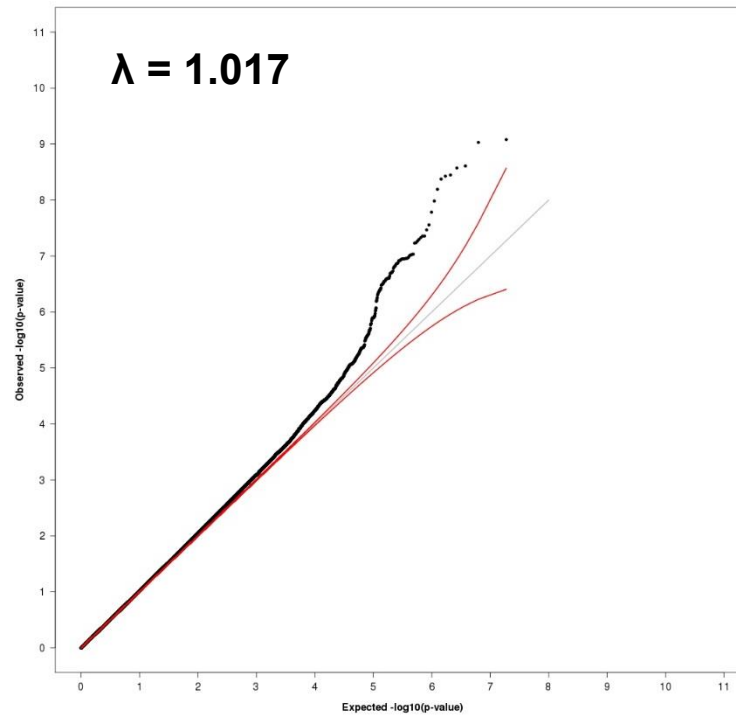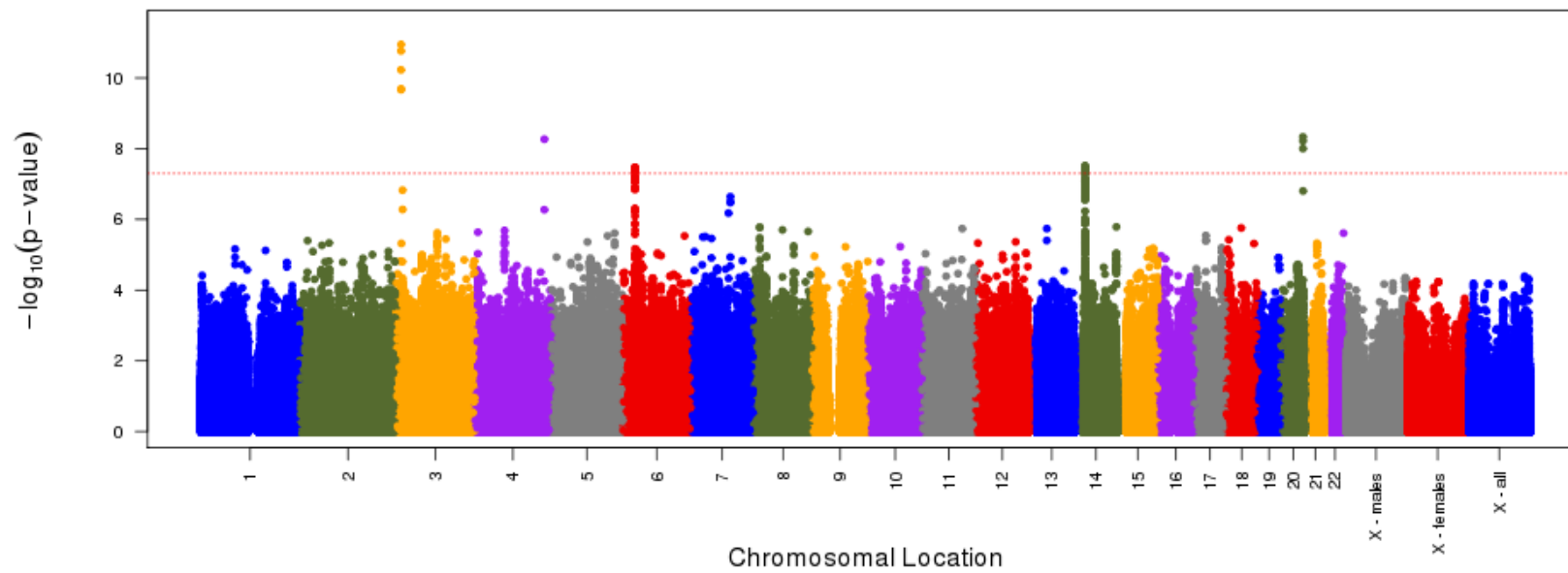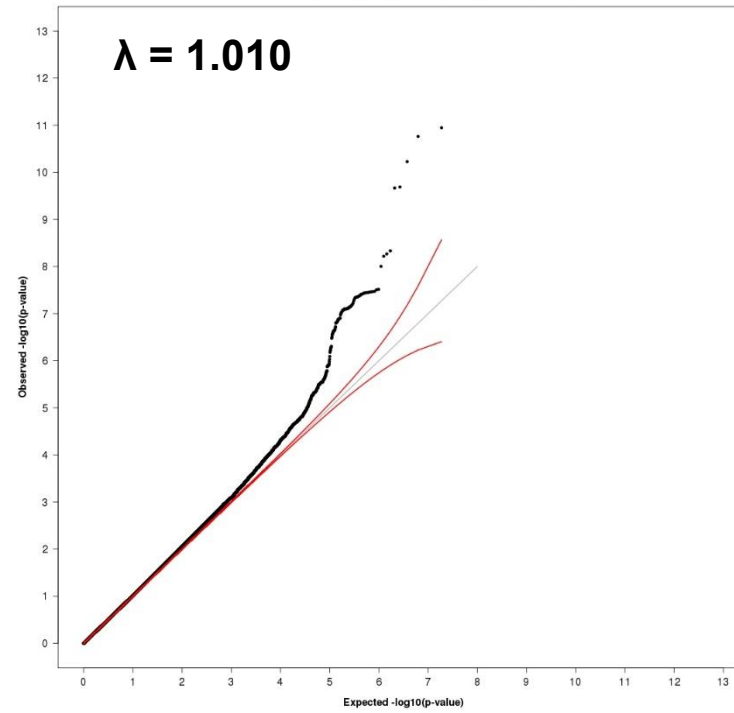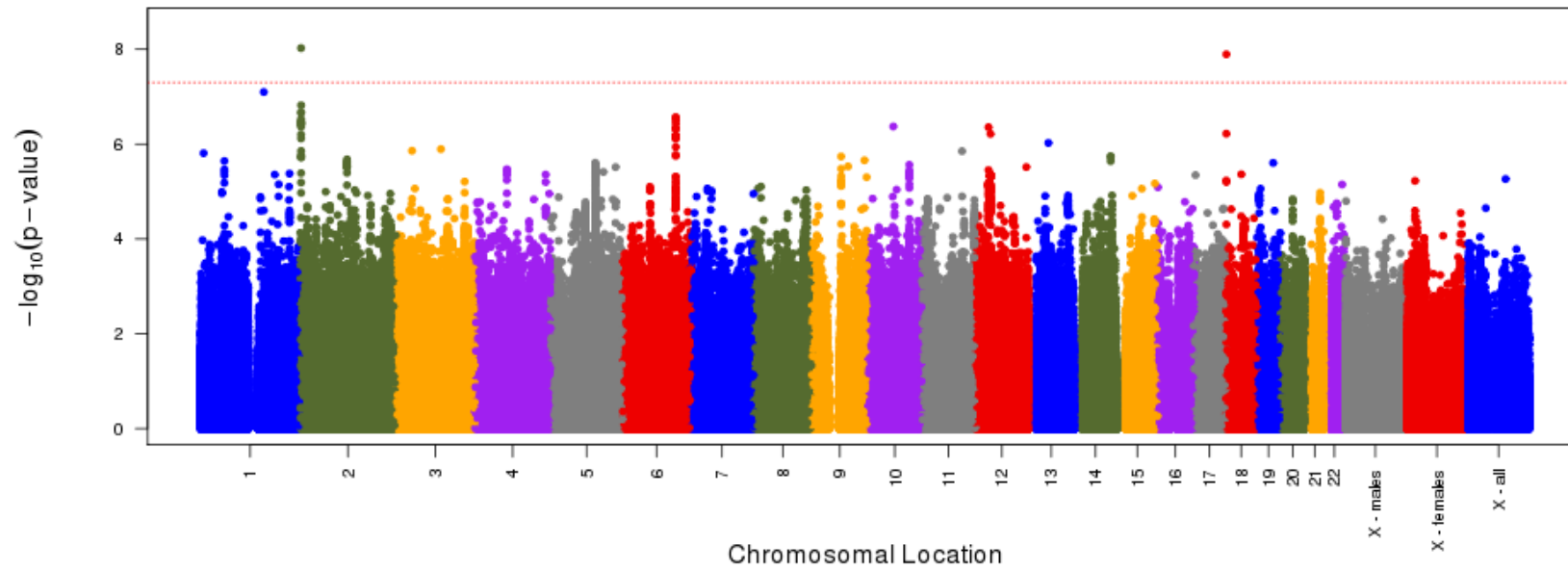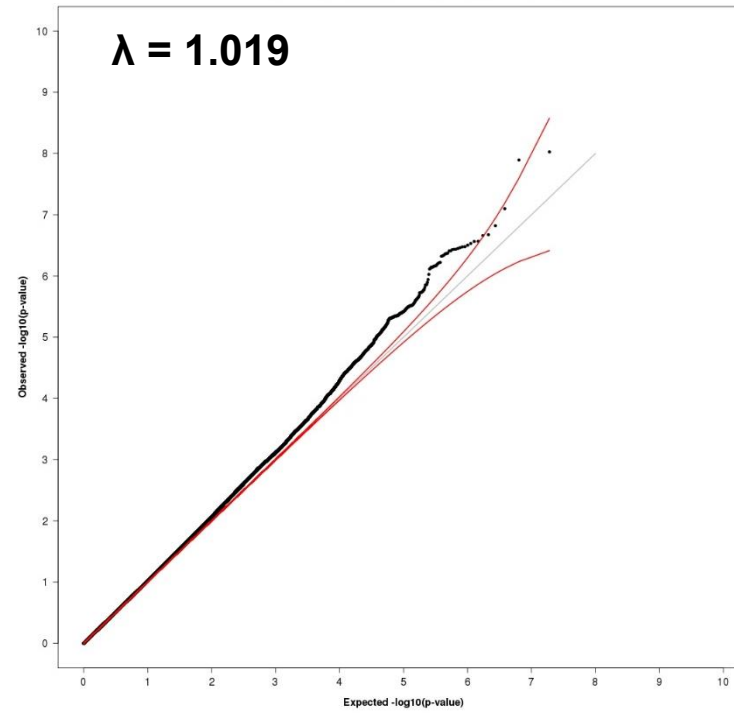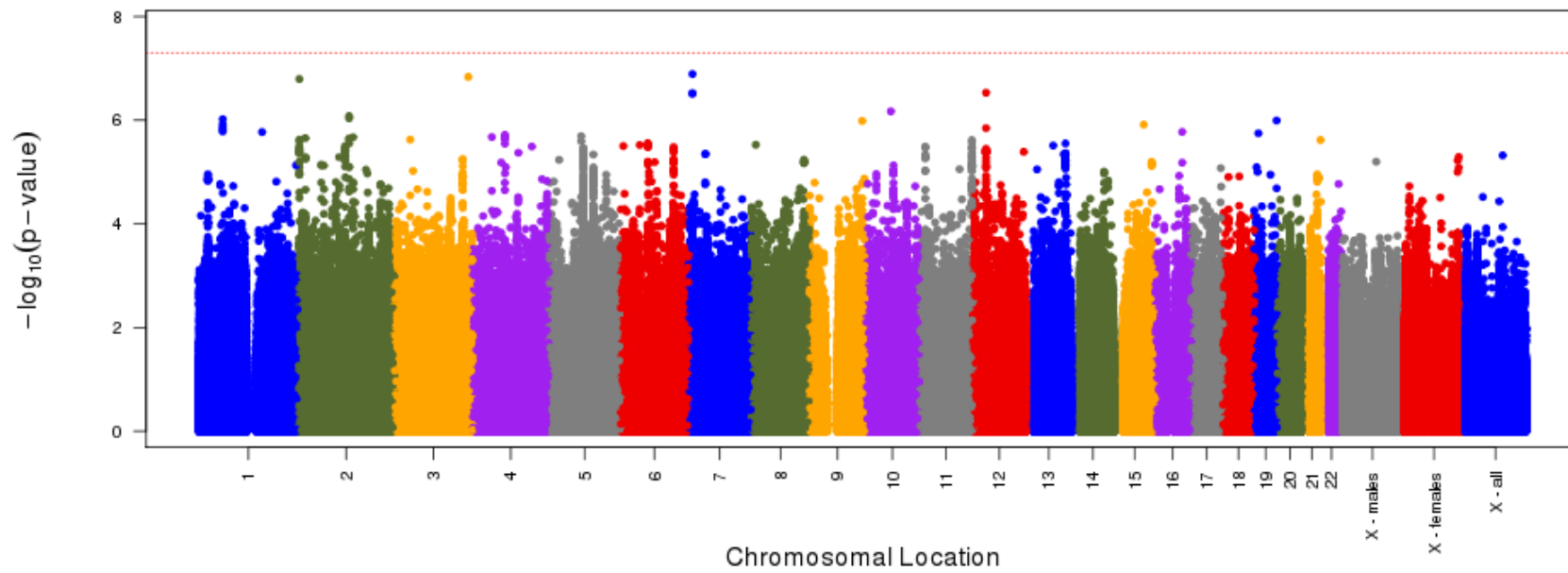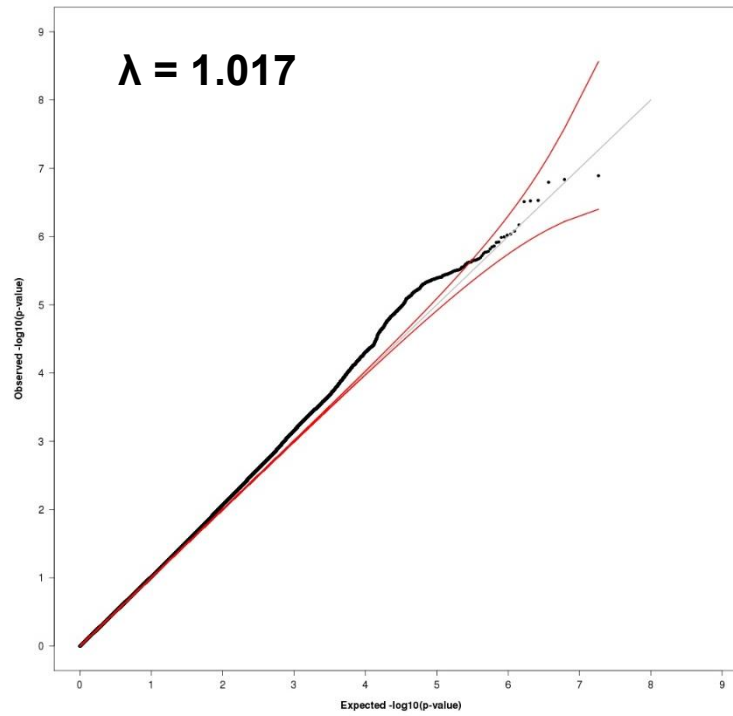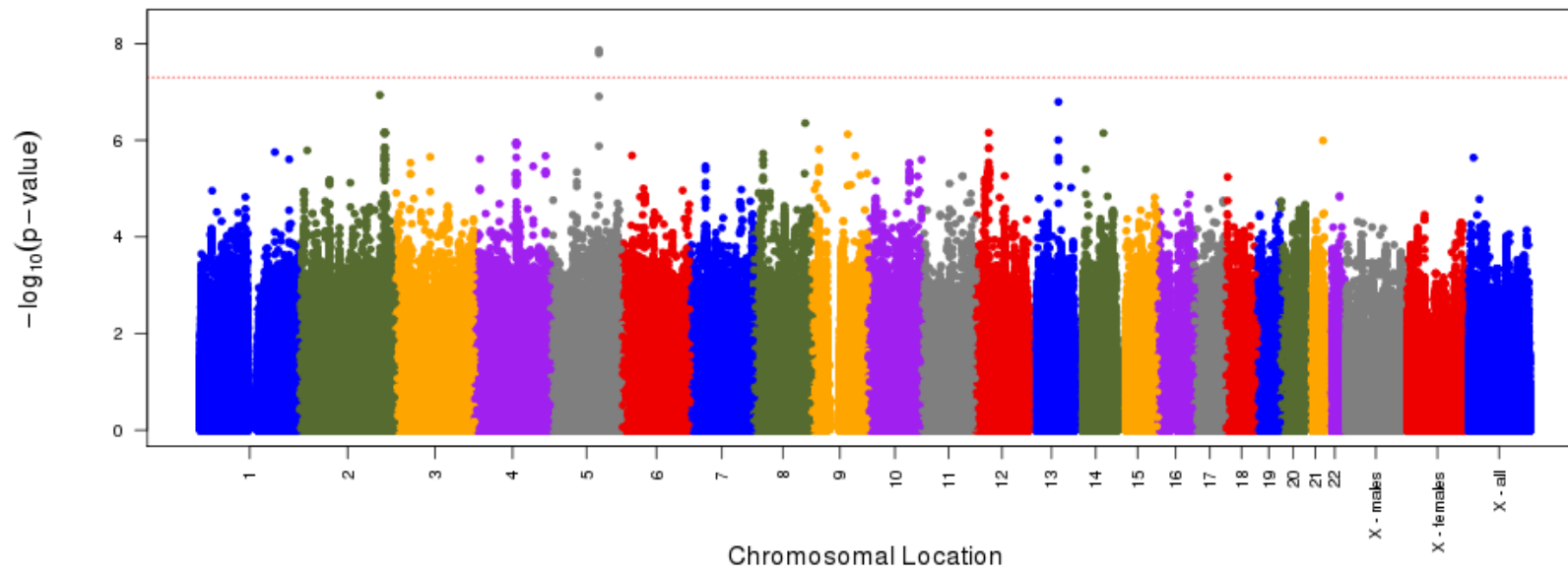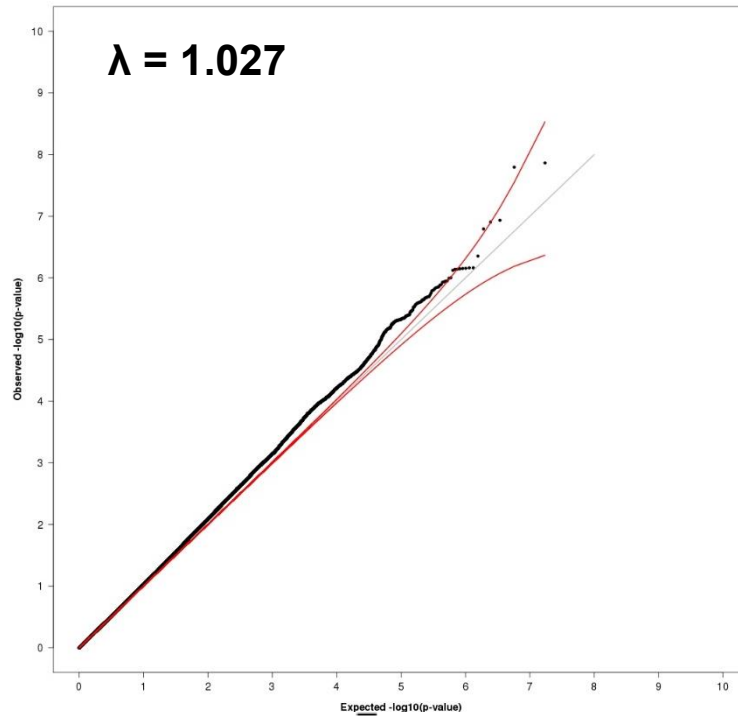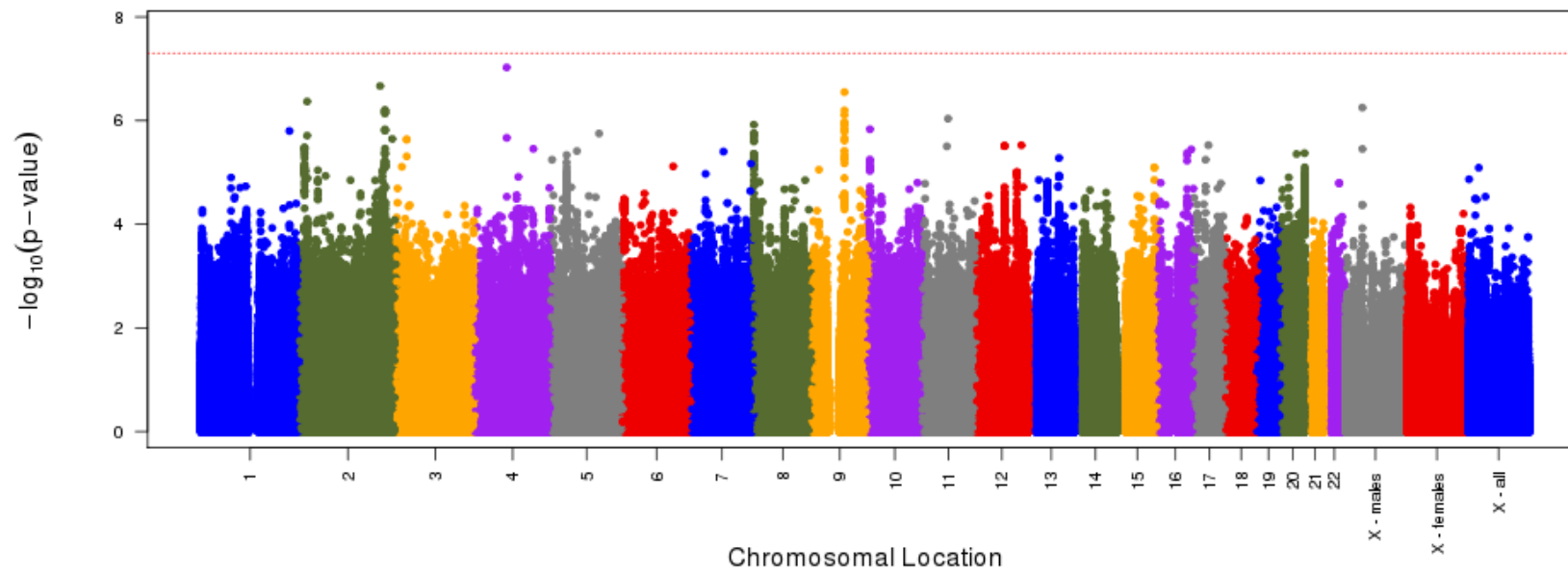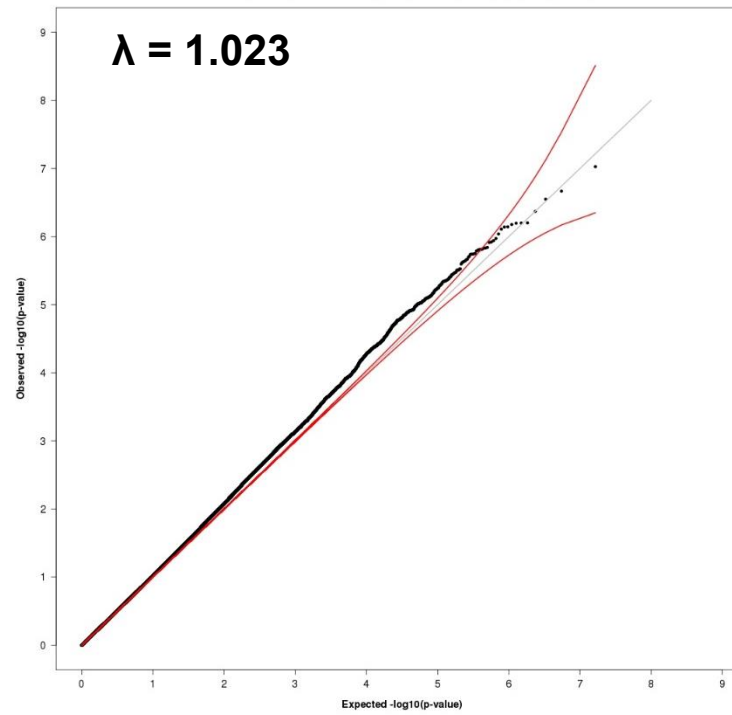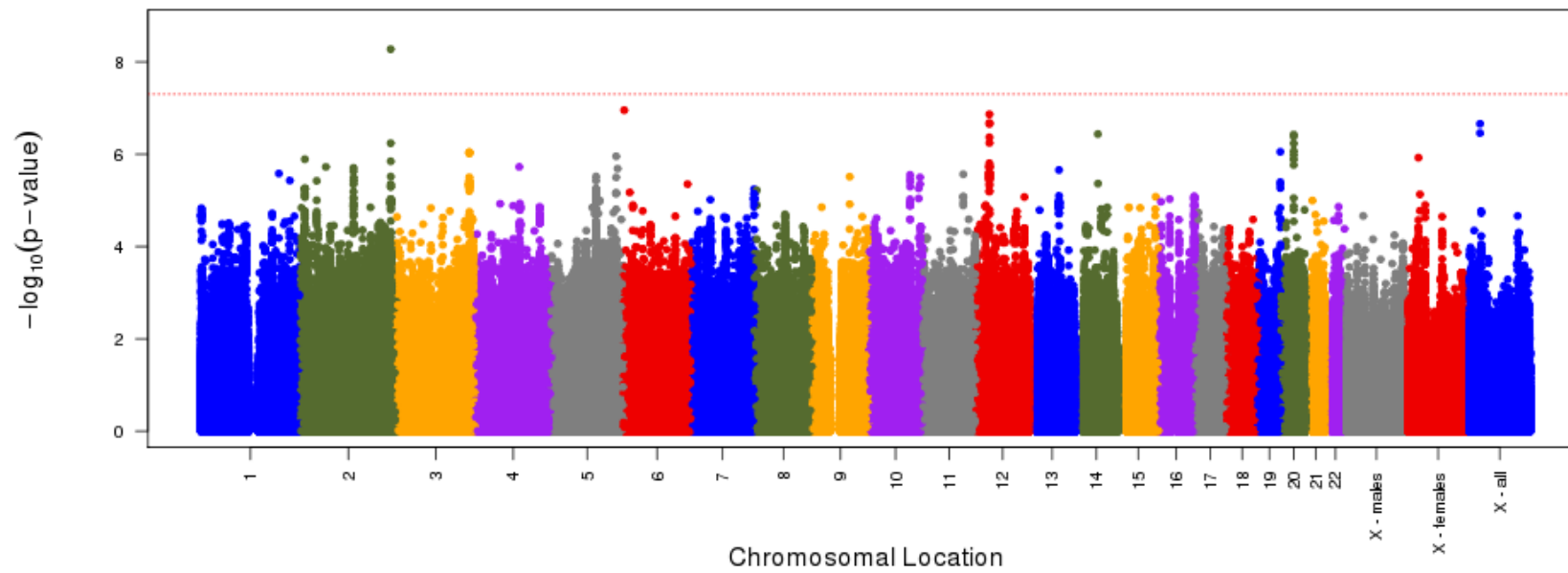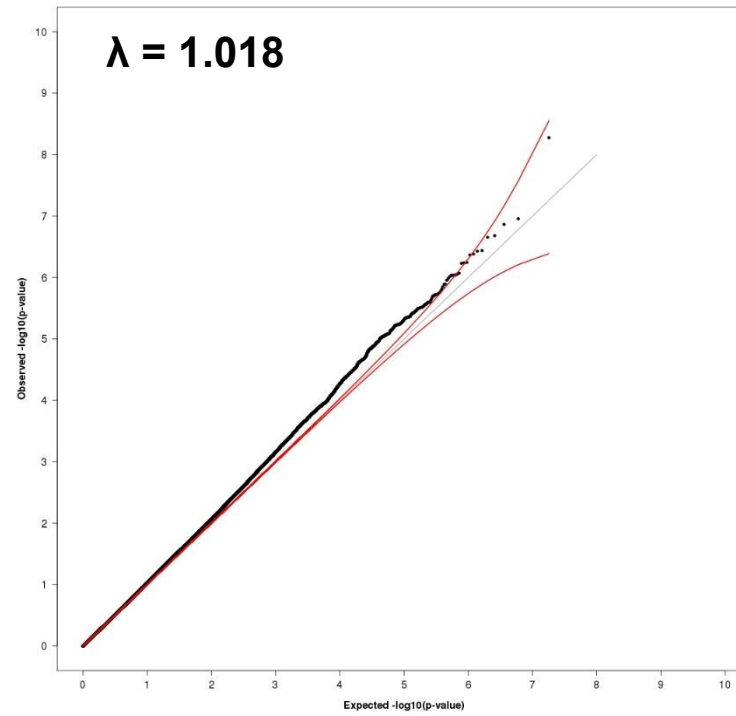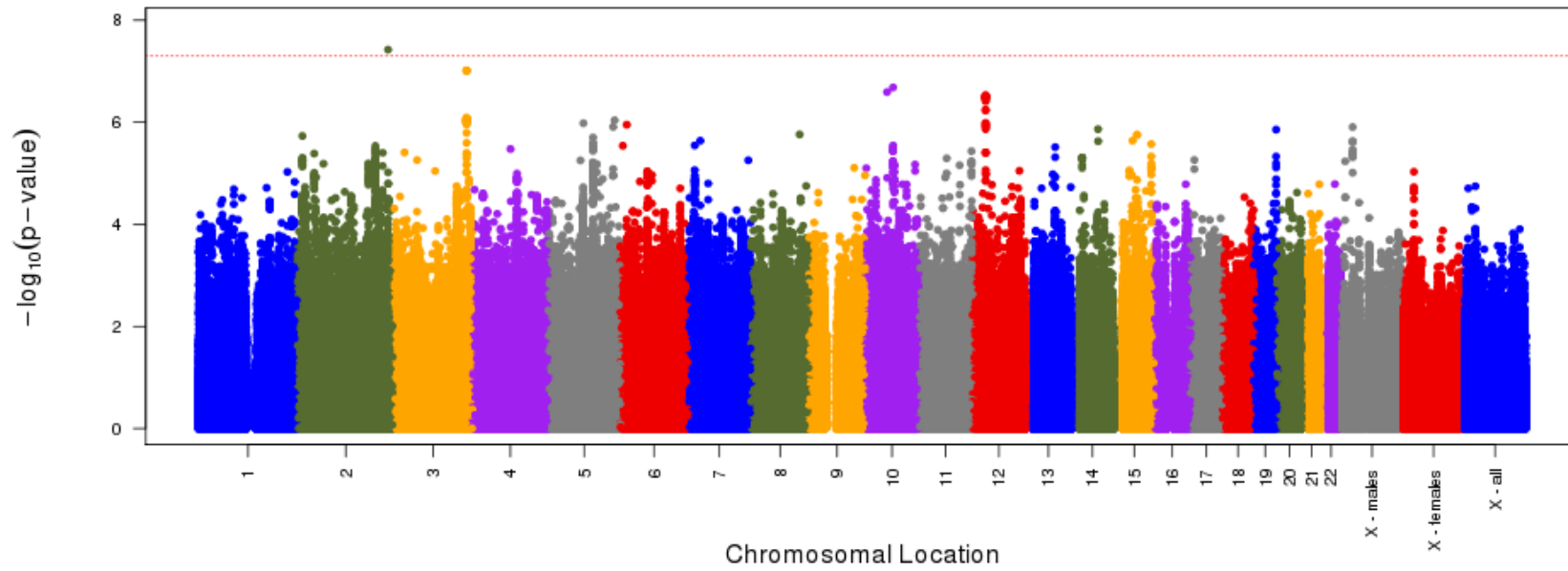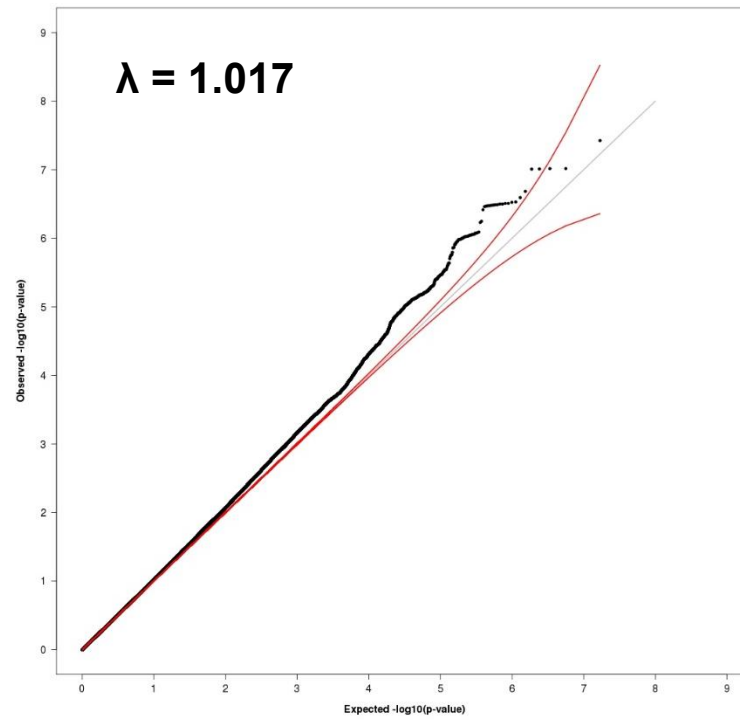
Figure S2. Regional chromosomal location plots and forest plots by cohort of newly discovered DKD associations

