

Mendelian pathway analysis of laboratory traits reveals distinct roles for ciliary subcompartments in common disease pathogenesis

Theodore George Drivas^{1,2}, Anastasia Lucas¹, Xinyuan Zhang¹, Marylyn DeRiggi Ritchie¹

¹ Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19194, USA

² Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

Summary: Rare monogenic disorders of the primary cilium, termed ciliopathies, are characterized by extreme presentations of otherwise-common diseases, such as diabetes, hepatic fibrosis, and kidney failure. However, despite a revolution in our understanding of the cilium's role in rare disease pathogenesis, the organelle's contribution to common disease remains largely unknown. We hypothesized that common genetic variants affecting Mendelian ciliopathy genes might also contribute to common complex diseases pathogenesis more generally. To address this question, we performed association studies of 16,875 common genetic variants across 122 well-characterized ciliary genes with 12 quantitative laboratory traits characteristic of ciliopathy syndromes in 378,213 European-ancestry individuals in the UK BioBank. We incorporated tissue-specific gene expression analysis, expression quantitative trait loci (eQTL) and Mendelian disease information into our analysis, and replicated findings in meta-analysis to increase our confidence in observed associations between ciliary genes and human phenotypes. 73 statistically-significant gene-trait associations were identified across 34 of the 122 ciliary genes that we examined (including 8 novel, replicating associations). With few exceptions, these ciliary genes were found to be widely expressed in human tissues relevant to the phenotypes being studied, and our eQTL analysis revealed strong evidence for correlation between ciliary gene expression levels and patient phenotypes. Perhaps most interestingly our analysis identified different ciliary subcompartments as being specifically associated with distinct sets of patient phenotypes, offering a number of testable hypotheses regarding the cilium's role in common complex disease. Taken together, our data demonstrate the utility of a Mendelian pathway-based approach to genomic association studies, and challenge the widely-held belief that the cilium is an organelle important mainly in development and in rare syndromic disease pathogenesis. The continued application of techniques similar to those described here to other phenotypes/Mendelian diseases is likely to yield many additional fascinating associations that will begin to integrate the fields of common and rare disease genetics, and provide insight into the pathophysiology of human diseases of immense public health burden.

Contact: theodore.drivas@gmail.com

INTRODUCTION

Found on nearly every human cell type, the primary cilium is a small, non-motile projection of the cell's apical surface that acts as an antenna for the reception and integration of signals from the extracellular environment.¹ Signaling at the cilium is mediated by cell surface receptors that are specifically trafficked to the cilium through a complex and highly regulated series of interactions.^{2,3} Receptors destined for the ciliary space must first interact with a protein subcomplex known as the BBSome, which allows receptors to dock at the base of the cilium (the basal body) and

navigate through the diffusion barrier formed by the ciliary transition zone to enter the inner ciliary space.^{4,5} Once inside, cargo is specifically trafficked along the microtubule core of the cilium through a process known as intraflagellar transport (IFT), before being ultimately returned to the extraciliary space to complete signal transduction.⁶

We now know that the interruption of any one of these critical ciliary pathways can lead to the development of devastating syndromic Mendelian disorders, collectively known as ciliopathies.⁷ The list of confirmed ciliopathies is constantly expanding, and currently includes disorders such as Alström syndrome, Bardet-Biedl syndrome, Joubert syndrome, Meckel-Gruber syndrome, Nephronophthisis, Senior-Løken syndrome, Sensenbrenner syndrome, and others.^{8–16} Much of the pathology of the ciliopathy syndromes can be attributed to the adverse effects of ciliary dysfunction on cell signaling pathways critical to embryologic development.^{7,17} Furthermore, although each ciliopathy is defined by its own unique constellation of characteristic phenotypic findings, many ciliopathies share similar or overlapping features.

Phenotypes that are often observed in individuals with ciliary disease include renal failure, hepatic fibrosis, obesity, dyslipidemia, and diabetes.^{8–10} Interestingly, these very same phenotypes are characteristic of common diseases that affect large proportions of the population. Recent evidence suggests that the development of these phenotypes in ciliopathy patients may be driven by the perturbation of ciliary processes required for signaling through common disease-relevant pathways not previously known to be dependent on the cilium for signal transduction, such as insulin, IGF-1, PDGF, and TGF β .^{2,18–28} Together, these observations raise the possibility that the same ciliary genes causative of rare Mendelian disease may also be involved in the pathogenesis of common complex diseases, such as kidney failure or dyslipidemia, in the general population.

To investigate this hypothesis, we set out to identify associations between common genetic variants in 122 well-characterized ciliary genes and 12 quantitative laboratory traits relevant to common disease in a large biobank cohort. We incorporated tissue-specific gene expression analysis, expression quantitative trait loci (eQTL), and Mendelian disease information into our analysis, and replicated findings in meta-analysis to increase our confidence in observed associations between ciliary genes and human phenotypes (Figure 1). The results of our analysis revealed 73 statistically significant associations between 34 ciliary genes and diverse laboratory traits (including 8 novel, replicating associations), and also identified distinct ciliary subcompartments associated with different disease processes and demonstrated pleiotropic effects for many ciliary genes across a number of phenotypes and organ systems. Our data challenge the widely-held belief that the cilium is an organelle important mainly in development and in rare syndromic disease pathogenesis, and establish a framework for the investigation of the common disease contributions of other Mendelian disease pathways.

RESULTS

34 ciliary gene loci are associated with diverse quantitative laboratory traits

We examined 16,875 genotyped or imputed variants within genomic loci defined by the boundaries of 122 well-defined ciliary genes in 378,213 European ancestry individuals in the UK BioBank. We tested each variant for association with each of 12 quantitative laboratory traits: Alanine aminotransferase (ALT), Alkaline Phosphatase (AlkPhos), Aspartate aminotransferase (AST), Total Cholesterol (Cholesterol), Serum Creatinine (Creatinine), Gamma glutamyltransferase (GGT), Serum Glucose (Glucose), Glycated haemoglobin (A1c), HDL cholesterol (HDL), LDL cholesterol (LDL), Triglycerides, and Urea using linear regression, adjusting our model for age, sex, and ancestry PCs 1-10. ALT, AlkPhos, AST, and GGT are all

markers of liver disease, Creatinine and Urea are markers of kidney disease, perturbations in HDL, LDL, Total Cholesterol, or Triglyceride levels are the hallmark of dyslipidemia, while elevations in Glucose or A1c levels can be diagnostic of diabetes and impaired glucose homeostasis. The characteristics of the subjects and phenotypes included in the discovery analysis are summarized in Table S1. To replicate novel findings, large scale meta-analyses of the same 12 quantitative laboratory traits were carried out, as described in the methods section, for genetic variants contained within the same 122 ciliary genes as analyzed in our discovery set. Results of our discovery analysis and replication meta-analyses are displayed in Figures 2-5.