chr2:3745215 – rs12615970 – COLEC11 – CKD min

rs12615970 – CKD min

Cohort	EAF		ES (95% CI)
CACTI	.126699		0.90 (0.40, 2.06)
EDC	.114606		0.61 (0.34, 1.10)
EDIC	.126419		0.89 (0.48, 1.66)
FD	.1387		0.80 (0.68, 0.93)
France	.137077		0.83 (0.59, 1.17)
GWU_GOKIND	.132239		1.00 (0.61, 1.65)
Italy	.122767		0.84 (0.40, 1.77)
Joslin	.134103		0.67 (0.52, 0.86)
RomDiane	.127016		0.85 (0.38, 1.92)
Scotland	.11494		0.71 (0.54, 0.94)
Steno	.141644		0.64 (0.44, 0.93)
UK-ROI	.133887		0.83 (0.60, 1.14)
WESDR	.123773		0.67 (0.44, 1.04)
Overall (I-squared = 0.0%, p = 0.942)			0.76 (0.69, 0.84)

.34   1   2.94

chr3:926345 – rs115061173 – *LINC01266* – ESRD vs ctrl min
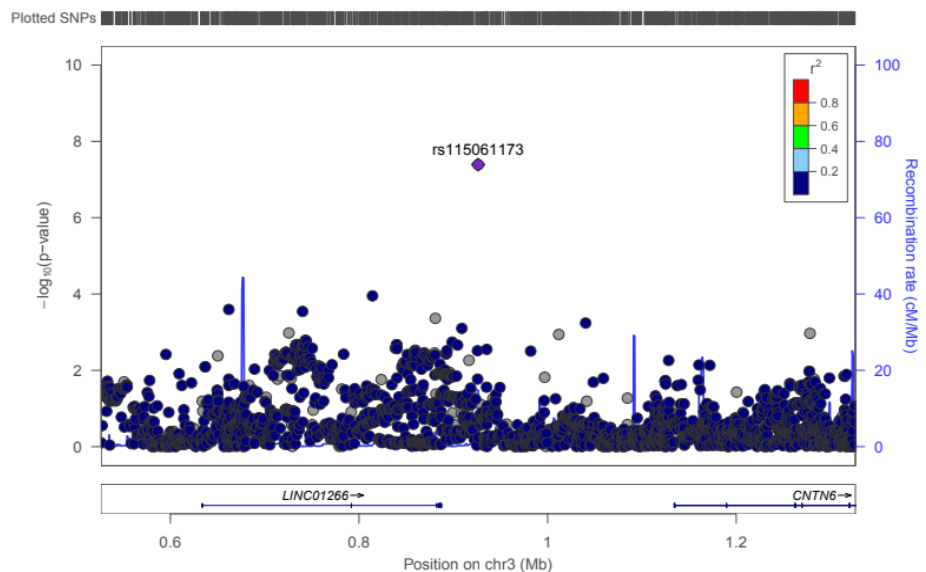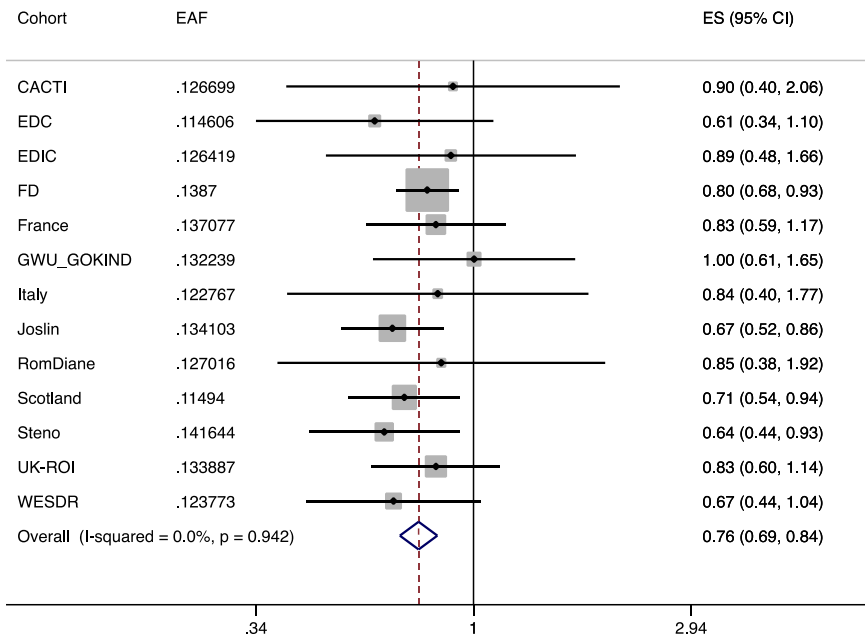
rs115061173 – ESRD vs. ctrl min

Cohort	EAF		ES (95% CI)
FinnDiane	.004891		6.03 (1.46, 25.00)
France	.017945		17.39 (4.12, 73.42)
UK-ROI	.018004		8.04 (2.06, 31.37)
Overall (I-squared = 0.0%, p = 0.570)			9.36 (4.15, 21.08)

.0136   1   73.4

Decreased risk   Increased risk
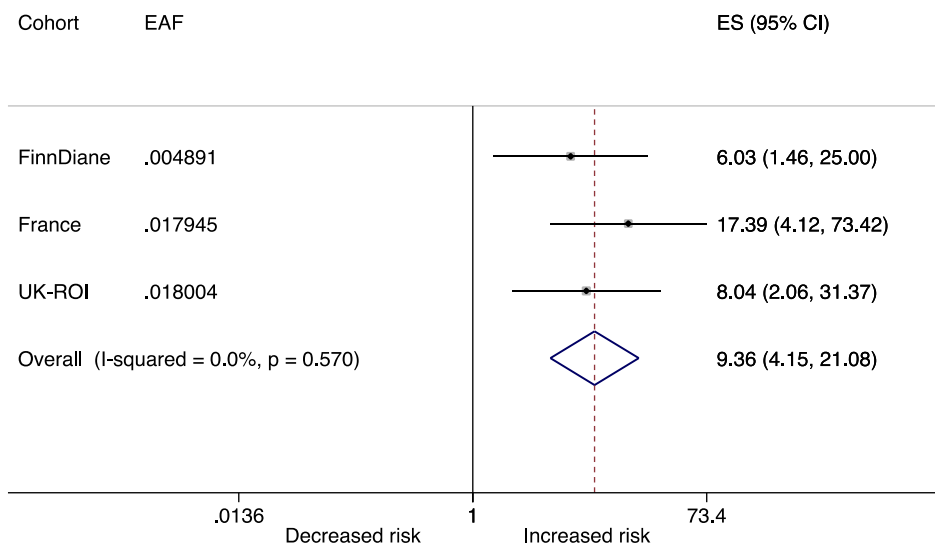
chr2:228121101 – rs55703767 – *COL4A3* – DN min

rs55703767 – DN min

Study	EAF		ES (95% CI)
EDC	.23		0.82 (0.55, 1.21)
EDIC	.21		0.63 (0.42, 0.92)
FD	.19		0.81 (0.71, 0.93)
FRANCE2_Merged	.22		0.77 (0.61, 0.98)
GWU-GOKIND	.22		0.94 (0.67, 1.31)
ITALY	.25		1.17 (0.78, 1.77)
JOSLIN	.2		0.73 (0.61, 0.87)
LATDIANE	.25		1.01 (0.47, 2.17)
LITDIANE	.24		1.73 (0.61, 4.90)
ROMDIANE	.22		0.61 (0.36, 1.01)
SCOTLAND	.21		0.81 (0.63, 1.05)
STENO	.2		0.76 (0.60, 0.97)
SWEDEN	.22		0.85 (0.49, 1.48)
UK-ROI	.19		0.70 (0.58, 0.86)
WESDR	.24		0.83 (0.60, 1.13)
Overall (I-squared = 0.0%, p = 0.601)			0.79 (0.73, 0.84)

Decreased risk   Increased risk

chr3:11910635 – rs142823282 – *TAMM41* – Micro full

rs142823282 – Micro full

Cohort	EAF		ES (95% CI)
Joslin	.015959		1.40 (0.57, 3.46)
Scotland	.017362		24.97 (10.94, 56.99)
Overall (I-squared = 95.3%, p = 0.000)			6.74 (3.66, 12.38)

Decreased risk   Increased risk

64
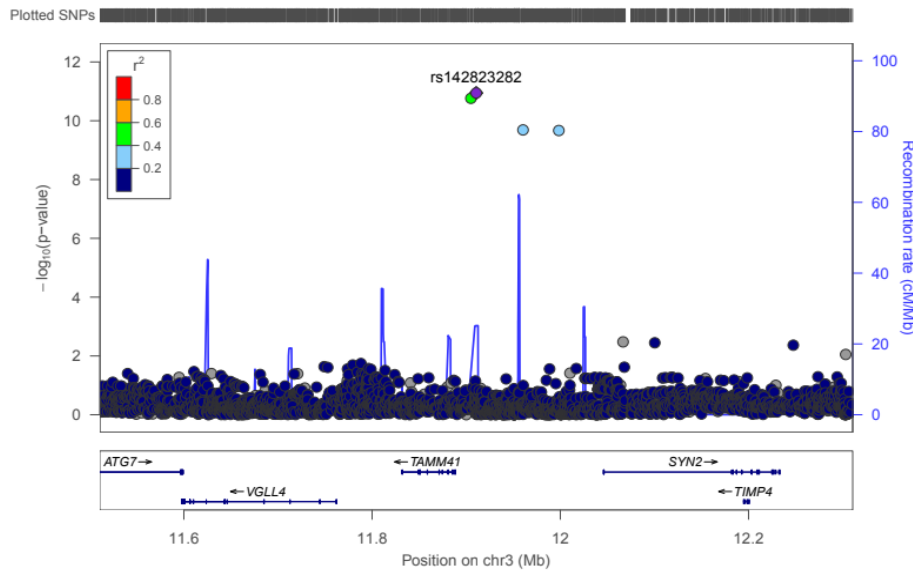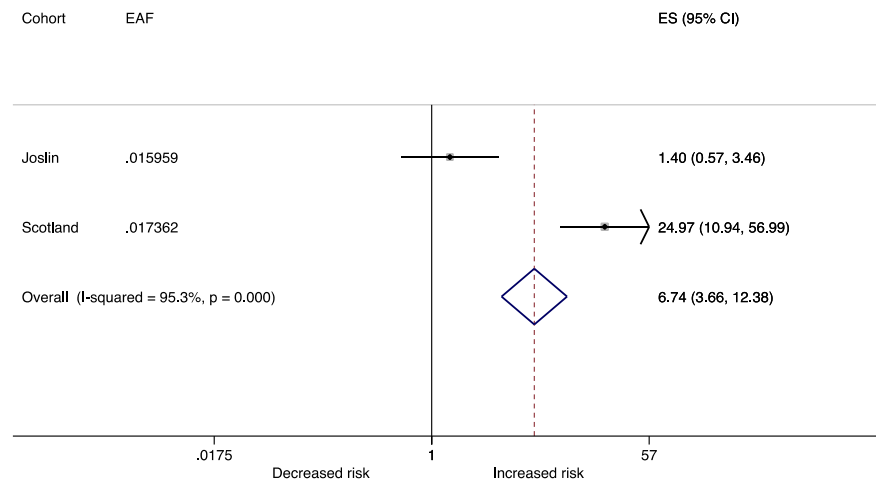
chr4:71358776 – rs191449639 – *MUC7* – DN min

rs191449639 – DN min

Cohort	EAF		ES (95% CI)
FinnDiane	.003502		2.57 (0.58, 11.31)
Scotland	.007543		3869.23 (504.37, 29682.47)
Overall (I-squared = 96.9%, p = 0.000)			32.42 (9.77, 107.59)

3.4e-05 Decreased risk 1 Increased risk 29682
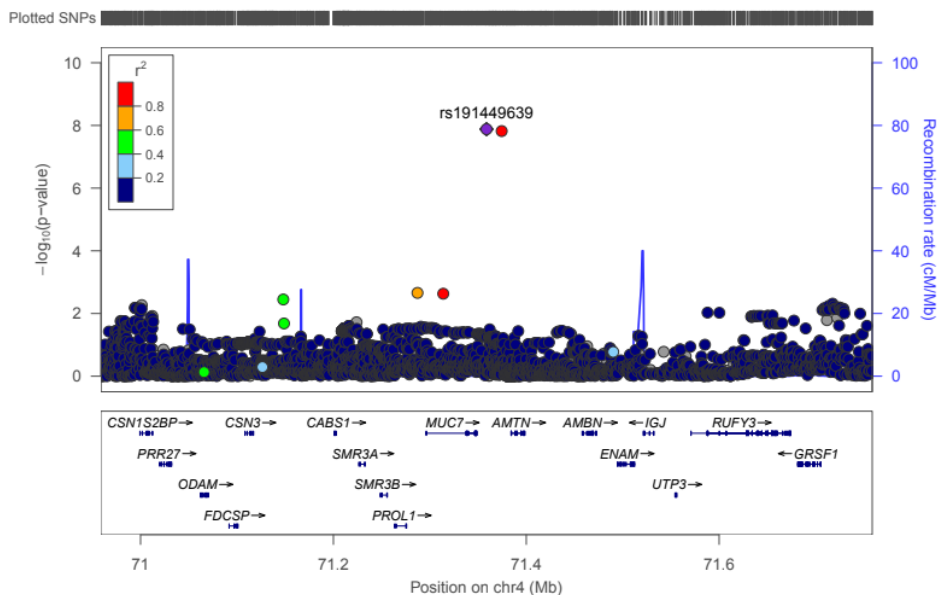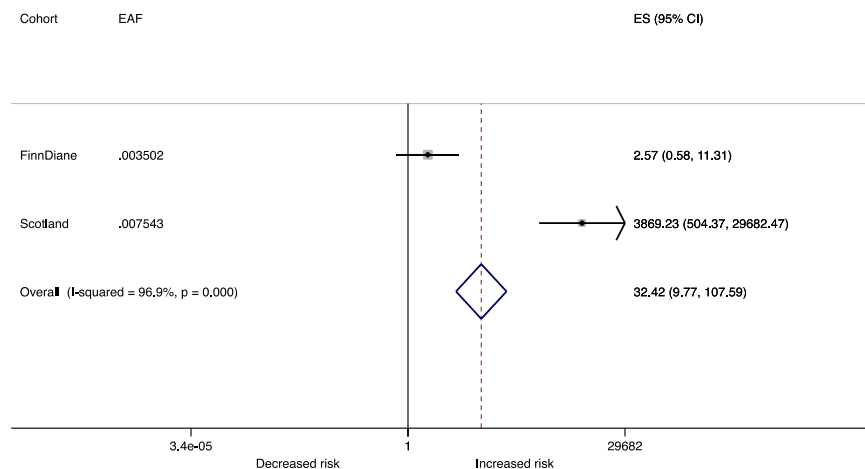
chr5:121774582 – rs149641852 – *SNCAIP* – CKD extreme min

rs149641852 – CKD extreme min

Cohort	EAF		ES (95% CI)
FD	.008086		5.26 (1.77, 15.63)
France	.015511		63.97 (11.33, 361.07)
UK-ROI	.016152		6.29 (1.65, 23.97)
Overall (I-squared = 67.4%, p = 0.046)			9.01 (4.22, 19.25)

.00277 Decreased risk 1 Increased risk 361
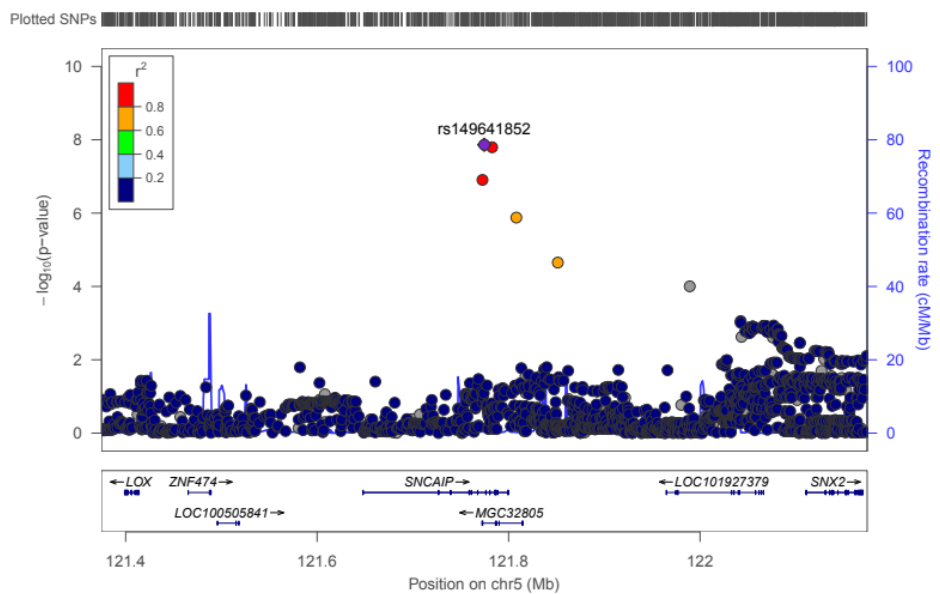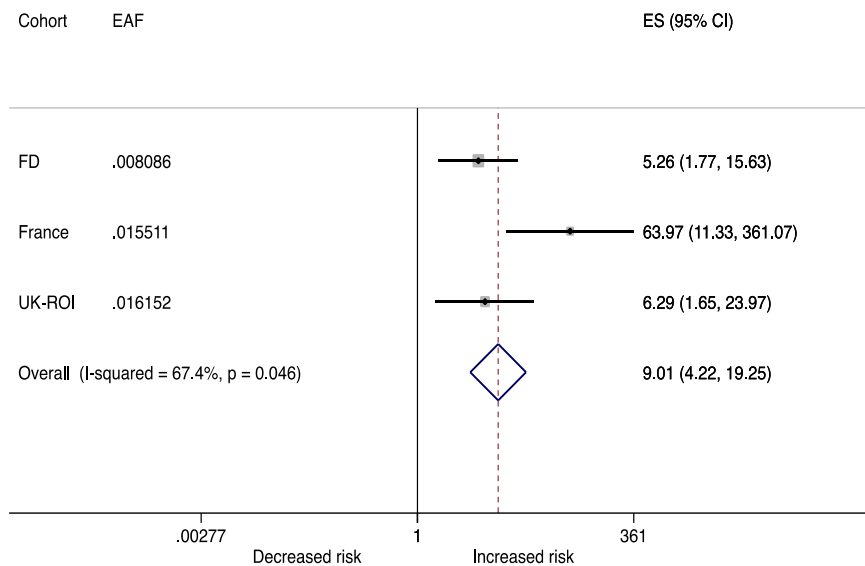
chr6:30865279 – rs118124843 – *DDR1* – Micro full

rs118124843 – Micro full

Cohort	EAF		ES (95% CI)
Joslin	.011969		4.74 (2.29, 9.82)
Scotland	.010477		3.21 (1.70, 6.04)
Overall (I-squared = 0.0%, p = 0.426)			3.80 (2.36, 6.12)

.102 — 1 — 9.82
Decreased risk · Increased risk
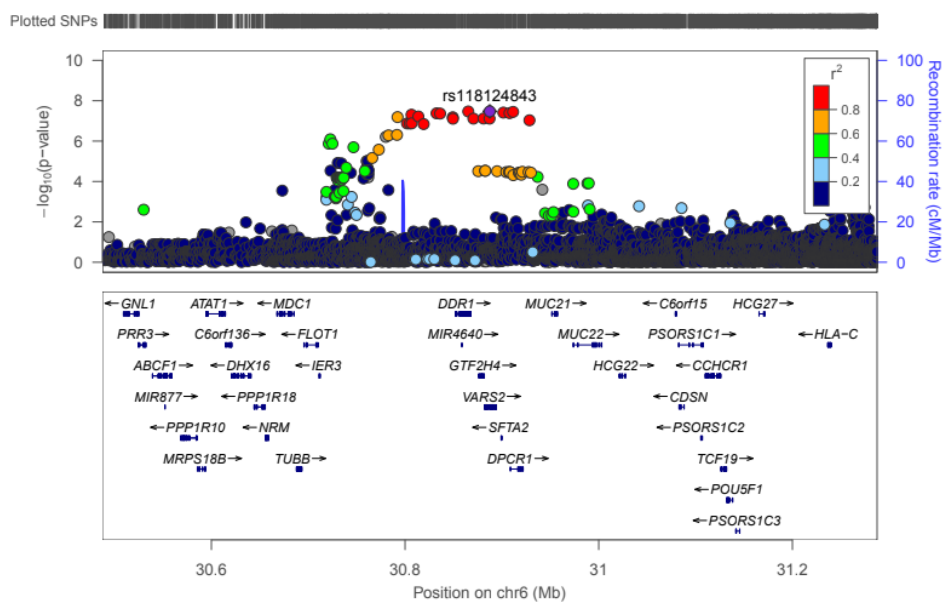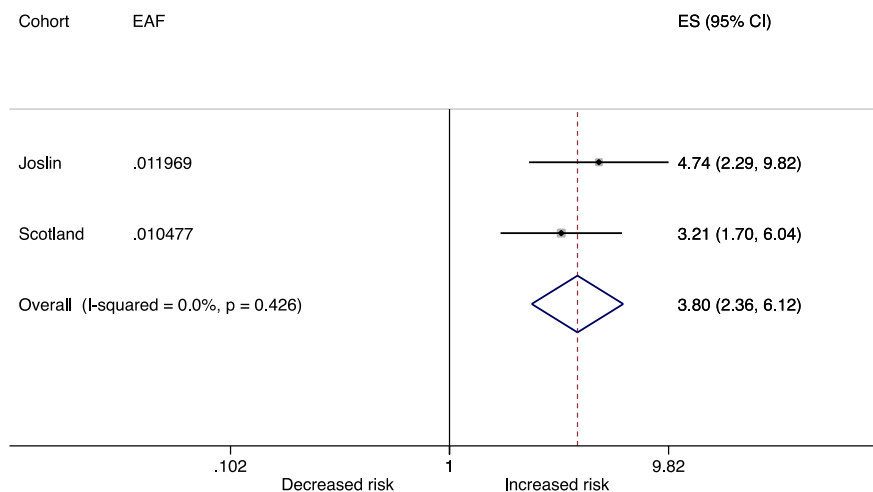
chr8:128100029 – rs551191707 – *PRNCR1* – ESRD vs macro min
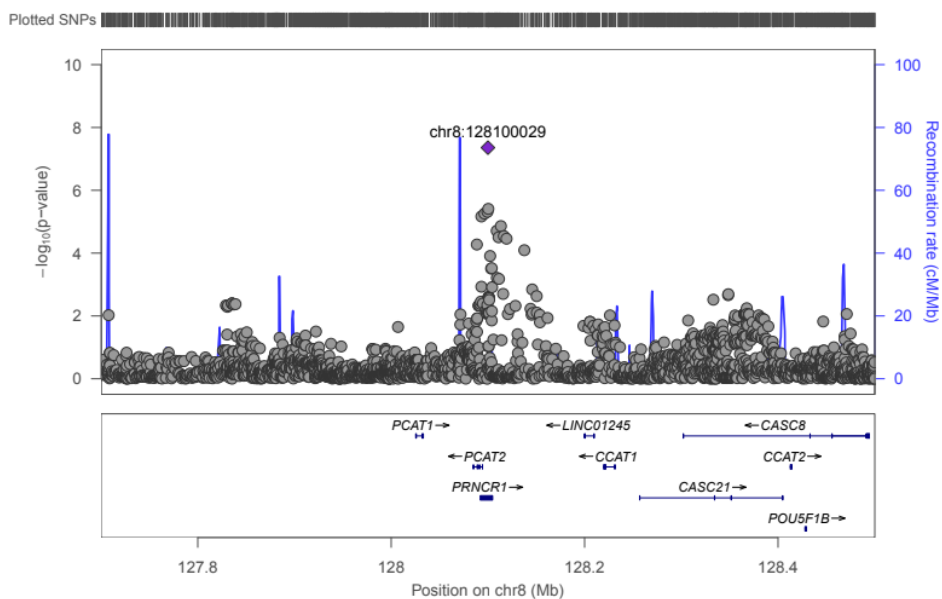
rs551191707 – ESRD vs. macro min