Our discovery analysis identified 549 variants within 34 different ciliary genes reaching genome-wide significance ($p < 5e-8$) for association with at least one of the 12 phenotypes (Figure 2-5, Table S2-13). Many ciliary gene loci were significantly associated with multiple traits, a phenomenon known as pleiotropy (with one locus, containing the gene *IFT172*, significantly associated with nine different traits). 73 significant gene-trait associations were identified altogether, and of these 27 (37%) represent replication of previously-known associations from the literature, whereas 46 (63%) represent previously-unreported associations (novel findings). 8 of these 46 (17%) unreported associations were found to replicate in our meta-analyses. These results are summarized in Table S14. Of the 8 replicating novel associations, five were for kidney-related phenotypes, which is likely reflective of the fact that these phenotypes had the largest samples sizes in our meta-analysis (the meta-analysis of kidney related traits had a sample size of ~800,000 individuals, whereas no other meta-analyzed trait had a sample size greater than 370,000). All replicating novel associations and interesting known associations are discussed in more detail below.

Tissue-specific analysis demonstrates widespread expression of significant ciliary genes

For all genes with significant associations in our discovery analysis, we examined tissue-specific expression in the GTEx version 8 database²⁹ to determine if each gene was expressed in tissues relevant to the detected phenotype association (Figure 6, Table S15). As the cilium is a nearly ubiquitous organelle,³⁰ we were not surprised to find that the majority of ciliary genes examined (20 of 34, 59%) were broadly expressed in all tissues (Figure 6). Four genes, however, stood out as being poorly expressed in all or nearly all examined tissues – *CENPF*, *RP1*, *RP1L1*, and *WDPCP*. Why these four specific ciliary genes were the only ones found to be poorly expressed is not entirely clear.

Trait-significant variants are enriched for and correlate with eQTLs of ciliary genes

The majority of the 549 ciliary gene variants statistically significantly associated with laboratory traits lay within untranslated regions of the genome. To explore the possibility that these variants might be exerting their effect on laboratory traits by modulating ciliary gene expression, we performed an eQTL analysis of each of the 73 significant gene-trait pairs (Figures S1-S34). Utilizing the GTEx version 8 database of cis-eQTL variants,²⁹ we plotted the association peaks for each significant gene-trait pair in chromosomal space, overlaying the relevant eQTL information for each ciliary candidate gene using eQTPlot.^{31,32} We simultaneously generated plots illustrating enrichment of ciliary candidate gene eQTLs among trait-significant variants, and generated p-p plots (plotting, for each variant, the p-value of association with candidate gene expression (p_{eQTL}) against the p-value of association with laboratory trait level (p_{trait})) to illustrate and identify significant correlations between variants' effects on candidate gene expression and their effects on laboratory traits, as determined by linear regression and calculation of the Pearson correlation coefficient and p-value of correlation.

Interestingly, we found that for 70 of the 73 (96%) significant gene-trait pairs, the detected association peaks were significantly enriched for eQTLs of the candidate ciliary gene ($p < 7e-4$).

Furthermore, for 48 of the 70 (69%) significant gene-trait pairs that demonstrated eQTL enrichment, there was a significant correlation between p_{trait} and p_{eQTL} (Pearson correlation coefficient >0.25 , p-value of correlation $<3.5\text{e-}4$). These data indicate that the majority of association signals are being driven by genetic variants that are similarly associated with significant changes in expression of candidate ciliary genes. Together, these findings are consistent with the hypothesis that the detected laboratory trait associations are driven by changes in ciliary gene expression.

Novel Associations

Our discovery analysis identified 46 previously-unreported genome-wide significant associations, 8 of which were found to replicate in our meta-analyses at either the variant or gene level:

BBS2

We found a significant association between variants within the *BBS2* gene and HDL cholesterol levels (most significant association with rs118024138, p-value $5.73\text{e-}37$). This same locus was found to be significantly associated with HDL in our meta-analysis (Figure 2). *BBS2* encodes a BBSome-associated protein critical in the regulation of cargo transport through the ciliary space.³³ Biallelic pathogenic variants in the gene have been associated with Bardet-Biedl syndrome, characterized by retinal degeneration, kidney failure, obesity, impaired glucose handling, and, interestingly, dyslipidemia.^{8,34,35} Furthermore, *BBS2* was found to be significantly expressed in both liver and adipose tissue (Figure 6), with a majority of variants significantly associated with HDL levels also found to be significantly associated with *BBS2* expression in the GTEx database (Figure S3), lending a number of lines of supporting evidence to this novel association.

CLUAP1

A significant association was detected between variants within the *CLUAP1* gene and serum creatinine levels (most significant association for variant rs9790, p-value $2.56\text{e-}12$), with variant-level replication of this finding in our meta-analysis (Figure 3). Although *CLUAP1* exists in a gene-dense region of the genome, the peak of association was confined specifically to within the *CLUAP1* gene boundaries (Figure S8). We found *CLUAP1* to be significantly expressed in kidney tissues (Figure 6), with a majority of the variants significantly associated with creatinine also significantly associated with *CLUAP1* expression levels in the GTEx eQTL database (Figure S8). *CLUAP1* encodes the intraflagellar transport (IFT)-associated protein CLUAP1/IFT38, a component of the ciliary IFT-B complex important for anterograde transport of cargo along the ciliary axoneme. Biallelic pathogenic variants in the gene have been associated with an overlap syndrome similar to Joubert syndrome and Orofaciodigital syndrome in one individual.³⁶ While this individual was not found to have evidence of renal disease, renal disease is a common feature of Joubert syndrome patients more generally.⁹ Interestingly, this individual was also noted to be obese (a common feature of ciliopathy syndromes); the *CLUAP1* locus has been published as associated with BMI and type 2 diabetes in previous GWAS studies.^{37–40}

CSNK1D

A single variant within the *CSNK1D* gene, rs11653735, was found to be significantly associated with urea levels with a p-value of $2.08\text{e-}10$, with this same variant found to replicate in our meta-analysis (Figure 3). *CSNK1D* encodes the protein kinase CK1 δ that has been found to play a critical role in regulating ciliogenesis.⁴¹ Variants in *CSNK1D* have not been implicated in any classic ciliopathy syndrome. The rs11653735 variant is a deep intronic variant, and if/how it affects CK1 δ function is unclear. Intriguingly, the genomic region surrounding *CSNK1D* contains the gene *CCDC57*, a poorly characterized gene associated with the centriole and cilium that was not included in our study.⁴² The peak of association for urea levels spans the entirety of the *CCDC57* gene (Figure S9), with a strong correlation between variants associated with urea and those

associated with *CCDC57* expression (Figure S35), raising the possibility that this locus is associated with urea levels through perturbation of either or both of these ciliary genes, *CCDC57* and *CSNK1D*.