Cohort	EAF		ES (95% CI)
EDC	.12		2.61 (0.98, 6.96)
FD	.14		1.48 (1.14, 1.91)
FRANCE2_Merged	.1		2.53 (1.28, 5.03)
JOSLIN	.1		1.36 (0.84, 2.23)
SCOTLAND	.09		2.20 (0.84, 5.76)
UK-ROI	.11		2.56 (1.54, 4.26)
WESDR	.09		1.36 (0.53, 3.49)
Overall (I-squared = 13.7%, p = 0.325)			1.69 (1.40, 2.04)

.144 — 1 — 6.96
Decreased risk · Increased risk

66
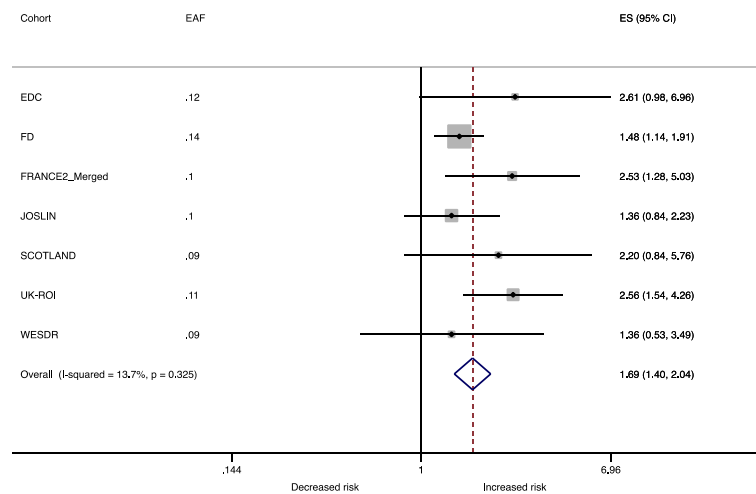
chr7:99728546 – rs77273076 – *MBLAC1* – Micro min

rs77273076 – Micro full

Cohort	EAF		ES (95% CI)
FinnDiane	.005779		11.84 (3.12, 45.00)
Scotland	.008895		8.11 (3.23, 20.37)
Overall (I-squared = 0.0%, p = 0.647)			9.16 (4.29, 19.56)

.0222   Decreased risk   1   Increased risk   45
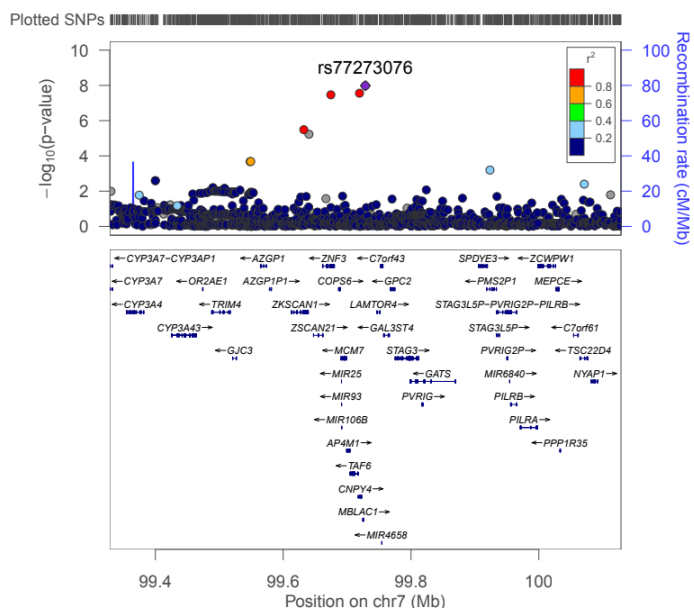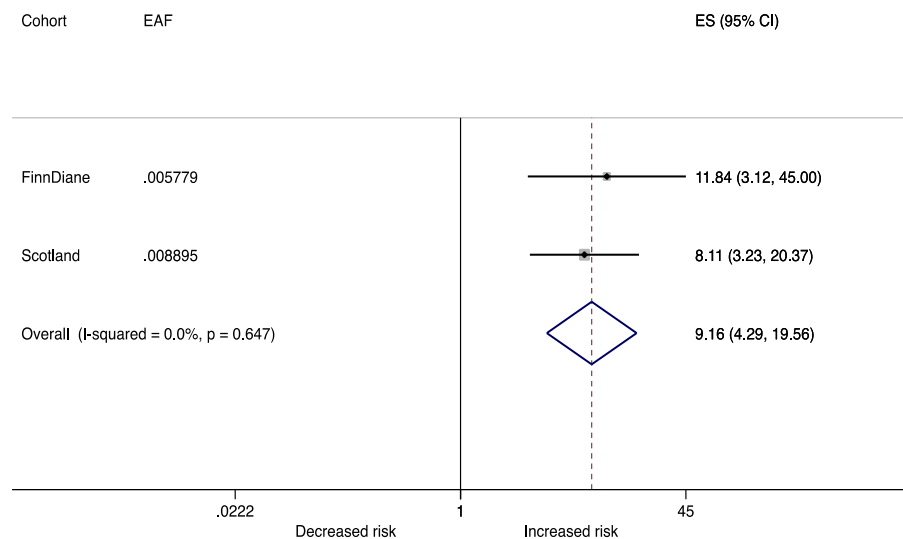
chr11:16937846 – rs183937294 – *PLEKHA7* – Micro min

rs183937294 – Micro min

Cohort	EAF		ES (95% CI)
Joslin	.006582		11.13 (2.39, 51.74)
Scotland	.008037		23.43 (6.45, 85.11)
Overall (I-squared = 0.0%, p = 0.467)			17.22 (6.41, 46.26)

.0117   Decreased risk   1   Increased risk   85.1

67
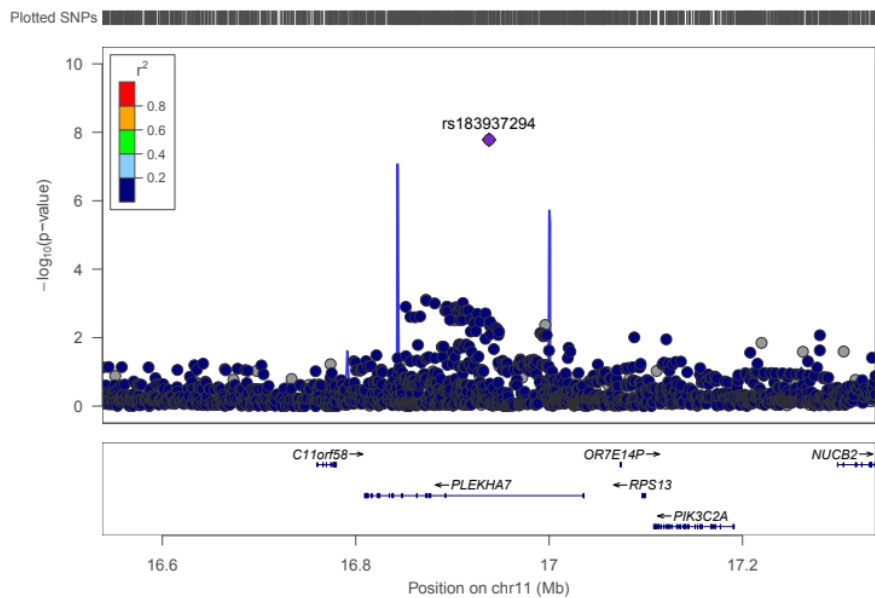
chr14:26004712 – rs61983410 – *STXBP6* – Micro full

rs61983410 – Micro full

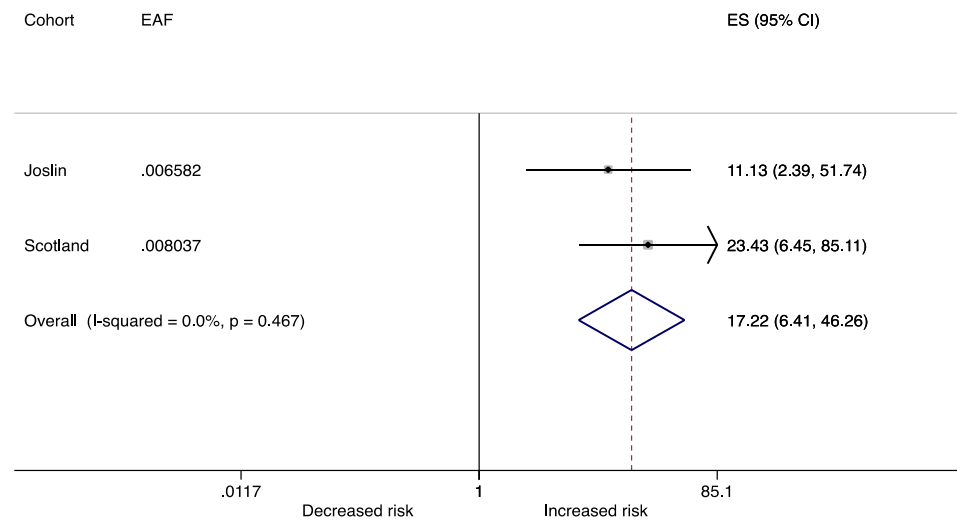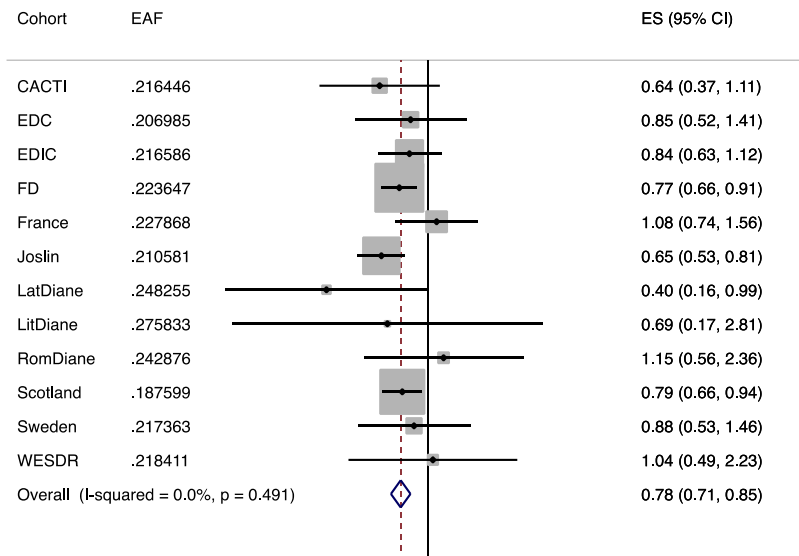| Cohort	EAF	ES (95% CI)
CACTI	.216446	0.64 (0.37, 1.11)
EDC	.206985	0.85 (0.52, 1.41)
EDIC	.216586	0.84 (0.63, 1.12)
FD	.223647	0.77 (0.66, 0.91)
France	.227868	1.08 (0.74, 1.56)
Joslin	.210581	0.65 (0.53, 0.81)
LatDiane	.248255	0.40 (0.16, 0.99)
LitDiane	.275833	0.69 (0.17, 2.81)
RomDiane	.242876	1.15 (0.56, 2.36)
Scotland	.187599	0.79 (0.66, 0.94)
Sweden	.217363	0.88 (0.53, 1.46)
WESDR	.218411	1.04 (0.49, 2.23)
Overall (I-squared = 0.0%, p = 0.491)		0.78 (0.71, 0.85)
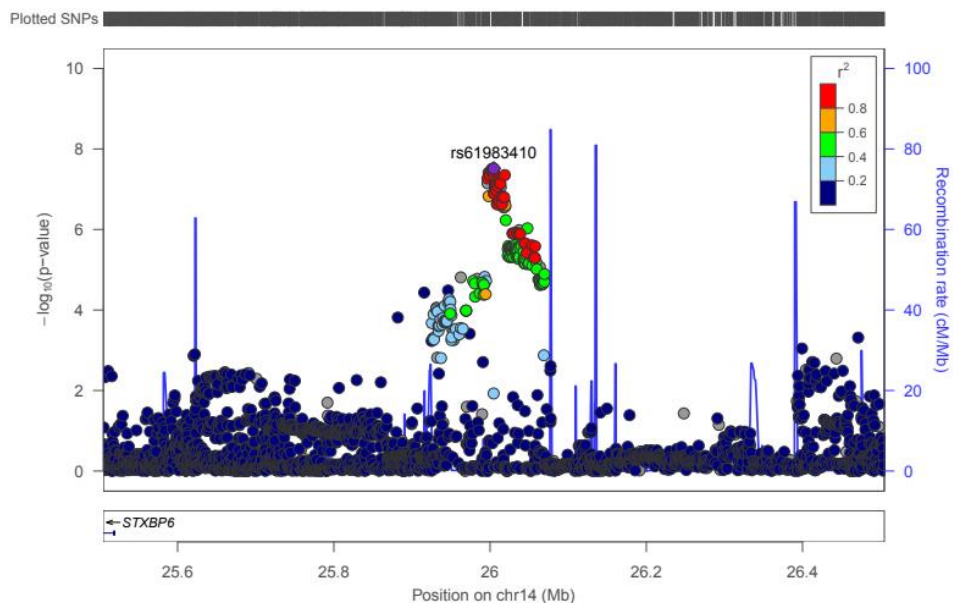
chr18:1811108 – rs185299109 – 18p11 – CKD min

rs185299109 – CKD full

Cohort	EAF	ES (95% CI)
Joslin	.008203	6.68 (1.65, 27.07)
Scotland	.005833	86.71 (18.01, 417.38)
Overall (I-squared = 82.5%, p = 0.017)		20.75 (7.30, 58.99)

.0024   1   417

Decreased risk   Increased risk

68
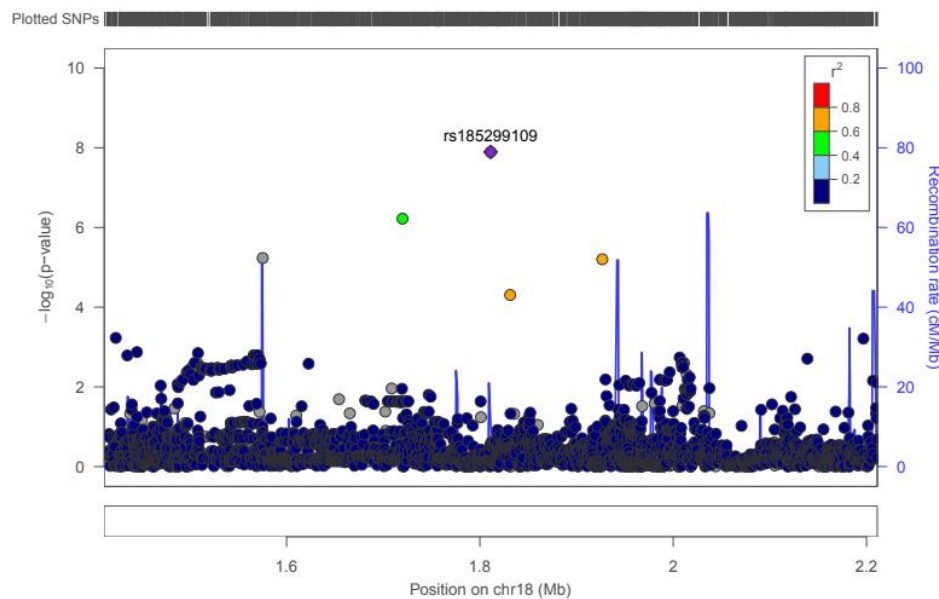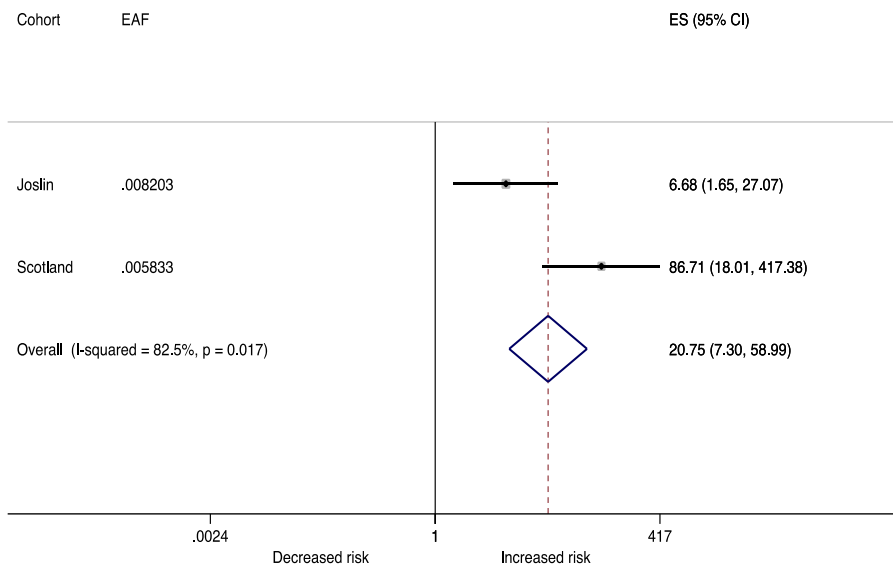
chr14:73740250 – rs113554206 – *PAPLN* – Macro full
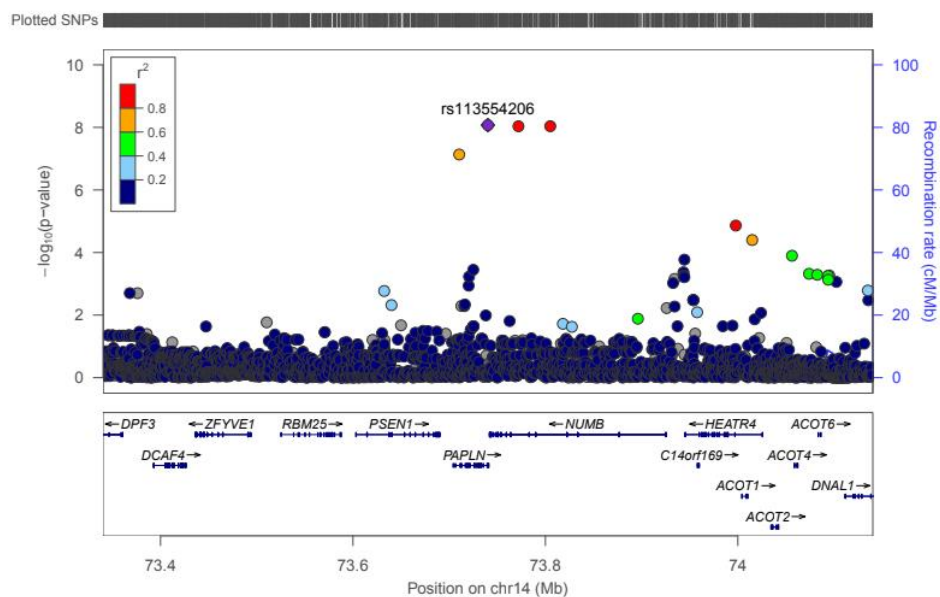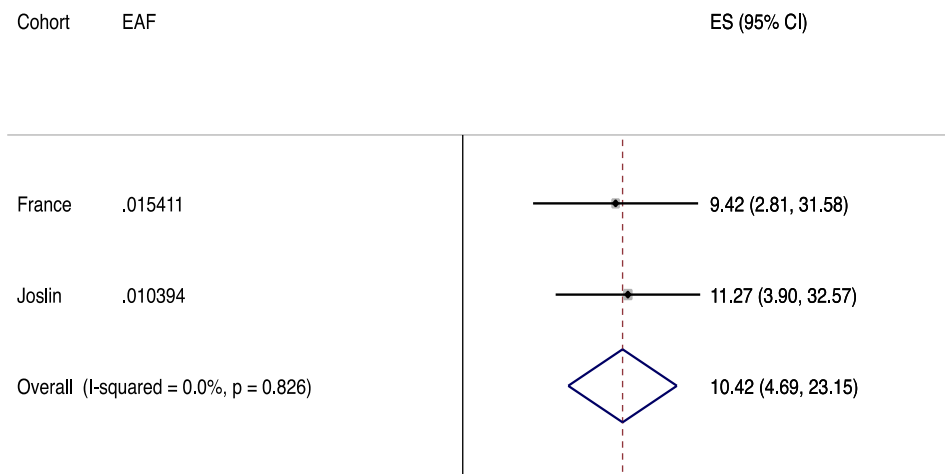
chr20:55837263 – rs144434404 – *BMP7* – Micro min

rs113554206 – Macro full

Cohort	EAF	ES (95% CI)
France	.015411	9.42 (2.81, 31.58)
Joslin	.010394	11.27 (3.90, 32.57)
Overall (I-squared = 0.0%, p = 0.826)		10.42 (4.69, 23.15)

rs144434404 – Micro min

Cohort	EAF	ES (95% CI)
Joslin	.008253	6.29 (1.69, 23.37)
Scotland	.012241	6.94 (3.38, 14.24)
Overall (I-squared = 0.0%, p = 0.897)		6.78 (3.61, 12.74)

.0428   1   23.4

Decreased risk   Increased risk
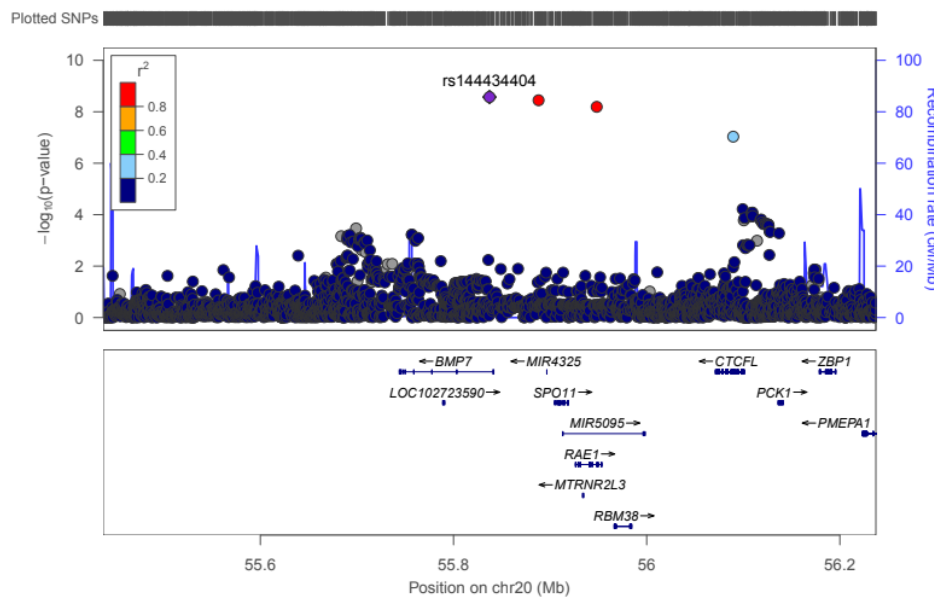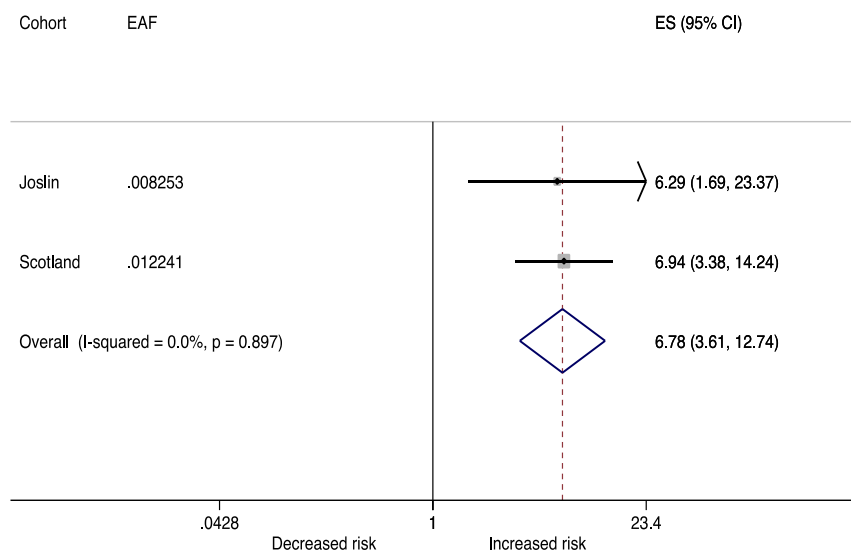
Figure S3. Correlation of expression of *COL4A3* with degree of fibrosis and eGFR in microdissected kidney samples.
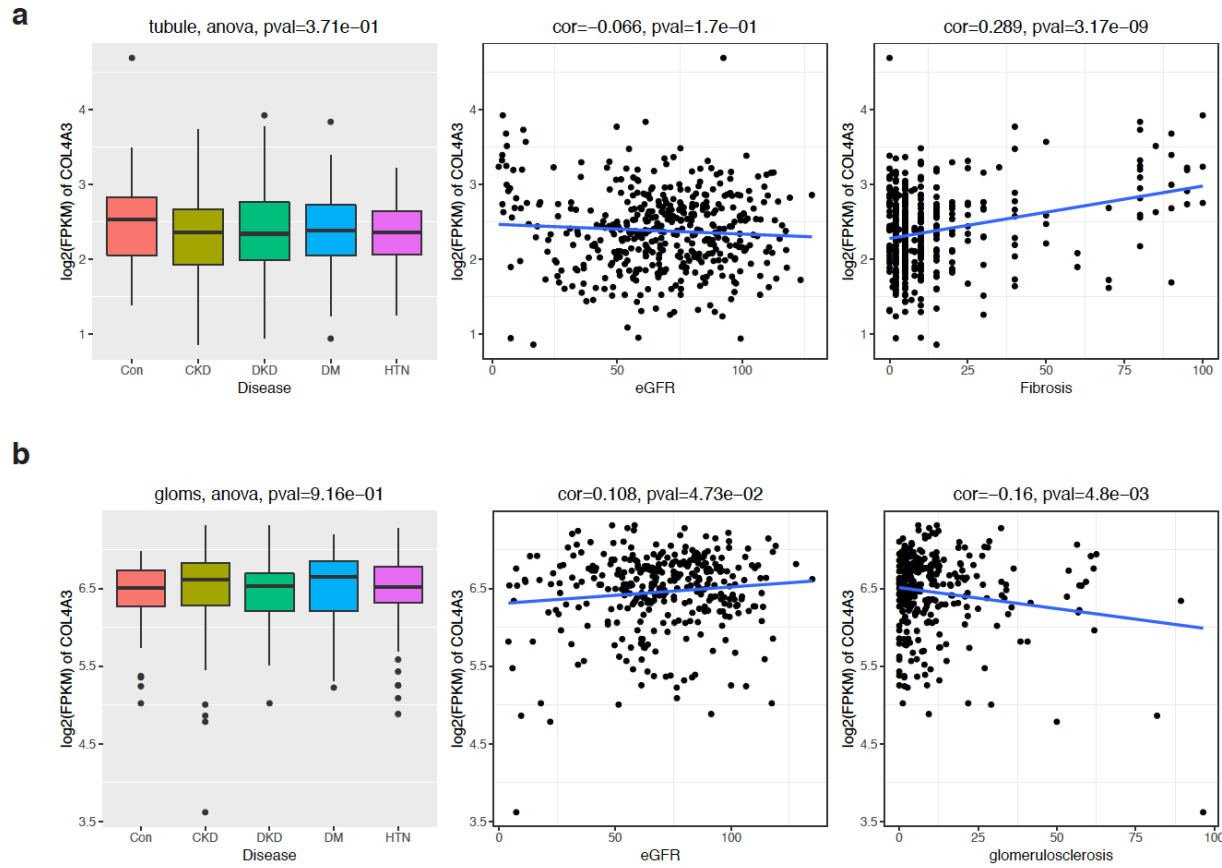
Figure S4. Genotype – phenotype associations at the lead loci when stratified by mean HbA$_{1c}$ <7.5% in the FinnDiane study.
Only loci with a minor allele count ≥10 in each stratum are shown.

Figure S5. Genotype – phenotype associations at the lead rs55703767 (COL4A3) locus when stratified by mean HbA1c <7.5% in up to 3226 individuals with type 2 diabetes (T2D) from the GoDARTS.

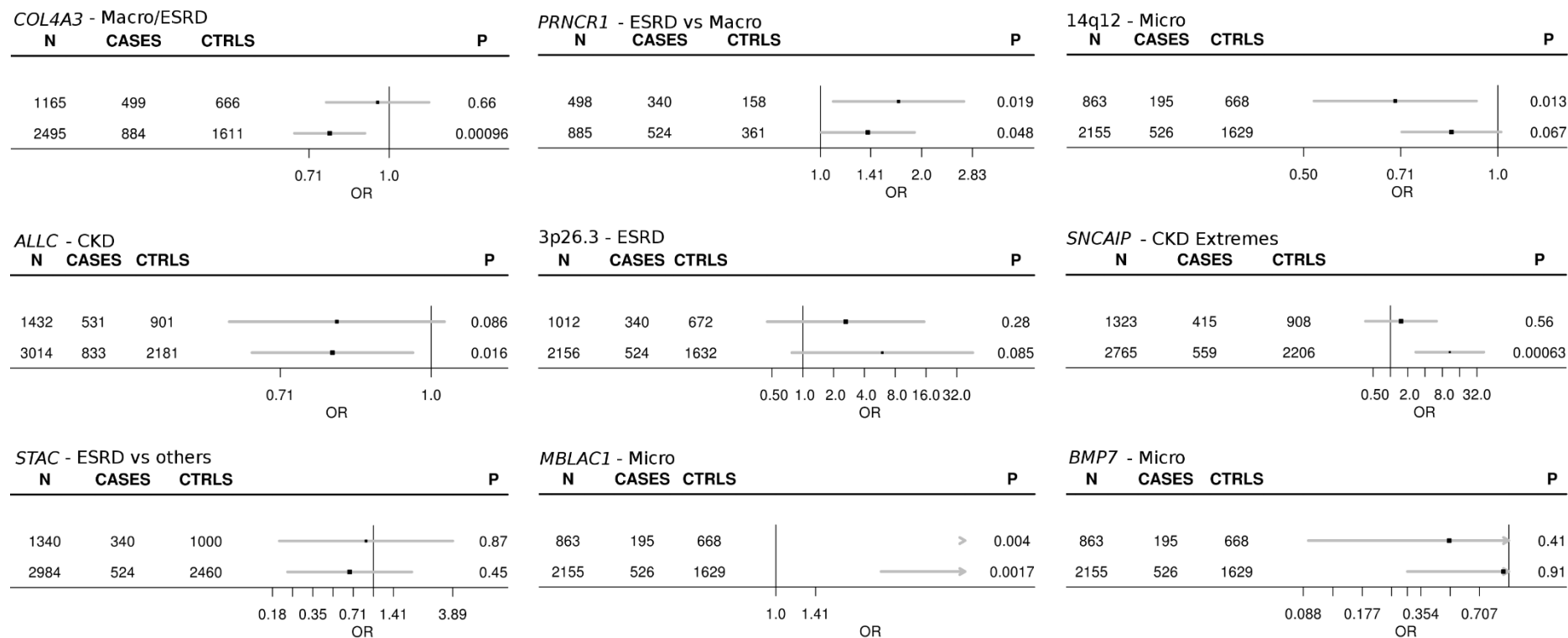For All vs. ctrl phenotype, 1632 individuals (848 cases, 784 controls) had HbA1c<7.5%, and 1572 individuals (874 cases, 698 controls) had HbA1c>=7.5%.

Pheno	P		P HbA1c<7.5	P HbA1>7.5
All vs. ctrl	0.003		0.02	0.07
DN	0.028		0.03	0.26
Macro vs. ctrl	0.058		0.01	0.55
CKD-DN	0.851		0.57	0.77