IQCB1

We found a number of variants within and surrounding the *IQCB1* locus to be significantly associated with both serum creatinine (a known association, most significant p-value of association in our discovery analysis 3.05e-09 for SNP rs9823335) and urea levels (a novel association, most significant p-value of association in our discovery analysis 2.56e-09 for rs75382826). Both of these associations were found to replicate at the variant and gene level in our meta-analysis (Figure 3). *IQCB1* is significantly expressed in the kidney (Figure 6), and nearly every variant in the *IQCB1* locus that was found to be significantly associated with urea/creatinine levels was also found to be significantly associated with *IQCB1* expression, with strong evidence for correlation between levels of *IQCB1* expression and creatinine/urea levels (Figure S16). *IQCB1* encodes a well-characterized protein of the same name that is a critical component of the ciliary transition zone. Rare pathogenic biallelic variants in *IQCB1* are a major cause of Senior-Løken syndrome, a disorder characterized by retinal degeneration and progressive kidney failure.¹⁴ The novel association of the *IQCB1* locus with blood urea levels corroborates the known association with creatinine/glomerular filtration rate, and adds further evidence supporting the role of *IQCB1* in kidney function more generally.

KIF17

We identified a very strong novel association between variants within the *KIF17* locus with AlkPhos levels (most significant p-value of association 4.09e-47 for rs61778523), which replicated at the gene level in our meta-analysis (Figure 4). Interestingly, we did not find *KIF17* to be significantly expressed in hepatic tissue (Figure 6), and there did not appear to be a significant correlation between variants associated with *KIF17* expression levels and AlkPhos levels (Figure S17). *KIF17* encodes a motor protein important in driving IFT-B-mediated anterograde ciliary transport, but no Mendelian disease has yet been associated with dysfunction of this gene.⁴³ Thus altogether there is minimal corroborating evidence to support *KIF17* as the causative gene of this novel association.

LUZP1

We found variants within the *LUZP1* gene to be significantly associated with both GGT levels (most significant p-value 1.41e-08 for variant rs56087807) and serum creatinine levels (most significant p-value 5.23e-11 for variant rs1208930), with both of these findings replicating at both the variant and gene level in our meta-analysis (Figures 3-4). The association between common variants near the *LUZP1* locus and GGT levels has been reported,⁴⁴ but the association with creatinine is novel. *LUZP1* was found to be significantly expressed in all tissues relevant to these phenotypes (Figure 6), and nearly every variant significantly associated with creatinine/GGT levels was also found to be significantly associated with *LUZP1* expression, with evidence for correlation between levels of *LUZP1* expression and creatinine/GGT levels (Figure S18). *LUZP1* encodes a ciliary basal body protein that has recently been shown to play important roles in regulating ciliogenesis.^{45,46} While there is no Mendelian disease yet known to result from disruption of the *LUZP1* gene, recent work has shown reduced *LUZP1* expression in fibroblasts derived from Townes-Brockes syndrome (characterized by renal disease, among other findings) patients,⁴⁷ suggesting that *LUZP1* may play a role in the pathogenesis of this disorder.

PIBF1

A significant, novel association was detected between variants within the *PIBF1* gene and serum creatinine levels (most significant p-value 1.08e-09 for variant rs111440455), with this finding

replicating at both the variant and gene level in our meta-analysis (Figure 3). We found *PIBF1* to be significantly expressed in the kidney (Figure 6), with nearly every variant significantly associated with creatinine levels also found to be significantly associated with *PIBF1* expression, with strong evidence for correlation between levels of *PIBF1* expression and creatinine levels (Figure S21). *PIBF1* encodes the protein PIBF1/CEP90, a component of the basal body essential for ciliogenesis and has been implicated as a cause of Joubert syndrome with kidney disease.^{48–50} Altogether there appear to be a number of lines of evidence supporting this novel association.

WDPCP

A number of variants within the *WDPCP* gene were found to be significantly associated with GGT (most significant p-value 1.79e-10 for variant rs7566031), total cholesterol (most significant p-value 1.21e-08 for variant rs7566031), and LDL cholesterol (most significant p-value 1.79e-12 for variant rs7566031) levels (Figures 2,4). None of these associations have been previously reported for any variant within 100kb of the *WDPCP* locus, and the association with LDL was specifically found to replicate, at both the variant and gene level, in our meta-analysis. *WDPCP* was only found to be significantly expressed in adipose tissue, of all the tissues investigated (Figure 6), and although nearly every variant significantly associated with LDL levels was also significantly associated with *WDPCP* expression levels (Figure S33), the nature of the association peak would suggest that the peak within the *WDPCP* locus may represent the tail of an association peak centered over the neighboring gene, *EHBP1*, which has been previously reported as associated with LDL and total cholesterol levels.^{51,52} That being said, the *EHBP1* gene has not been implicated in any Mendelian genetic disorders, whereas bi-allelic pathogenic variants in *WDPCP* have been shown to cause Bardet-Biedl syndrome, a disorder characterized, in part, by hepatic fibrosis and dyslipidemia, both phenotypes characterized by perturbed GGT and LDL cholesterol levels.^{8,35}

Replication of known associations

Of the 27 significant gene-trait associations we identified that replicate previously-known associations from the literature, 11 have been reported as mapping to the ciliary candidate genes that we set out to study: *CENPF* with Creatinine,^{53,54} *CEP164* with HDL,^{44,55} *POC5* with Cholesterol LDL and HDL,^{51,52,56,57} *PROSER3* with HDL,⁵² *RP1* with Cholesterol and LDL,^{51,58,59} *RP1L1* with Triglycerides,^{51,52} *SDCCAG8* with Creatinine,^{53,55,60} and *SPATA7* with Creatinine⁵⁴ (Figures 2-3, S5-6, S22-23, S25-28). Our study adds further evidence to the hypothesis that these ciliary genes play an important role in affecting these laboratory phenotypes in the general population.

The remaining 16 significant, previously-reported associations mapped to 5 ciliary genes in our study (*CEP170*, *DYNC2LI1*, *IFT172*, *TRIM32*, *TTC8*), all of which exist within gene-dense regions of the genome. In these cases, the previously-reported mapped gene for the associated locus was different than the ciliary gene being investigated. In all cases, we found that the peak of association spanned multiple genes, making the identification of a single causative gene difficult. Four of these genes are discussed in more detail below.