0.90 0.95 1.0 1.05
OR (95% CI)
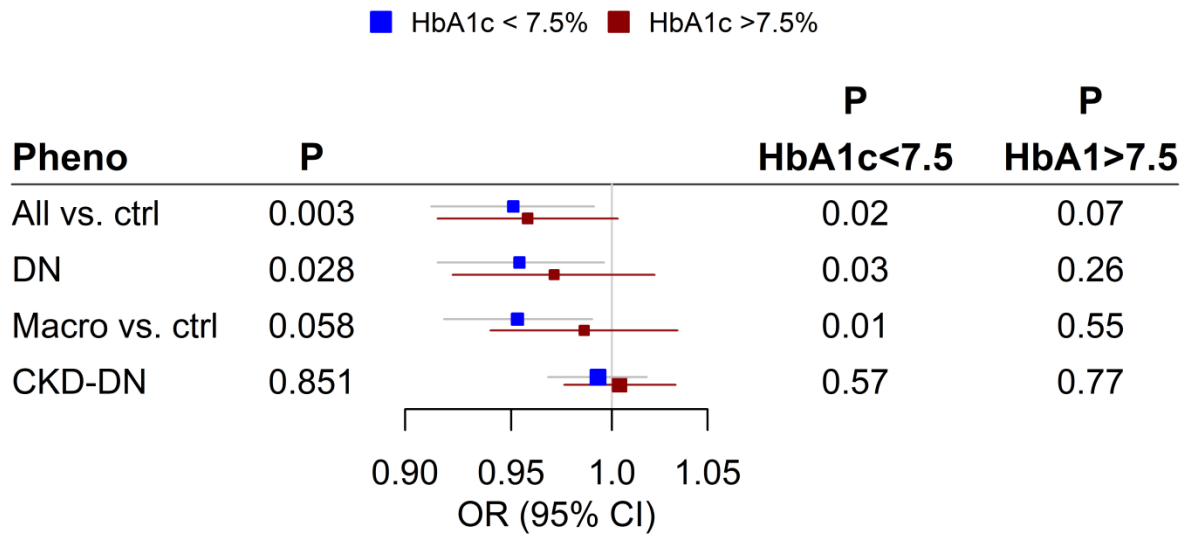
Figure S6. Fishplots comparing significance and directionality between minimal and fully adjusted models for each of the 10 phenotype definitions. Fishplots comparing the significance and directionality between the minimal and fully adjusted models for each of the 10 phenotype definitions. P-values are signed according to consistency in the direction of effect between the two GWAS under comparison, such that the –log(P) of SNPs with effect sizes in the same direction are plotted on quadrant 1 (the head and body of the fish), and the –log(P) of SNPs with effect sizes in opposite directions are plotted in quadrant 3 (the tail of the fish).
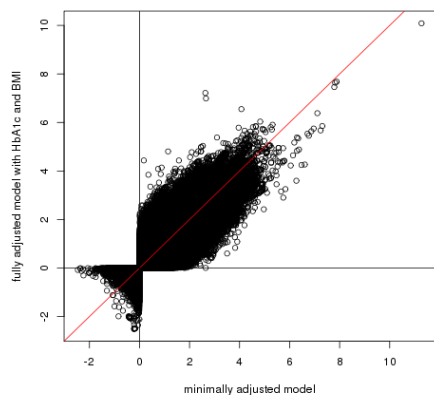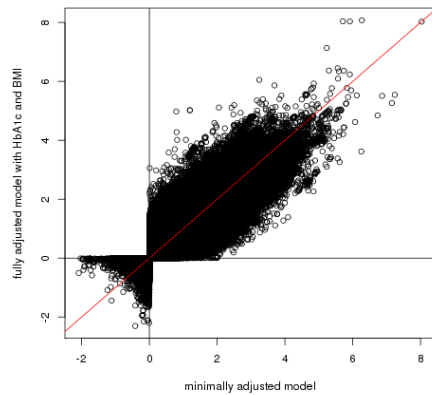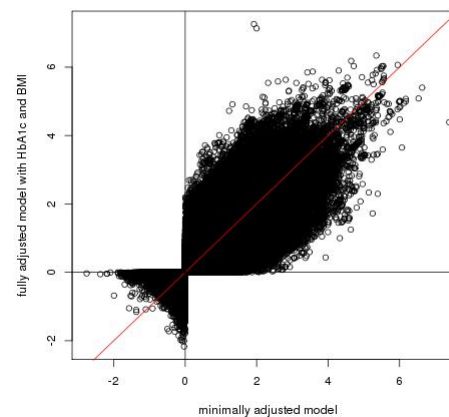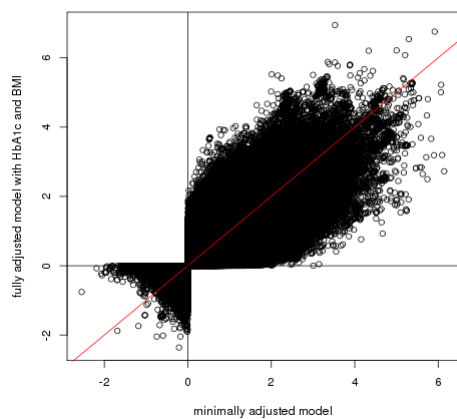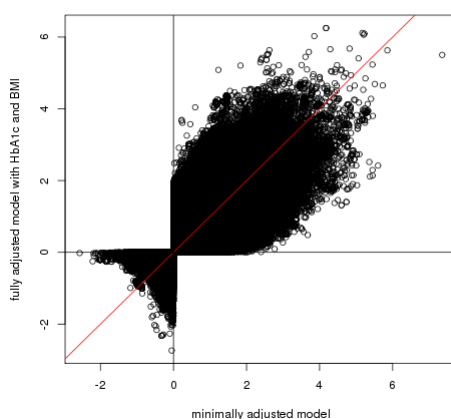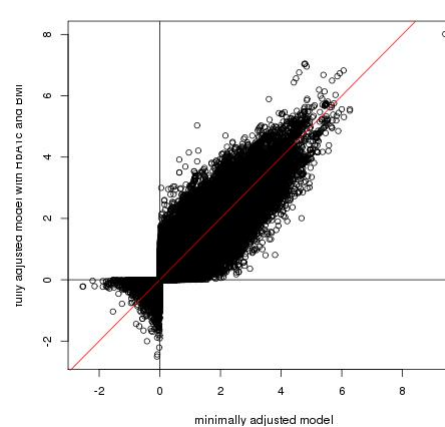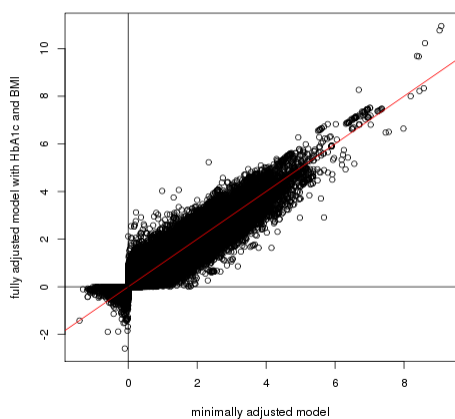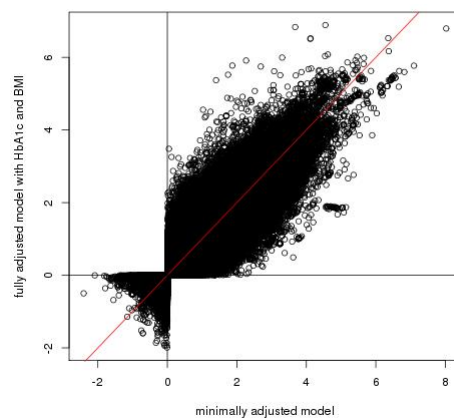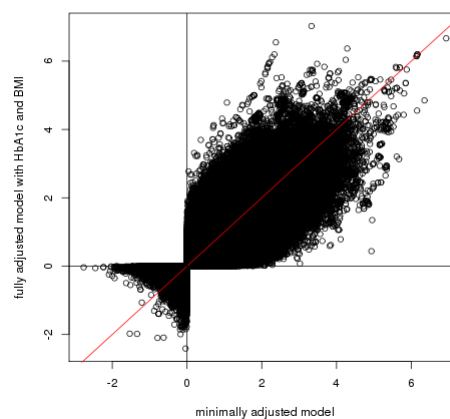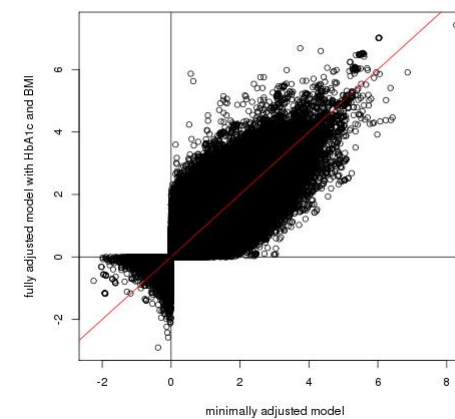
Figure S7: Association at previously reported loci (p<5×10⁻⁸) for renal complications in individuals with diabetes. *AFF3 and RGMA-MCTP2* were originally reported for ESRD (T1D) (Sandholm et al., 2012); *CDCA7/SP3* for ESRD in women (T1D) (Sandholm et al., 2013); *ERBB4* for DN (T1D) (Sandholm et al., 2012); *GABRR1* for microalbuminuria (T2D)(Van Zuydam et al., 2018); *GLRA3* for albuminuria (T1D) (Sandholm et al., 2014); *PRKAG2* and *UMOD* for eGFR (Pattaro et al., 2016; Van Zuydam et al., 2018); and *SCAF8/CNKSR3* for DN (T2D) (Iyengar et al., 2015).
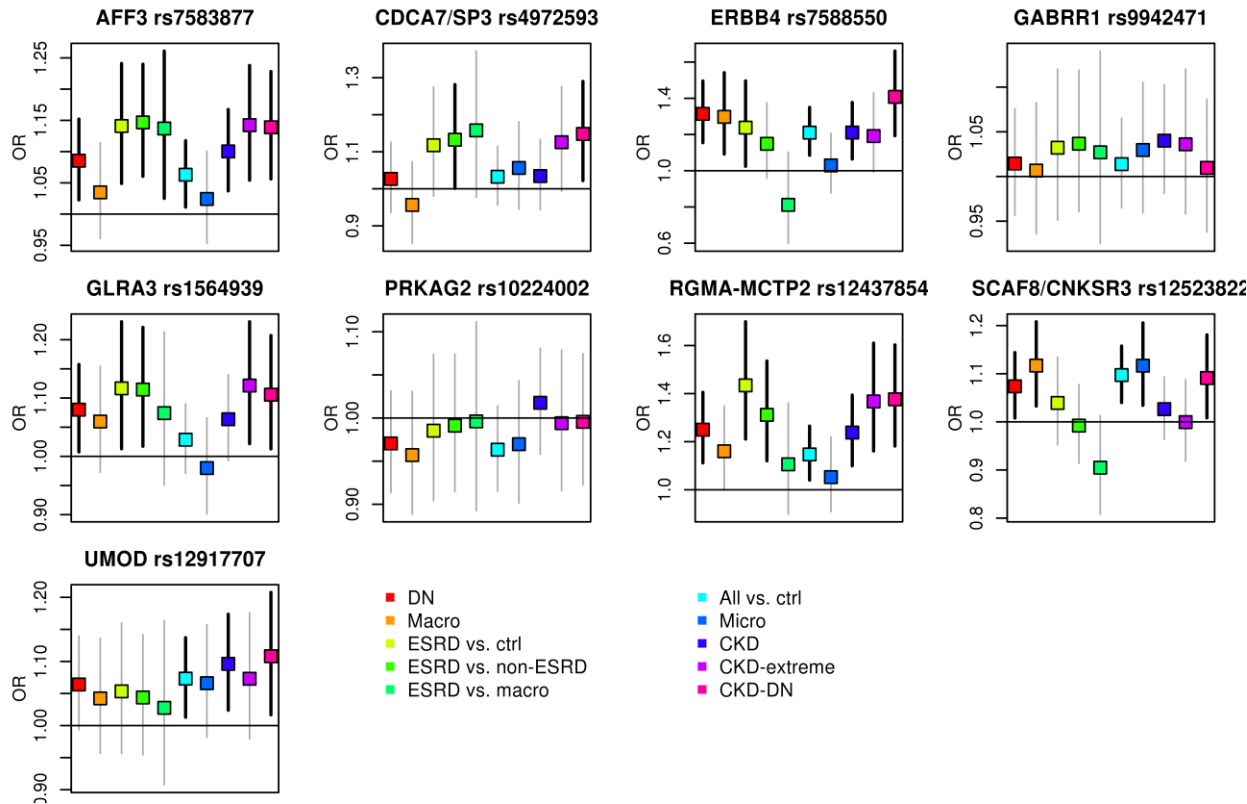
Figure S8: Forest plots of the associations at the previously reported lead loci from the GENIE consortium with largely overlapping studies. A: *RGMA-MCTP2* rs12437854. **B:** *AFF3* rs7583877. **C:** *ERBB4* rs7588550. Meta-analysis results for *RGMA-MCTP2*: Previous P = 2.0×10^{-9}, OR = 1.80 (95% confidence interval 1.48, 2.17), Current P = 7.4×10^{-4}, OR = 1.31 (1.12, 1.54); Meta-analysis results for *AFF3*: Previous p=1.20×10^{-8}, OR = 1.29 (1.18, 1.40), Current p=5.97x10-4, OR=1.15 (1.06, 1.24). Meta-analysis results for *ERBB4*: Previous P = 2.1×10^{-7}, OR = 0.66 (0.56, 0.77), Current P = 3.5×10^{-5}, OR = 0.76 (0.67, 0.87).
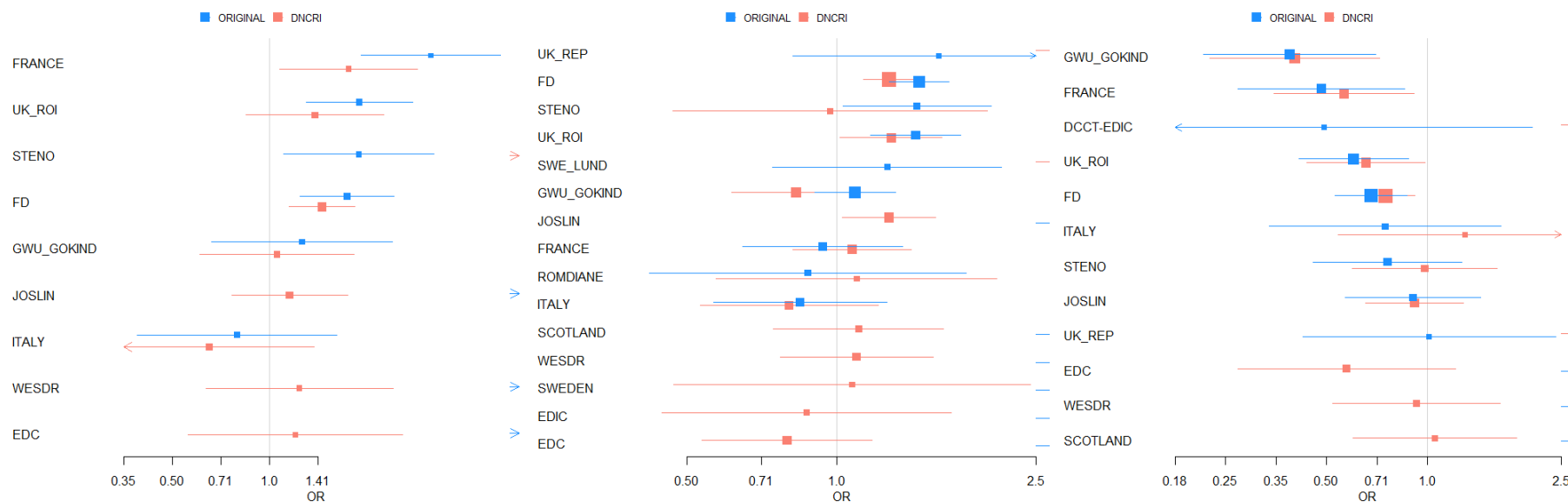
Figure S9: Meta-analysis results for the loci that have previously been associated with DKD, or with eGFR or AER in the general population. Figure shows OR [95% CI] for the 25 loci with $p<0.05$ for at least one sub-phenotype; associations with $p<0.05$ are indicated with black confidence intervals. Results are plotted so that odds ratio (OR)>1 indicates association in the same direction with the original report (for eGFR, this means that the allele associated with higher risk of DN is associated with lower eGFR). A total of 69 loci were evaluated, including loci for DKD (5 loci: *AFF3*, *RGMA-MCTP2*, *ERBB4* (Sandholm 2012), *CDCA7/SP3* (Sandholm2014), *SCAF8/CNKSR3* (Iyengar 2015)), for albuminuria in individuals with diabetes (*GLRA3* (Sandholm 2013), 3 suggestive loci *CUBN, HST6ST1* and *RAB38* (Teumer 2016)), for eGFR in individuals with diabetes (*UMOD*, Pattaro et al. 2016 and Van Zuydam et al. 2018, *PRKAG2* Van Zuydam et al. 2018) or without diabetes (61 loci, Gorski 2017). Associations at *AFF3, RGMA-MCTP2, ERBB4, SCAF8/CNKSR3*, and *UMOD* remained significant after correction for 69 tested loci ($p<7\times10^{-4}$).

Forest plots of odds ratios (OR) across phenotypes for multiple loci: AC005534.6 rs6459680, AFF3 rs7583877, ARL15 rs111366116, CACNA1S rs3850625, CDCA7/SP3 rs4972593, DPEP1 rs164748, ERBB4 rs7588550, GLRA3 rs1564939, INO80 rs2928148, LPHN2 rs10874312, NAT8 rs4500972, PIK3R1 rs113246091, RERG rs7956634, RGMA-MCTP2 rs12437854, SCAF8/CNKSR3 rs12523822, SDCCAG8 rs2802729, SLC22A2 rs2279463, SLC6A13 rs11062167, SYPL2 rs10127790, TFDP2 rs7640665, TSPAN9 rs67551338, UMOD rs12917707, UMOD rs77924615, VEGFA rs1317983, WDR37 rs80282103.

Legend:
- DN
- Macro
- ESRD vs. ctrl
- ESRD vs. non-ESRD
- ESRD vs. macro
- All vs. ctrl
- Micro
- CKD
- CKD-extreme
- CKD-DN