CEP170* and *TTC8

Interestingly, in the case of *CEP170* and *TTC8*, the previously-mapped gene for each association peak was a different ciliary gene, and in both cases were ciliary genes that we also identified as significantly associated with the same trait. We found variants within the *CEP170* gene to be significantly associated with creatinine (Figure 3), with the association at this locus previously reported as mapping to the neighboring *SDCCAG8* gene (also found in our study); we found variants within the *TTC8* gene to also be significantly associated with creatinine (Figure 3), with the association at this locus previously reported as mapping to the neighboring *SPATA7* gene

(also found in our study). All four of these ciliary genes are significantly expressed in the kidney (Figure 6). The association peak at the *SPATA7/TTC8* locus spans five genes altogether, with our eQTL analysis more supportive of *SPATA7*, rather than *TTC8*, as the potentially causative gene at this locus (Figures S28, S32). The association peak at the *SDCCAG8/CEP170* locus, on the other hand, spans only these two ciliary genes, with our eQTL analysis showing strong evidence for correlation between expression levels of both of these genes and creatinine levels (Figures S7, S27). In fact, our eQTL analysis shows that each variant associated with changes in expression of *SDCCAG8* has the same direction of effect on *CEP170* expression.

DYNC2LI1* and *IFT172

Two additional ciliary genes with significant, previously-reported associations bear mention: *DYNC2LI1* and *IFT172*. Both of these genes, critical components of the ciliary IFT machinery,^{61,62} are located in gene-dense regions of the genome that have previously been associated with a number of patient phenotypes. The *DYNC2LI1* gene neighbors the genes *ABCG5* and *ABCG8*, both known to be important in cholesterol metabolism,^{51,52,59,63} while the *IFT172* gene abuts the *GCKR* gene, which has been extensively associated with blood glucose, triglyceride, LDL, GGT, AlkPhos, and creatinine levels.^{44,51–54,64–69}

We found the *DYNC2LI1* locus to be significantly associated with LDL and total cholesterol, as has been previously published, but also found significant associations with A1c (most significant p-value 4.50e-13 for rs116520905), Creatinine (most significant p-value 1.14e-08 for rs116520905), and ALT levels (most significant p-value 4.29e-14 for rs56266464) – all of them previously unreported, and none of which replicated in our meta-analysis (Figures 2-5). The *DYNC2LI1* gene is known to be causative of two Mendelian ciliopathy syndromes, Ellis-van Creveld syndrome and short rib polydactyly syndrome, both syndromes sometimes characterized by kidney and liver disease.^{62,70} eQTL analysis was not particularly supportive of the *DYNC2LI1* gene being the causative gene in the locus for any of these associations (Figure S10), and altogether it seems that these associations may be driven by the significant variants' effects on neighboring genes such as *ABCG5* and *ABCG8*.

The *IFT172* locus, on the other hand, was found to be significantly associated with nine separate phenotypes (A1c, AlkPhos, Cholesterol, Creatinine, GGT, Glucose, LDL, Triglycerides, Urea)(Figures 2-5), all of which had been previously reported and mapped to the neighboring *GCKR* gene. *IFT172* is a critically important gene known to be causative of at least five severe ciliopathy syndromes characterized by renal disease, obesity, impaired glucose handling, dyslipidemia, and hepatic fibrosis.^{61,71,72} We found *IFT172* to be ubiquitously expressed in all tissues we examined (Figure 6), and our eQTL analysis revealed significant correlations between p_{trait} and p_{eQTL} for *IFT172* and a number of hepatic, liver, and lipid phenotypes (Figure S13), providing further evidence linking *IFT172* to the associated phenotypes. Thus altogether it remains difficult to confidently assign causality at this locus to either *GCKR* or *IFT172* alone.

DiCE/pathway analysis reveals divergent phenotype associations for different ciliary compartments

To integrate the multiple lines of evidence we had gathered linking each ciliary gene to each significantly associated phenotype, we employed a Diverse Convergent Evidence⁷³ (DiCE) analysis to estimate the strength of available corroborating data supporting a given gene-trait association. Each gene-trait pair with a significant association in our discovery analysis was given a DiCE score based on the reproducibility of the association, tissue expression analysis, eQTL analysis, and associated Mendelian disease phenotypes. The smallest score received was a 3 (given to *KIF17*-AlkPhos and *NIN*-Creatinine), whereas four genes received the maximum possible score of 11 (*IFT172*-Glucose, Triglycerides; *IQCB1*-Creatinine; *SDCCAG8*-Creatinine;

TRIM32-Creatinine). These DiCE scores were used to generate a schematic (Figure 7) illustrating the strength of evidence supporting each ciliary gene's association with a given phenotype domain.

Using this approach, interesting patterns emerged in the data implicating different ciliary compartments in different disease processes. The DiCE score for genes affecting each phenotype domain were summed, by ciliary compartment, to generate a 4x4 contingency table (Table S16). A chi-square test performed on this data revealed significant difference across groups ($p = 2.728e-05$), primarily driven by the relative overabundance of basal body-associated genes and underabundance of IFT-associated genes within the kidney phenotype group (Figure S41). Of the 12 genes important to ciliary basal body function, 9 (75%) were found to associate with kidney phenotypes, whereas only 3 (25%) were found to associate with either liver or lipid traits, and only 1 (8%) was found to associate with glucose traits (Figure 7). At the same time, IFT-related genes accounted for 50% of the genes significantly associated with glucose traits and 42% of genes significantly associated with liver and lipid traits, but for only 21% of genes associated with kidney traits (Figure 7). Taken together, these data indicate that components of the ciliary basal body appear to play an important role in the normal function of the kidney, while ciliary genes involved in IFT appear more critical to liver-, lipid-, and glucose-related traits.

Yet another pattern that emerged in the data regarded differences in pleiotropy across different phenotypes. A majority of genes associated with either glucose (83%), lipid (67%), or liver-related (57%) traits demonstrated pleiotropic effects, being significantly associated with multiple traits across these three domains, whereas only 5 of 19 (26%) genes associated with kidney function demonstrated any pleiotropy at all (Figure 7). Interestingly, all five of the genes significantly associated with kidney phenotypes that were found to exhibit pleiotropy were specifically also associated with liver traits, with three (*DYNC2LI1*, *IFT172*, and *RP1L1*) exhibiting pleiotropic effects across all four phenotypic domains.

Taken together, the data yielded by our DiCE/Pathway analysis, seem to indicate that there exist divergent ciliary genetic pathways, one more specific to kidney function, and the other more specific to liver, lipid, and glucose physiology.

DISCUSSION

Our study set out to investigate associations between common variants in genes important to primary cilium structure/function and 12 diverse quantitative laboratory traits associated with common complex diseases. Our analysis identified 73 statistically-significant gene-trait associations across 34 of the 122 ciliary genes that we examined. With few exceptions, these ciliary genes were found to be widely expressed in human tissues relevant to the phenotypes being studied. In the vast majority of cases, significant trait-associated variants were also found to act as eQTLs for the ciliary genes being investigated, with strong evidence for correlation between gene expression levels and patient phenotypes. Strikingly, in many cases, the phenotypes significantly associated with the ciliary gene being studied were features of the Mendelian ciliopathies caused by rare loss of function variants in the same gene. Together, our data provide a number of lines of evidence supporting the cilium's role in the pathogenesis of common disease, and challenge the widely-held belief that the cilium is an organelle important mainly in development and in rare syndromic disease pathogenesis.

The finding that the primary cilium is involved in common disease pathogenesis is an important milestone in our understanding of this organelle's role in human physiology. Just as the cilium had been overlooked for decades before being rediscovered as a major player in rare Mendelian disease, the discovery of the cilium's role in common complex disease pathogenesis may represent a new rethinking of the importance of this tiny organelle. Perhaps even more exciting, however, is the possibility that common pathologic pathways might underlie the development of both rare syndromic and common complex disease. This possibility is further highlighted by perhaps the most interesting and unexpected finding of our study – that different ciliary subcompartments are specifically associated with distinct sets of patient phenotypes. This concept, that even within a single organelle, distinct subsets of genes/proteins may act together to affect patient phenotypes in divergent ways, is not new to Mendelian disease, but is an interesting and fascinating finding in common disease genetics. In fact, the recognition that particular ciliary pathways/complexes may be specifically involved in divergent common disease processes leaves the door open for novel targeted therapeutics and diagnostic approaches for common disease, informed by our extensive knowledge of ciliopathy pathogenesis and rare disease pathways. The further characterization of these previously unknown links between the cilium and common complex disease will establish a new field of study that has the potential to fundamentally transform our understanding of common disease pathogenesis.

An unexpected result of our analysis was the identification of four ciliary genes that were not significantly expressed in any tissues we examined in the GTEx database – *CENPF*, *WDPCP*, *RP1*, and *RP1L1* (Figure 5, Table S15). In the case of *CENPF* and *WDPCP*, this is a surprising finding, as rare variants in each of these genes are known to cause complex, multisystem ciliopathy disorders,^{74–76} suggesting broad expression across multiple tissue types. We cannot rule out the possibility that these genes are expressed broadly only during fetal development but not adulthood, that they are expressed only in a small but critical subset of cells, or that the corresponding mRNAs are unstable, making their detection in the GTEx database unlikely.

RP1 and *RP1L1*, on the other hand, represent more interesting cases. These two genes are highly homologous to each other, with mutations in each being well-established causes of Mendelian disorders characterized by retinal degeneration.^{77,78} In both cases, the genes are thought to be expressed exclusively in the retina.⁷⁸ In our discovery analysis, both gene loci were found to be significantly associated with numerous laboratory phenotypes spanning multiple organ systems (*RP1*: AST, GGT, LDL, Cholesterol; *RP1L1*: A1c, AlkPhos, ALT, AST, Cholesterol, Creatinine, and GGT), which is not consistent with retina-only expression of either gene. In both cases, the variants significantly associated with each phenotype were also found to significantly affect *RP1/RP1L1* expression levels based on our eQTL analysis, with evidence for correlation between gene expression levels and laboratory trait levels (Figures S25-26). *RP1* exists in a locus with only one nearby gene, *SOX17*; interestingly, there did not appear to be a significant correlation between variants associated with *SOX17* expression and any laboratory trait level (Figure S36), suggesting that the association between variants in this locus and AST, GGT, and LDL may, in fact, be mediated by changes in *RP1* expression, through currently unclear mechanisms. *RP1L1* exists in a locus with a number of neighboring genes – *C8orf74*, *PINX1*, *PRSS55*, *SOX7*– and although our eQTL analysis found that the variants significantly associated with each trait also associated with the expression of many of these genes, by far the strongest correlation was with *RP1L1* expression (figure S26, S37-40). Clearly, further work is needed to understand if and how genetic variants in these two apparently retina-specific genes are affecting non-retinal phenotypes.

Another interesting consequence of our analysis was the identification of gene-dense genomic loci, containing at least one ciliary gene, that were found to strongly associate with laboratory

traits. It has traditionally been difficult to attribute causality to a single gene in loci such as this, which display high degrees of linkage disequilibrium across multiple genes. In two cases, the loci we identified contained two different ciliary genes (*CEP170* and *SDCCAG8*; *TTC8* and *SPATA7*) (Figures S7, S27, S28, S32). These examples suggest that a single genomic locus might be associated with a given phenotype through the perturbation of expression of multiple genes operating within the same pathway. In fact, the primary cilium may serve as an excellent model for observing such associations and testing this hypothesis.

In a different case, we identified variants within the critically important ciliary IFT gene *IFT172* as being significantly associated with nine separate phenotypes – A1c, AlkPhos, Cholesterol, Creatinine, GGT, Glucose, LDL, Triglycerides, and Urea (Figure S13). *IFT172* neighbors the gene *GCKR*, which has traditionally been considered the mostly likely causal candidate gene in this locus. *GCKR* encodes a well-studied regulatory protein important in controlling glucose flux through the glycolytic pathway, and has been shown *in vitro* to have significant effects on triglyceride and glucose metabolism.⁷⁹ Previous fine mapping work of the *GCKR* locus has identified a relatively common missense variant in the gene as the putative causal variant responsible for the strong association of this locus with these multiple phenotypes.⁸⁰ However, there is minimal evidence that the protein encoded by the *GCKR* gene actually plays an important role in disease pathogenesis; knockout mice completely deficient in *GCKR* are normoglycemic except under extreme dietary conditions, and express no other apparent health phenotypes,^{81,82} and domestic cats have been shown to be completely deficient in *GCKR* mRNA and protein product as a species.⁸³ It seems unlikely that a gene that, in other species, can be completely abolished without obvious health effects could be solely responsible for so many robust human phenotype associations. *IFT172*, on the other hand, is a critically important gene known to be causative of at least five severe ciliopathy syndromes characterized by renal disease, obesity, impaired glucose handling, dyslipidemia, and hepatic fibrosis,^{61,71,72} and seems a much more likely candidate gene for the association peak based solely on its known pathogenic potential. Furthermore, unlike *GCKR*, which is expressed primarily in the liver,⁸⁴ *IFT172* was found to be ubiquitously expressed in all tissues we examined, with our eQTL analysis revealing significant correlations between p_{trait} and p_{eQTL} for *IFT172* and a number of hepatic, liver, and lipid phenotypes. This situation is reminiscent of the ongoing discussion surrounding the *FTO* locus – countless GWAS and fine mapping studies have linked the locus surrounding the *FTO* gene to body mass index and body fat percentage, but recent evidence suggests that the neighboring ciliary gene *RPGRIP1L*, may at least be partially driving these associations.^{85–87} Clearly more work is needed to disentangle the contributions of both *GCKR* and *IFT172* to the many patient phenotypes associated with this genomic locus.