Figure S10. Expression of quantitative trait loci (eQTL) analysis in microdissected tubule samples. Boxplots showing normalized gene expression by stratified by homozygous common (red), heterozygous (green), and homozygous rare (blue) genotype. We identified nominal associations for rs55703767 in tubule samples with *IRS1* (a) and in glomerular samples with *RP11-395N3.2* and *AGFG1* (b). We also found nominal associations of rs61983410 with the gene encoding Cathepsin G, *CTSG*, in both eQTL analysis of whole kidney samples (c) and microdissected tubule samples (d).
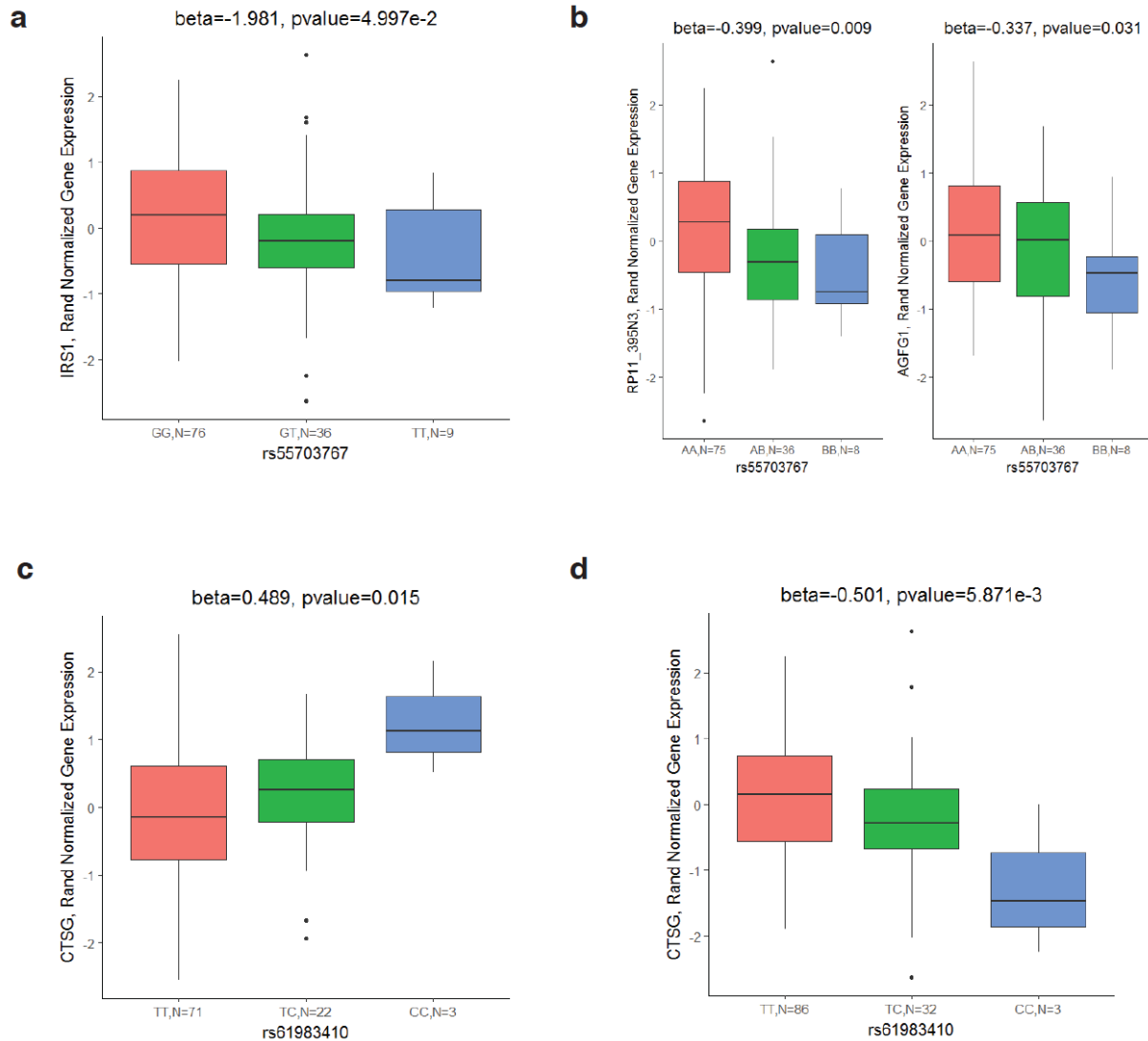
Figure S11. Functional annotation of *TAMM41*. ChIP-seq data derived from healthy adult human kidney samples (Bernstein et al., 2010) shows that variants associated with microalbuminuria are located close to H3K27ac, H3K9ac, H3K4me1, and H3K4me3 signals, suggesting that this locus is an active regulator of *TAMM41*.
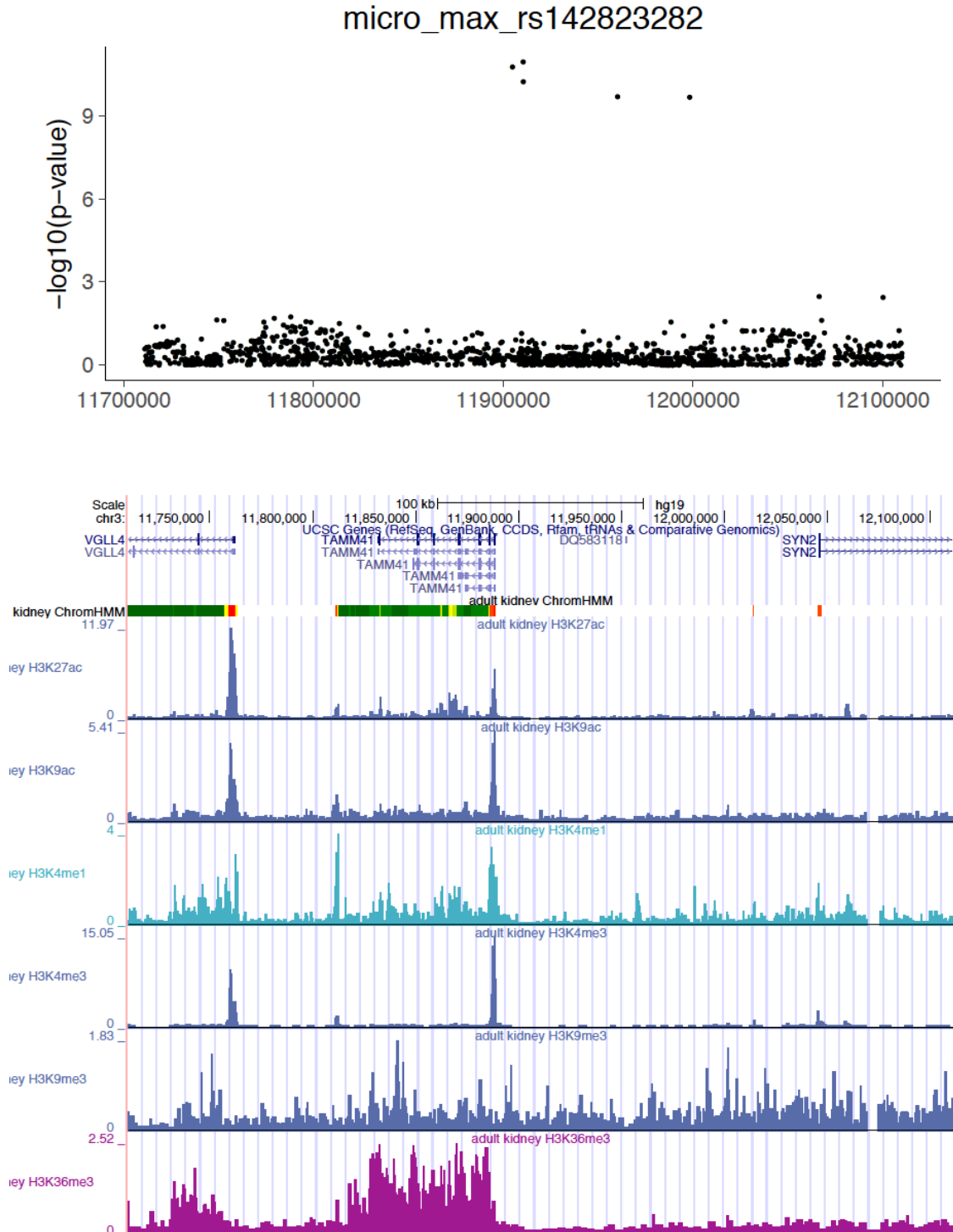
Supplemental Methods: Cohort descriptions

CACTI: The Coronary Artery Calcification in Type 1 Diabetes (CACTI) study enrolled 656 subjects with diabetes diagnosed before age 30 years, treated with insulin within 1 year of diagnosis, and diabetes duration of at least 10 years on enrollment.[1]

DCCT/EDIC: The Diabetes Control and Complications Trial (DCCT) was a multi-center randomized clinical trial to compare intensive and conventional insulin therapy on the development and progression of early vascular and neurological complications of type 1 diabetes (T1D). Renal outcomes were defined as time in years from DCCT baseline until the event. AERs were measured annually in DCCT and every other year in the post-study Epidemiology of Diabetes Interventions and Complications (EDIC) cohort. Persistent microalbuminuria was defined as the time to two consecutive AER >30 mg/24 hours (>20.8 µg/min); severe nephropathy was the time to AER >300 mg/24 hours (>208 µg/min) with prior persistent microalbuminuria, or ESRD. 22% developed persistent microalbuminuria during follow-up (268 events, 976 censored), while 10% developed severe nephropathy (132 events, 1,172 censored).[2, 35]

EDC: The Pittsburgh Epidemiology of Diabetes Complications (EDC) is a historical cohort study based on incident cases of childhood onset (prior to age 17 years) T1D, diagnosed or seen within one year of diagnosis (1950-80) at Children's Hospital of Pittsburgh.[4] The cohort, which has been shown to be epidemiologically representative of the Allegheny County, Pennsylvania, T1D population,[36] was first assessed for the EDC study between 1986 and 1988 (mean participant age and diabetes duration were 28 and 19 years, respectively). Subsequently, biennial examinations were conducted for 10 years, with a further detailed examination at 18 and 25 years from enrollment. All EDC study participants provided informed consent, and all study procedures were approved by the University of Pittsburgh Institutional Review Board (IRB). Microalbuminuria was defined as albumin excretion rate (AER) 20-200 µg/min (30-300 mg/24 hours), overt nephropathy as AER >200 µg/min (>300 mg/24 hours) and albuminuria as >20 µg/min (>30 mg/24 hours) in at least two of three validated timed urine collections. End-stage renal disease was defined as receiving dialysis or renal transplantation.

FinnDiane: Finnish Diabetic Nephropathy Study (FinnDiane) is an ongoing nationwide Finnish multicenter study of adult participants with T1D described previously.[5, 6] The participants were invited to the study by their attending physician who filled a questionnaire on the medical status of the patient and performed a clinical examination. A subset of the patients participated at one or more follow-up visits with a similar setting. Additional health related information was obtained from Finnish hospital discharge registry and from the patients' medical records. Further patients were included to the FinnDiane study through collaboration with the Finnish National Institute for Health and Welfare; for these participants, health related data was obtained from the hospital discharge registry and from the medical records. For this study, participants were limited to those with T1D diagnosed prior to age 40 years and with insulin treatment begun within 2 calendar years from diabetes onset. Disease status was defined by urine albumin excretion rate (AER) or urine albumin to creatinine ratio (ACR) in at least two out of three consecutive urine collections at local centers: microalbuminuria was defined as AER 20-200 µg/min or 30-300 mg/24h or an ACR of 2.5-25 mg/mmol for men and 3.5-35 mg/mmol for women in overnight, 24-hour or spot urine collections, respectively. Similarly, the limit for macroalbuminuria was AER >200 µg/min or >300 mg/24h or ACR > 25 mg/mmol for men and >35 mg/mmol for women. ESRD was defined as ongoing dialysis treatment or receipt of transplanted kidney. Control patients with normal AER were required to have T1D duration of at least 15 years.[5, 6]

France-Belgium: The GENEDIAB ('Génétique de la Néphropathie Diabétique, Genetics of Diabetic Nephropathy) and Genesis subjects were recruited in France, and in France-Belgium, respectively. Patients with T1D were selected on the following criteria: 1) age at diabetes onset before age 35 years, and 2) definitive insulin use within one year after diagnosis. Diabetic nephropathy was classified according to the highest three AER measurements within the last 5 years. Categories included: 1) controls (normoalbuminuria), 2) incipient nephropathy (microalbuminuria), 3) established nephropathy (proteinuria), and 4) advanced nephropathy (serum creatinine >150 mol/L and/or renal replacement therapy).[7, 8]

GoKinD US: Genetics of Kidneys in Diabetes US Study (GoKinD): The GoKinD study consists of a DKD case-control cohort of individuals diagnosed with T1D prior to 31 years of age who began insulin treatment within 1 year of T1D diagnosis. Controls were 18-59 years of age, with T1D for at least 15 years but without DKD. DKD definition includes individuals with end-state renal disease (ESRD), dialysis or kidney transplant and persistent macroalbuminuria (at least 2 out of 3 tests positive for albuminuria by dipstick ≥1+, or ACR >300 µg albumin/mg of urine creatinine). Cases were defined as people 18-54 years of age, with T1D for at least 10 years and DKD. Individuals were recruited at two study centers, George Washington University and the Joslin Diabetes Center using differing methods.[9] The Joslin GoKinD subjects were analyzed jointly with subjects from the Joslin Microalbuminuria and 50-years medalists (see below).

The InterDiane Consortium: The International Diabetic Nephropathy Consortium (InterDiane) was initiated in 2010 based on the protocol of the FinnDiane Study. The aim of the study is to identify risk factors for diabetic nephropathy and other chronic complications in patients with T1D. The participating studies follow the main protocol of the FinnDiane Study and use the same standardized questionnaires for data acquisition. T1D was defined as diabetes onset <40 years with insulin treatment initiated within one year of diagnosis. The main renal phenotype information has been collected at a baseline visit but in some countries prospective patient visits have been performed and additional phenotype information has been gathered. The last available phenotype information has been used in the analyses. Patients included fulfil the harmonized case and control criteria of the present study. InterDiane centers included in this study come from Romania, Austria, Latvia and Lithuania.

- **AusDiane: The Austrian Diabetic Nephropathy Study (AusDiane)** was initiated in 2012 in the state of Salzburg in Austria, and is part of the InterDiane Consortium (please see also the InterDiane cohort description). The patients have been studied during a regular visit at two hospitals (Department of Internal Medicine 1, Paracelsus Medical University Hospital Salzburg and Diakonissen-Wehrle Hospital Salzburg). Recruitment was done consecutively in the outpatient departments of these two hospitals. Clinical data were collected mainly as part of

the Type 1 diabetes Registry of the state of Salzburg. Patients have been studied repeatedly every 1 to 1.5 years to improve the phenotype. The last available phenotype is used for the analysis. This study comprises 71 patients with normal AER and diabetes duration ≥15 years, 13 with microalbuminuria, 4 with macroalbuminuria and 2 with ESRD and with GWAS data available and passing the inclusion criteria. Renal status was assessed by morning urine samples at least once every year. The study received ethical approval from the local ethics committee (Ethikkommission Salzburg). Written consent was obtained prior to participation in the study.

- **The Latvian Diabetic Nephropathy Study (LatDiane)** was initiated in 2012 and is part of the InterDiane Consortium (please see also the InterDiane cohort description). Recruitment of patients took place in Pauls Stradins University Hospital (Riga). The patients were recruited from the Endocrinology department of Pauls Stradins University Hospital and from out-patient clinics of Riga and Riga district (cities Jelgava, Jurmala, Ogre, Salaspils ect). The study comprises 80 patients with normal AER and diabetes duration ≥15 years, 33 with microalbuminuria, 18 with macroalbuminuria and 7 with ESRD and with GWAS data available and passing the inclusion criteria. Patients from out-patients clinics of Riga and Riga district were invited for a separate recruitment visit following the invitation of their endocrinologist. Patients undergoing treatment or correction of therapy in Endocrinology department of Pauls Stradins University Hospital were recruited in the department. Renal status was assessed based on available data of albuminuria (albumin content in 24-hour urine or albumin/creatinine in morning spot urine). In addition, during the recruitment visit, morning spot urine was collected from all patients, and sent for albumin/creatinine measurement. For patients without available data on measurements of albuminuria before recruitment to the LatDiane Study, albumin/creatinine determination in morning spot urine was repeated also several weeks after recruitment. Follow-up visits are planned for 2018. The study received ethical approval from the Latvian Central Ethics Committee. Written consent was obtained prior to participation in the study.[11]

84

- **The Lithuanian Diabetic Nephropathy Study (LitDiane)** was initiated in 2013 and is part of the InterDiane Consortium (please see also the InterDiane cohort description). Patients with T1D have been collected in a single center at the Hospital of Lithuanian University of Health Sciences (HLUHS) in Kaunas. Patients were included in the study from out-patient and inpatient departments of Endocrinology clinic of HLUHS during separate study visit. Medical records were reviewed for each patient and prospective visits are performed once a year. Renal status was classified based on the urinary albumin excretion rate (AER) in at least two out of three consecutive urine collections as: normal AER (<30mg/24h in a 24-hour urine collection), incipient diabetic nephropathy (microalbuminuria; AER ≥30 and <300mg/24h) or overt diabetic nephropathy (macroalbuminuria; AER ≥300mg/24h). Patients on dialysis or with a kidney transplant were considered to have end-stage renal disease (ESRD). As a measure of renal function estimated GFR (eGFR) was calculated with the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) formula. At the time of analysis, the study comprised 39 patients with normal AER, 32 with microalbuminuria, 9 with macroalbuminuria and 10 with ESRD and with GWAS data available and passing the inclusion criteria. The study received ethical approval from the Kaunas Regional Biomedical Research Ethics Committee (No. BE-2-16, 13-March-2013). Written consent was obtained prior to participation in the study.

- **The Romanian Diabetic Nephropathy Study (RomDiane)** was initiated in 2010 in Romania as the pilot study of the InterDiane Consortium. Patients have been studied in a cross-sectional manner in two centers in Bucharest and one in Craiova between 2010 and 2012. Renal status was assessed based on the AER or ACR in two out of three consecutive urine collections at local centers. This study comprises 89 patients with normal AER and diabetes duration ≥15 years, 48 with microalbuminuria, 70 with macroalbuminuria and 28 with ESRD, and with GWAS data available and passing the inclusion criteria. The study received ethical approval from the local ethics committee. Written consent was obtained prior to participation in the study.[12]

Italy: Subjects with T1D were recruited at the Complications of Diabetes Unit of the San Raffaele Scientific Institute, Milan, Italy. Diabetic nephropathy was defined as a median AER >200 μg min-1 in three overnight collections of sterile urine in patients with T1D for at least 10 years, concomitant diabetic retinopathy and absence of clinical or laboratory evidence of cardiac failure or other renal or urinary tract disease. Patients without nephropathy had a median AER <20 μg/min.[5]

Joslin Cohort: There were 2,271 Joslin patients with T1D included in this study. These patients were derived from three cohorts included in the ongoing Joslin Kidney Study.[10] Recruitment of 1,600 patients into the 1st Joslin Kidney Study on Natural History of Microalbuminuria in T1D took place between 1991 and 1993, and the cohort was followed through 2004. Recruitment of 1,108 patients into the 2nd Joslin Kidney Study on Natural History of Early Renal Decline in T1D took place between 2003 and 2012 and the follow-up of this cohort is still ongoing. The Joslin Proteinuria Cohort that included 630 patients was assembled from among those who developed proteinuria while attending the Joslin Clinic between 1991 and 2004. The follow-up of this cohort is still ongoing. In the analysis of data for this study, the kidney phenotypes of patients at the enrollment into the Joslin Kidney Study were considered. Genotyping data were available for 244 patients with ESRD, 475 patients with proteinuria, 470 patients with microalbuminuria and 1,189 patients with normoalbuminuria.

SDRNT1BIO: The Scottish Diabetes Research Network Type 1 Bioresource is a prospective cohort study of 6,127 individuals from across Scotland. Participants aged 16 years and over with a clinical diagnosis of T1D and insulin use within a year of onset were recruited from primary and secondary care across Scotland between 2010 and 2013. Serum, plasma, whole blood and urine samples were collected at study day allowing eGFR and albuminuria status to be obtained. Further retrospective and prospective biochemistry, co-morbidity and lifestyle data were linked from routine electronic health care records, providing serial estimates of renal status.[13]

Steno: Patients with T1D attending the outpatient clinic at Steno Diabetes Center were invited to participate in a study of genetic risk factors for diabetes complications. T1D was

considered present if the age at onset of diabetes was ≤35 years and time to definite insulin therapy ≤1 year. DKD was defined by persistent albuminuria (>300 mg/24 h) in two out of three consecutive measurements, presence of retinopathy, and absence of other kidney or urinary tract disease. Absence of DKD (controls) was defined as persistent normoalbuminuria (<30 mg/24 h) after more than 15 years of T1D in patients not treated with ACE inhibitors or angiotensin-II receptor blockers. ESRD was defined as chronic dialysis or kidney transplantation.[15]

Sweden: All patients with T1D were Swedish and diagnosed before 30 years of age. The patients with macroalbuminuria (urinary AER ≥ 200 µg min-1 in at least two consecutive overnight samples) were defined as case. The patients with AER <20 µg min-1 were considered as control.[16]

UK-ROI: In the United Kingdom (UK) GoKinD, Warren 3 and All Ireland (UK-ROI) study, data were collected under a parallel protocol to that of the GoKinD study in the United States (see above). Briefly, all individuals are white with parents and grandparents born in the UK or Ireland and who had T1D diagnosed before 31 years of age. Cases have DKD diagnosed by the onset of proteinuria (>0.5 g/24 hr) >10 years since diagnosis of diabetes; controls are diabetic individuals without evidence of proteinuria (or microalbuminuria) >15 years after onset of diabetes.[18]

WESDR: The Wisconsin Epidemiologic Study of Diabetic Retinopathy was an epidemiologic study of subjects with diabetes diagnosed before 30 years of age and taking insulin. Outcomes collected included proteinuria on a urine dipstick test.[19]

References

1.   Dabelea, D., G. Kinney, J.K. Snell-Bergeon, J.E. Hokanson, R.H. Eckel, J. Ehrlich, S. Garg, R.F. Hamman, M. Rewers, and S. Coronary Artery Calcification in Type 1 Diabetes, *Effect of type 1 diabetes on the gender difference in coronary artery calcification: a role for insulin resistance? The Coronary Artery Calcification in Type 1 Diabetes (CACTI) Study.* Diabetes, 2003. 52(11): p. 2833-9.