Altogether, our data demonstrate the utility of a Mendelian disease-based approach to common variant association studies, and show that Mendelian disease genes can play an important role in common disease pathogenesis. Our data implicate the primary cilium as an important player in common disease, challenging the widely-held belief that the cilium is an organelle important mainly in development and in rare syndromic disease pathogenesis. It is not altogether surprising that such a critically important organelle, causative of rare ciliopathies characterized by phenotypes ranging from kidney failure to hepatic fibrosis, has important roles in the function of these same organ systems more generally; it stands to reason that a pathway important in rare disease might also be important for common disease affecting the same organ systems. The continued application of techniques similar to those described here to other phenotypes/Mendelian diseases is likely to yield many additional fascinating associations that will begin to integrate the fields of common and rare disease genetics, and provide insight into the pathophysiology of human diseases of immense public health burden.

ACKNOWLEDGMENTS

We would like to thank Michael P. Hart and Pinar S. Gurel for their careful reading of the manuscript and helpful comments.

FUNDING

TGD is supported in part by the NIH T32 training grant 5T32GM008638-23. MDR is supported in part by NIH R01 AI077505.

METHODS

Selection of cilium genes for analysis

122 genes with well-characterized roles in ciliary biology¹⁷ were selected for analysis. In all cases, gene transcript coordinates were defined using reported transcriptional start and stop sites that yielded the maximum transcript length for each gene. A complete list of the genes included in this study, along with the genomic coordinates used, is included in Table S17. All genes were classified as being components of ciliary sub-compartments (basal body, transition zone, BBsome, or IFT-associated) based on literature review^{17,88}. Mendelian disease associations for each gene were obtained from the Online Mendelian Inheritance in Man (OMIM)⁸⁹ database and literature review.

Preparation of UKBB Genotype data

The UKBB cohort release version 2 has deep genetic and phenotypic data on ~500,000 subjects across the United Kingdom. Subjects were genotyped on two similar genotype arrays across 106 batches and imputed to 96 million variants.⁹⁰ Quality control of genotype data was performed largely as previously described.⁹⁰ Subjects with poor quality genotype data were excluded from our analysis. Genetically related subjects (second-degree or closer with pi-hat larger than 0.25) were also excluded. European ancestry subjects were extracted using a combination of self-reported European ancestry (UKBB Data-Field: 21000) and genetic principal component analysis (UKBB Data-Field: 22006). We next filtered variants to retain only those with a minor allele frequency between 0.4 and 0.01, and which fell within the genomic regions defined by our set of 122 ciliary genes. This final set of variants was subjected to linkage disequilibrium (LD) pruning with an r^2 threshold of 0.5 using a 50 variant window shifted by 5 variant steps. After quality control 378,213 subjects and 16,874 genetic variants were included in our discovery association analyses. Principle components of genomic data that were subsequently used as covariates for association studies were provided in the UKBB data release.

Preparation of UKBB Phenotype data

For all 378,213 subjects passing genotype quality control, laboratory measurement information from the UKBB cohort release version 2 was extracted for the following phenotypes: Alanine aminotransferase (ALT; UKBB Data-Field: 30620), Alkaline Phosphatase (AlkPhos; UKBB Data-Field: 30610), Aspartate aminotransferase (AST; UKBB Data-Field: 30650), Cholesterol (UKBB Data-Field: 30690), Creatinine (UKBB Data-Field: 30700), Gamma glutamyltransferase (GGT)(UKBB Data-Field: 30730), Glucose (UKBB Data-Field: 30740), Glycated haemoglobin (A1c; UKBB Data-Field: 30750), HDL cholesterol (HDL; UKBB Data-Field: 30760), LDL cholesterol (LDL; UKBB Data-Field: 30780), Triglycerides (UKBB Data-Field: 30870), and Urea (UKBB Data-Field: 30670). For each subject, only baseline initial encounter laboratory measurements were

included. For each phenotype, outlier subjects with laboratory measurements greater than three standard deviations from the mean were excluded. For each phenotype, values were subjected to BoxCox⁹¹ or natural log transformation to maximize the normalcy of distribution. Following these quality control steps, the following number of subjects, per phenotype, were included in association analyses: A1c (n= 354,174), AlkPhos (n= 357,629), ALT (n= 354,959), AST (n= 355,215), Cholesterol (n= 359,024), Creatinine (n= 358,650), GGT (n= 354,772), Glucose (n= 324,597), HDL (n= 327,335), LDL (n= 358,417), Triglycerides (n= 354,118), Urea (n= 357,364). Summary statistics for the laboratory data used for our analysis can be found in Table S1. The UKBB data was accessed using application #32133.

Discovery Association Studies

Association studies were performed using the genotype and phenotype data prepared as described above using the PPlatform for the Analysis, Translation, and Organization of large-scale data (PLATO)^{92,93}, a standalone program developed by the Ritchie lab for the performance of phenome-wide linear/logistic regression of genetic variants against participant phenotypes. Linear regression, assuming an additive genetic model, was performed across all 16,874 ciliary genetic variants and the 12 phenotypes listed above. Regression models were adjusted by age, sex, and the first ten principle components of the corresponding genomic data for all 378,213 European ancestry subjects.

Meta-analyses

Transethnic meta-analyses were carried out using the Stouffer method in the program METAL⁹⁴ using publicly-available GWAS^{44,95–97} and meta-analysis^{53,56,58,98–100} summary statistics. Studies were analyzed together only if performed on the same trait, with the exception of Glucose (where studies looking at both fasting and non-fasting glucose levels were analyzed together) and Creatinine (where studies looking at both serum creatine and estimated glomerular filtration rate (eGFR, a metric derived from the serum creatinine, along with age, sex, and race) were analyzed together). In all cases, variants were filtered to retain only those which fell within the genomic regions defined by our set of 122 ciliary genes. Where necessary, variant nomenclature was standardized to ensure that identical variants were appropriately analyzed across studies. The studies included for meta-analysis contained no overlapping sample sets, with the exception of the meta-analyses of Creatinine/eGFR, and Urea. In both these cases, there was an overlap of 6,492 individuals between the large meta-analysis of eGFR and Urea traits (n = 567,460) conducted by Wuttke et al.⁵³ (which contained samples from an earlier release of the BioVu dataset), and a GWAS performed on the current BioVU dataset⁹⁷ (urea n = 33,322, creatinine n = 33,493). In both cases, we performed separate meta-analyses both including and excluding the current BioVu dataset (Figure S42) and found no differences in terms of the number of significant replicating loci identified. The meta-analyses including both of these data sets, with the small number of overlapping samples, are displayed in the main figures of the manuscript. The total transethnic sample size for each trait meta-analyzed was: A1c (n = 215,760), AlkPhos (n = 178,929), ALT (n = 208,562), AST (n = 208,165), Cholesterol (n = 369,671), Creatinine/eGFR (n = 831,567), GGT (n = 136,161), Glucose fasting/non-fasting (n = 229,080), HDL (n = 311,121), LDL (n = 312,410), Triglycerides (n = 346,901), and Urea (n = 798,670). A more detailed description of the studies included in our meta-analyses can be found in Table S18.

Significance Thresholds and Novel Findings

For all 12 traits in our discovery analysis, analyzed across 16,874 ciliary variants, the Bonferroni-corrected p-value significance threshold was calculated at 2.5e-7, but we chose to only report as significant those associations reaching the more stringent, commonly-accepted genome-wide p-value significance threshold of 5e-8. Our meta-analyses included many more variants than those in our discovery analysis, and thus we defined significant replication in our meta-analyses in two

ways; at the variant level, and at the gene level. At the variant level, analyzing only the union of variants present in both the discovery analysis and meta-analysis for a given trait, a variant-trait association was deemed to replicate if: (1) the variant was significantly associated with the trait in our discovery analysis with $p < 5e-8$ and (2) the p-value of association for the same variant, with the same trait, in the meta-analysis was below the Bonferroni-corrected significance threshold for the number of tests performed, defined as the number of variants shared between the discovery analysis and replication meta-analysis for a given trait. Listed here, for each trait, is the total number of variants shared between the discovery analysis and replication meta-analysis, along with the corresponding Bonferroni-adjusted significance threshold used to determine significant variant-level replication: A1c (14,271; $3.5e-6$), AlkPhos (15,197; $3.3e-6$), ALT (15,197; $3.3e-6$), AST (15,186; $3.3e-6$), Cholesterol (15,143; $3.3e-6$), Creatinine/eGFR (13,331; $3.8e-6$), GGT (15,171; $3.3e-6$), Glucose (15,132; $3.3e-6$), HDL (15,143; $3.3e-6$), LDL (15,143; $3.3e-6$), Triglycerides (15,143; $3.3e-6$), Urea (13,336; $3.8e-6$). The experiment-wise error rate for all tests across all phenotypes/variants in this analysis was calculated at $2.8e-7$. At the gene level, a gene-trait association was deemed to replicate if: (1) any variant within the gene was found to be significantly associated with the trait in our discovery set with $p < 5e-8$ and (2) the p-value of association between the same trait and any variant within the same gene in the meta-analysis was below the Bonferroni-corrected significance threshold for tests performed, defined using the total number of variants analyzed in the meta-analysis for a given trait. Listed here, for each trait, is the total number of variants in the replication meta-analysis, along with the corresponding Bonferroni-adjusted significance threshold used to determine significant gene-level replication: A1c (104,706; $4.8e-7$), AlkPhos (153,637; $3.3e-7$), ALT (153,186; $3.3e-7$), AST (149,729; $3.3e-7$), Cholesterol (92,118; $5.4e-7$), Creatinine/eGFR (42,925; $1.2e-6$), GGT (146,818; $3.4e-7$), Glucose (91,397; $5.5e-7$), HDL (92,106; $5.4e-7$), LDL (92,012; $5.4e-7$), Triglycerides (92,107; $5.4e-7$), Urea (42,701; $1.2e-6$). The experiment-wise error rate for all tests across all phenotypes/variants in this analysis was calculated at $4e-8$. We defined novel findings as those that (1) were significant in our discovery analysis with $p < 5e-8$, (2) replicated at the gene and/or variant level (as defined above) in meta-analysis, and (3) were at least 100kb away from any other variant previously reported as significantly associated with the given phenotype. The Bonferroni-corrected significance threshold for all 1,456,062 independent association tests performed in our entire study was calculated to be $3.4e-8$.

Discovery/Replication Meta-analysis Result Visualization

Results were visualized using a modified version of the Hudson R package¹⁰¹ to easily compare findings between our discovery analysis in the UKBB and replication meta-analyses of the same phenotype. For each Hudson plot, our discovery analysis results are plotted on the top half of the plot, with our meta-analysis results plotted in the bottom half of the plot. Each analyzed gene is given equal space, with all tested genetic variants for a given gene plotted at the midline of the gene block (i.e., variants are not represented in chromosomal space, but compressed to a single horizontal position per gene). For all Hudson plots, the study-wise Bonferroni-adjusted significance threshold ($3.4e-8$) is represented by a blue dashed line, while the Bonferroni-adjusted significance threshold for the trait being analyzed is represented by a red dashed line.

Tissue Specific Expression Studies

Tissue-specific expression data for each gene with significant associations in our discovery analysis was obtained from version 8 of the Genotype-Tissue Expression (GTEx) Project.²⁹ The GTEx Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 5/20/2020. The tissues analyzed were limited to those deemed to be relevant to the studied phenotypes, and included data from kidney, liver, pancreas, adipose, and skeletal muscle. A gene was considered

to be significantly expressed in a given tissue if the GTEx-derived transcripts per million (TPM) for the gene in that tissue was greater than 1.5. A heatmap of tissue-specific expression information for each gene was generated using the ggplot2 package for R.¹⁰²

eQTL Plots

The eQTpLot package was used to make all three eQTL plots; the R package is available on GitHub at <https://github.com/RitchieLab/eQTpLot>.^{31,32}