2.   Nathan, D.M. and D.E.R. Group, *The diabetes control and complications trial/epidemiology of diabetes interventions and complications study at 30 years: overview.* Diabetes Care, 2014. 37(1): p. 9-16. PMCIDPMC3867999

3.   The Diabetes Control and Complications Trial Research Group, *Implementation of treatment protocols in the Diabetes Control and Complications Trial.* Diabetes Care, 1995. 18(3): p. 361-76.

4.   Orchard, T.J., J.S. Dorman, R.E. Maser, D.J. Becker, A.L. Drash, D. Ellis, R.E. LaPorte, and L.H. Kuller, *Prevalence of complications in IDDM by sex and duration. Pittsburgh Epidemiology of Diabetes Complications Study II.* Diabetes, 1990. 39(9): p. 1116-24.

5.   Sandholm, N., R.M. Salem, A.J. McKnight, E.P. Brennan, C. Forsblom, T. Isakova, G.J. McKay, W.W. Williams, D.M. Sadlier, V.P. Makinen, E.J. Swan, C. Palmer, A.P. Boright, E. Ahlqvist, H.A. Deshmukh, B.J. Keller, H. Huang, A.J. Ahola, E. Fagerholm, D. Gordin, V. Harjutsalo, B. He, O. Heikkila, K. Hietala, J. Kyto, P. Lahermo, M. Lehto, R. Lithovius, A.M. Osterholm, M. Parkkonen, J. Pitkaniemi, M. Rosengard-Barlund, M. Saraheimo, C. Sarti, J. Soderlund, A. Soro-Paavonen, A. Syreeni, L.M. Thorn, H. Tikkanen, N. Tolonen, K. Tryggvason, J. Tuomilehto, J. Waden, G.V. Gill, S. Prior, C. Guiducci, D.B. Mirel, A. Taylor, S.M. Hosseini, D.E.R. Group, H.H. Parving, P. Rossing, L. Tarnow, C. Ladenvall, F. Alhenc-Gelas, P. Lefebvre, V. Rigalleau, R. Roussel, D.A. Tregouet, A. Maestroni, S. Maestroni, H. Falhammar, T. Gu, A. Mollsten, D. Cimponeriu, M. Ioana, M. Mota, E. Mota, C. Serafinceanu, M. Stavarachi, R.L. Hanson, R.G. Nelson, M. Kretzler, H.M. Colhoun, N.M. Panduru, H.F. Gu, K. Brismar, G. Zerbini, S. Hadjadj, M. Marre, L. Groop, M. Lajer, S.B. Bull, D. Waggott, A.D. Paterson, D.A. Savage, S.C. Bain, F. Martin, J.N. Hirschhorn, C. Godson, J.C. Florez, P.H. Groop, and A.P. Maxwell, *New susceptibility loci associated with kidney disease in type 1 diabetes.* PLoS Genet, 2012. 8(9): p. e1002921. PMCID3447939

6.   Thorn, L.M., C. Forsblom, J. Fagerudd, M.C. Thomas, K. Pettersson-Fernholm, M. Saraheimo, J. Waden, M. Ronnback, M. Rosengard-Barlund, C.-G.a. Bjorkesten, M.-R. Taskinen, P.-H. Groop, and on behalf of the FinnDiane Study Group, *Metabolic syndrome in type 1 diabetes: Association with diabetic nephropathy and glycemic control (the FinnDiane study).* Diabetes Care, 2005. 28(8): p. 2019-2024.

7.   Marre, M., X. Jeunemaitre, Y. Gallois, M. Rodier, G. Chatellier, C. Sert, L. Dusselier, Z. Kahal, L. Chaillous, S. Halimi, A. Muller, H. Sackmann, B. Bauduceau, F. Bled, P. Passa, and F. Alhenc-Gelas, *Contribution of genetic polymorphism in the renin-angiotensin system to the development of renal complications in insulin-dependent diabetes: Genetique de la Nephropathie Diabetique (GENEDIAB) study group.* J Clin Invest, 1997. 99(7): p. 1585-95. PMCIDPMC507978

8.   Hadjadj, S., F. Pean, Y. Gallois, P. Passa, R. Aubert, L. Weekers, V. Rigalleau, B. Bauduceau, A. Bekherraz, R. Roussel, B. Dussol, M. Rodier, R. Marechaud, P.J. Lefebvre, M. Marre, and S. Genesis France-Belgium, *Different patterns of insulin resistance in relatives of type 1 diabetic patients with retinopathy or nephropathy: the Genesis France-Belgium Study.* Diabetes Care, 2004. 27(11): p. 2661-8.

9.   Pezzolesi, M.G., G.D. Poznik, J.C. Mychaleckyj, A.D. Paterson, M.T. Barati, J.B. Klein, D.P.K. Ng, G. Placha, L.H. Canani, J. Bochenski, D. Waggott, M.L. Merchant, B. Krolewski, L. Mirea, K. Wanic, P. Katavetin, M. Kure, P. Wolkow, J.S. Dunn, A. Smiles, W.H. Walker, A.P. Boright, S.B. Bull, t.D.E.R. Group, A. Doria, J.J. Rogus, S.S. Rich, J.H. Warram, and A.S. Krolewski, *Genome-Wide Association*

Scan for Diabetic Nephropathy Susceptibility Genes in Type 1 Diabetes. Diabetes, 2009. 58(6): p. 1403-1410.

10.   Krolewski, A.S., *Progressive renal decline: the new paradigm of diabetic nephropathy in type 1 diabetes.* Diabetes Care, 2015. 38(6): p. 954-62. PMCIDPMC4439536

11.   Sviklane, L., E. Olmane, Z. Dzerve, K. Kupcs, V. Pirags, and J. Sokolovska, *Fatty liver index and hepatic steatosis index for prediction of non-alcoholic fatty liver disease in type 1 diabetes.* J Gastroenterol Hepatol, 2018. 33(1): p. 270-276.

12.   Pop, A., D. Clenciu, M. Anghel, S. Radu, B. Socea, E. Mota, M. Mota, N.M. Panduru, and G. RomDiane Study, *Insulin resistance is associated with all chronic complications in type 1 diabetes.* J Diabetes, 2016. 8(2): p. 220-8.

13.   Akbar, T., S. McGurnaghan, C.N.A. Palmer, S.J. Livingstone, J. Petrie, J. Chalmers, R.S. Lindsay, J.A. McKnight, D.W.M. Pearson, A.W. Patrick, J. Walker, H.C. Looker, and H.M. Colhoun, *Cohort Profile: Scottish Diabetes Research Network Type 1 Bioresource Study (SDRNT1BIO).* Int J Epidemiol, 2017. 46(3): p. 796-796i. PMCIDPMC5582633

14.   Afkarian, M., I.B. Hirsch, K.R. Tuttle, C. Greenbaum, J. Himmelfarb, and I.H. de Boer, *Urinary excretion of RAS, BMP, and WNT pathway components in diabetic kidney disease.* Physiol Rep, 2014. 2(5): p. e12010. PMCIDPMC4098738

15.   Rossing, P., P. Hougaard, and H.H. Parving, *Risk factors for development of incipient and overt diabetic nephropathy in type 1 diabetic patients: a 10-year prospective observational study.* Diabetes Care, 2002. 25(5): p. 859-64.

16.   Ma, J., A. Mollsten, M. Prazny, H. Falhammar, K. Brismar, G. Dahlquist, S. Efendic, and H.F. Gu, *Genetic influences of the intercellular adhesion molecule 1 (ICAM-1) gene polymorphisms in development of Type 1 diabetes and diabetic nephropathy.* Diabet Med, 2006. 23(10): p. 1093-9. PMCIDPMC1618804

17.   Mollsten, A., I. Kockum, M. Svensson, S. Rudberg, A. Ugarph-Morawski, K. Brismar, J.W. Eriksson, and G. Dahlquist, *The effect of polymorphisms in the renin-angiotensin-aldosterone system on diabetic nephropathy risk.* J Diabetes Complications, 2008. 22(6): p. 377-83.

18.   McKnight, A.J., C.C. Patterson, N. Sandholm, J. Kilner, T.A. Buckham, M. Parkkonen, C. Forsblom, D.M. Sadlier, P.H. Groop, A.P. Maxwell, and U.K.G.S.G. Warren, *Genetic polymorphisms in nitric oxide synthase 3 gene and implications for kidney disease: a meta-analysis.* Am J Nephrol, 2010. 32(5): p. 476-81.

19.   Klein, R., B.E. Klein, S.E. Moss, and K.J. Cruickshanks, *The Wisconsin Epidemiologic Study of Diabetic Retinopathy: XVII. The 14-year incidence and progression of diabetic retinopathy and associated risk factors in type 1 diabetes.* Ophthalmology, 1998. 105(10): p. 1801-15.

20.   Todd, J.N., E.H. Dahlstrom, R.M. Salem, N. Sandholm, C. Forsblom, G. FinnDiane Study, A.J. McKnight, A.P. Maxwell, E. Brennan, D. Sadlier, C. Godson, P.H. Groop, J.N. Hirschhorn, and J.C. Florez, *Genetic Evidence for a Causal Role of Obesity in Diabetic Kidney Disease.* Diabetes, 2015. 64(12): p. 4238-46. PMCIDPMC4657582

21.   Li, J. and L. Ji, *Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix.* Heredity (Edinb), 2005. 95(3): p. 221-7.

22.   Mauer, M., B. Zinman, R. Gardiner, S. Suissa, A. Sinaiko, T. Strand, K. Drummond, S. Donnelly, P. Goodyer, M.C. Gubler, and R. Klein, *Renal and retinal effects of enalapril and losartan in type 1 diabetes.* N Engl J Med, 2009. 361(1): p. 40-51. PMCIDPMC2978030

23.   van Zuydam, N.R., E. Ahlqvist, N. Sandholm, H. Deshmukh, N.W. Rayner, M. Abdalla, C. Ladenvall, D. Ziemek, E. Fauman, N.R. Robertson, P.M. McKeigue, E. Valo, C. Forsblom, V. Harjutsalo, S. Finnish Diabetic Nephropathy, A. Perna, E. Rurali, M.L. Marcovecchio, R.P. Igo, Jr., R.M. Salem, N. Perico, M. Lajer, A. Karajamaki, M. Imamura, M. Kubo, A. Takahashi, X. Sim, J. Liu, R.M. van Dam, G. Jiang, C.H.T. Tam, A.O.Y. Luk, H.M. Lee, C.K.P. Lim, C.C. Szeto, W.Y. So, J.C.N.

Chan, G. Hong Kong Diabetes Registry Theme-based Research Scheme Project, S.F. Ang, R. Dorajoo, L. Wang, T.S.H. Clara, A.J. McKnight, S. Duffy, Warren, G. Genetics of Kidneys in Diabetes Study, M.G. Pezzolesi, G. Consortium, M. Marre, B. Gyorgy, S. Hadjadj, L.T. Hiraki, C. Diabetes, I. Complications Trial /Epidemiology of Diabetes, G. Complications Research, T.S. Ahluwalia, P. Almgren, C.A. Schulz, M. Orho-Melander, A. Linneberg, C. Christensen, D.R. Witte, N. Grarup, I. Brandslund, O. Melander, A.D. Paterson, D. Tregouet, A.P. Maxwell, S.C. Lim, R.C.W. Ma, E.S. Tai, S. Maeda, V. Lyssenko, T. Tuomi, A.S. Krolewski, S.S. Rich, J.N. Hirschhorn, J.C. Florez, D. Dunger, O. Pedersen, T. Hansen, P. Rossing, G. Remuzzi, S.U.m.f. Micro, C. Macrovascular hard endpoints for Innovative diabetes Tools, M.J. Brosnan, C.N.A. Palmer, P.H. Groop, H.M. Colhoun, L.C. Groop, and M.I. McCarthy, *A Genome-Wide Association Study of Diabetic Kidney Disease in Subjects With Type 2 Diabetes.* Diabetes, 2018. 67(7): p. 1414-1427. PMCIDPMC6014557

24. Ko, Y.A., H. Yi, C. Qiu, S. Huang, J. Park, N. Ledo, A. Kottgen, H. Li, D.J. Rader, M.A. Pack, C.D. Brown, and K. Susztak, *Genetic-Variation-Driven Gene-Expression Changes Highlight Genes with Important Functions for Kidney Disease.* Am J Hum Genet, 2017. 100(6): p. 940-953. PMCIDPMC5473735

25. Cohen, C.D., K. Frach, D. Schlondorff, and M. Kretzler, *Quantitative gene expression analysis in renal biopsies: a novel protocol for a high-throughput multicenter application.* Kidney Int, 2002. 61(1): p. 133-40.

26. Berthier, C.C., H. Zhang, M. Schin, A. Henger, R.G. Nelson, B. Yee, A. Boucherot, M.A. Neusser, C.D. Cohen, C. Carter-Su, L.S. Argetsinger, M.P. Rastaldi, F.C. Brosius, and M. Kretzler, *Enhanced expression of Janus kinase-signal transducer and activator of transcription pathway members in human diabetic nephropathy.* Diabetes, 2009. 58(2): p. 469-77. PMCIDPMC2628622

27. Schmid, H., A. Boucherot, Y. Yasuda, A. Henger, B. Brunner, F. Eichinger, A. Nitsche, E. Kiss, M. Bleich, H.J. Grone, P.J. Nelson, D. Schlondorff, C.D. Cohen, M. Kretzler, and D.N.A.B.C. European Renal c, *Modular activation of nuclear factor-kappaB transcriptional programs in human diabetic nephropathy.* Diabetes, 2006. 55(11): p. 2993-3003.

28. Irizarry, R.A., B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, and T.P. Speed, *Exploration, normalization, and summaries of high density oligonucleotide array probe level data.* Biostatistics, 2003. 4(2): p. 249-64.

29. Johnson, W.E., C. Li, and A. Rabinovic, *Adjusting batch effects in microarray expression data using empirical Bayes methods.* Biostatistics, 2007. 8(1): p. 118-27.

30. Parkin, J.D., J.D. San Antonio, V. Pedchenko, B. Hudson, S.T. Jensen, and J. Savige, *Mapping structural landmarks, ligand binding sites, and missense mutations to the collagen IV heterotrimers predicts major functional domains, novel interactions, and variation in phenotypes in inherited diseases affecting basement membranes.* Hum Mutat, 2011. 32(2): p. 127-43. PMCIDPMC4800984

31. Park, J., R. Shrestha, C. Qiu, A. Kondo, S. Huang, M. Werth, M. Li, J. Barasch, and K. Susztak, *Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease.* Science, 2018. 360(6390): p. 758-763. PMCIDPMC6188645

32. Bernstein, B.E., J.A. Stamatoyannopoulos, J.F. Costello, B. Ren, A. Milosavljevic, A. Meissner, M. Kellis, M.A. Marra, A.L. Beaudet, J.R. Ecker, P.J. Farnham, M. Hirst, E.S. Lander, T.S. Mikkelsen, and J.A. Thomson, *The NIH Roadmap Epigenomics Mapping Consortium.* Nat Biotechnol, 2010. 28(10): p. 1045-8. PMCIDPMC3607281

33. Gorski, M., P.J. van der Most, A. Teumer, A.Y. Chu, M. Li, V. Mijatovic, I.M. Nolte, M. Cocca, D. Taliun, F. Gomez, Y. Li, B. Tayo, A. Tin, M.F. Feitosa, T. Aspelund, J. Attia, R. Biffar, M. Bochud, E. Boerwinkle, I. Borecki, E.P. Bottinger, M.H. Chen, V. Chouraki, M. Ciullo, J. Coresh, M.C. Cornelis, G.C. Curhan, A.P. d'Adamo, A. Dehghan, L. Dengler, J. Ding, G. Eiriksdottir, K. Endlich, S. Enroth,

T. Esko, O.H. Franco, P. Gasparini, C. Gieger, G. Girotto, O. Gottesman, V. Gudnason, U. Gyllensten, S.J. Hancock, T.B. Harris, C. Helmer, S. Hollerer, E. Hofer, A. Hofman, E.G. Holliday, G. Homuth, F.B. Hu, C. Huth, N. Hutri-Kahonen, S.J. Hwang, M. Imboden, A. Johansson, M. Kahonen, W. Konig, H. Kramer, B.K. Kramer, A. Kumar, Z. Kutalik, J.C. Lambert, L.J. Launer, T. Lehtimaki, M. de Borst, G. Navis, M. Swertz, Y. Liu, K. Lohman, R.J.F. Loos, Y. Lu, L.P. Lyytikainen, M.A. McEvoy, C. Meisinger, T. Meitinger, A. Metspalu, M. Metzger, E. Mihailov, P. Mitchell, M. Nauck, A.J. Oldehinkel, M. Olden, B. Wjh Penninx, G. Pistis, P.P. Pramstaller, N. Probst-Hensch, O.T. Raitakari, R. Rettig, P.M. Ridker, F. Rivadeneira, A. Robino, S.E. Rosas, D. Ruderfer, D. Ruggiero, Y. Saba, C. Sala, H. Schmidt, R. Schmidt, R.J. Scott, S. Sedaghat, A.V. Smith, R. Sorice, B. Stengel, S. Stracke, K. Strauch, D. Toniolo, A.G. Uitterlinden, S. Ulivi, J.S. Viikari, U. Volker, P. Vollenweider, H. Volzke, D. Vuckovic, M. Waldenberger, J. Jin Wang, Q. Yang, D.I. Chasman, G. Tromp, H. Snieder, I.M. Heid, C.S. Fox, A. Kottgen, C. Pattaro, C.A. Boger and C. Fuchsberger, *1000 Genomes-based meta-analysis identifies 10 novel loci for kidney function.* Sci Rep, 2017. 7: p. 45040. PMCIDPMC5408227

34. McKelvey, R.D. and W. Zavoina, *A statistical model for the analysis of ordinal level dependent variables.* The Journal of Mathematical Sociology, 1975. 4(1): p. 103-120.

35. *Implementation of treatment protocols in the Diabetes Control and Complications Trial.* Diabetes Care, 1995. 18(3): p. 361-76.

36. Wagener, D.K., J.M. Sacks, R.E. LaPorte, and J.M. Macgregor, *The Pittsburgh study of insulin-dependent diabetes mellitus. Risk for diabetes among relatives of IDDM.* Diabetes, 1982. 31(2): p. 136-44.