eQTL Plot

For each candidate gene with a significant association in our discovery analysis, we obtained a list of all variants associated with the gene's expression level, across all tissues, from the GTEx version 8 dataset, as above. For each variant with significant associations with a given candidate gene's expression in multiple tissues, we selected data only from the tissue with the most significant association. We defined a variant as an expression quantitative trait locus (eQTL) for the candidate gene if its p-value of association with gene expression (p_{eQTL}) was <0.05 . We next defined a locus of interest (LOI) for each candidate gene to include the candidate gene's coordinates, along with 200kb of flanking genomic material on either side. We repeated our association studies, as described in "Discovery Association Studies" above, for each significant candidate gene-trait pair, this time analyzing all variants in the LOI. To create the eQTL Plot, we plotted each variant in the LOI in chromosomal space on the x-axis, with the p-value of association with the phenotypic trait (p_{trait}) on the y-axis. If a variant was determined to be an eQTL for the candidate gene, it was plotted as a colored triangle – in blue, for variants that had congruent directions of effect (e.g. the variant was associated with increased transcript levels of the candidate gene, and also with an increase in the quantitative phenotype being studied), or in red for variants that had incongruous directions of effect (e.g. the variant was associated with increased transcript levels of the candidate gene, but with a decrease in the quantitative phenotype being studied), with a color gradient corresponding to the magnitude of p_{eQTL} . The size of each triangle was set to correspond to the eQTL normalized effect size for the variant, as obtained from GTEx, while the directionality of each triangle was set to correspond to the direction of effect of the variant for the phenotypic trait. Variants that were not found to be eQTLs (no eQTL data available, or $p_{\text{eQTL}} > 0.05$) for the candidate gene were plotted as grey boxes. A depiction of the genomic positions of all genes within the LOI was added below the plot using the package Gviz for R.¹⁰³

eQTL Enrichment Plots

For variants within the LOI with $p_{\text{trait}} < 5e-8$, the proportion that were also candidate gene eQTLs was calculated and plotted, and the same was done for variants with $p_{\text{trait}} > 5e-8$. Significant enrichment for eQTLs among the variants with significant association with the phenotypic trait was determined by Fisher's exact test. Bonferroni-corrected significance thresholds for enrichment were set at $p < 7e-4$ (adjusting for testing for eQTL enrichment across the 73 gene-trait pairs identified as significant in our discovery analysis). Variants with congruent and incongruent directions of effect, as described above, were considered separately.

P-P Plots

Each variant within a given LOI, as defined above, was plotted with p_{eQTL} along the x-axis, and p_{trait} along the y-axis. Correlation between the two probabilities was visualized by plotting a best-fit linear regression over the points, with the line equation displayed on the plot. The Pearson correlation coefficient and p-value of correlation were computed and displayed on the plot as well. Separate plots were made and superimposed over each other for variants with congruent and incongruent directions of effect, as described above. Bonferroni-corrected significance thresholds for Pearson correlation were set at $p < 3.5e-4$ (adjusting for testing for correlation across the 73

gene-trait pairs identified as significant in our discovery analysis, once for congruent eQTLs and once for incongruent eQTLs).

DiCE Analysis

Each significant gene-trait pair was given a DiCE score⁷³ ranging from 1 to 11 summarizing the amount of evidence linking the candidate gene with the associated trait. One point was assigned for the initial discovery association. One point was assigned if the association replicated in our meta-analysis (at either the variant or gene level), and two points were assigned if the association had been previously reported (for any variant within 100kb of the ciliary candidate gene). One point was given if a gene was found to be expressed in at least one tissue relevant to the trait of interest, while two points were given if a gene was found to be expressed in all relevant tissue types (relevant tissue types considered were: Kidney (medulla) and Kidney (cortex) for Creatinine/Urea; Liver, Pancreas, and Muscle (skeletal) for Glucose/A1c; Liver, Adipose (subcutaneous) and Adipose (visceral) for Cholesterol/LDL/HDL/Triglycerides; Liver for AST/ALT/AlkPhos/GGT). If the association peak for a candidate gene was significantly enriched for candidate gene eQTLs, one point was assigned. If the Pearson correlation between p_{trait} and p_{eQTL} was greater than 0.25 with a p-value $< 5e-5$, one point was assigned. If both these outcomes were true, an additional one point was assigned. Lastly, if the associated trait was relevant to the Mendelian disease associated with the candidate gene, two additional points were assigned. As a single gene could be associated with multiple traits within a single trait domain (e.g. AlkPhos, ALT, AST, or GGT for Liver-related traits), the highest DiCE score within a trait domain was selected for each gene with multiple associations.

Figure Legends:

Figure 1. Schematic of analysis workflow for the data presented

Phenotype and genotype data from the UKBB release version 2 was extracted and processed as indicated. Linear association studies between ciliary gene variants and patient phenotypes was performed, with genome-wide significant associations ($p < 5e-8$) being studied further by tissue-specific expression, eQTL, and replication meta-analysis. Using this data, a DiCE/Pathway analysis was performed to identify ciliary subcompartments associated with specific traits.

Figure 2. Association study and meta-analysis results for lipid-related traits

Hudson plot illustrating the results of the discovery association analysis (top) and replication meta-analysis (bottom) of common variants within 122 ciliary gene with lipid-related traits. Data for cholesterol is in teal, HDL in red, LDL in purple, and triglycerides in orange. The study-wide Bonferroni-adjusted significance threshold ($p < 3.4e-8$) is shown as a blue dashed line, while the experiment-wide Bonferroni-adjusted significance threshold ($p < 3.0e-6$ for the discovery analysis, $p < 3.3e-6$ for the meta-analysis) is shown as a red dashed line. Each analyzed gene is given equal space along the horizontal axis, with all tested genetic variants for a given gene plotted at the midline of the gene block.

Figure 3. Association study and meta-analysis results for kidney-related traits

Hudson plot illustrating the results of the discovery association analysis (top) and replication meta-analysis (bottom) of common variants within 122 ciliary gene with kidney-related traits. Data for creatinine is in teal, and data for urea is in red. The study-wide Bonferroni-adjusted significance threshold ($p < 3.4e-8$) is shown as a blue dashed line, while the experiment-wide Bonferroni-adjusted significance threshold ($p < 3.0e-6$ for the discovery analysis, $p < 3.8e-6$ for the meta-analysis) is shown as a red dashed line. Each analyzed gene is given equal space along the

horizontal axis, with all tested genetic variants for a given gene plotted at the midline of the gene block.

Figure 4. Association study and meta-analysis results for liver-related traits

Hudson plot illustrating the results of the discovery association analysis (top) and replication meta-analysis (bottom) of common variants within 122 ciliary gene with liver-related traits. Data for AlkPhos is in teal, ALT in red, AST in purple, and GGT in orange. The study-wide Bonferroni-adjusted significance threshold ($p < 3.4 \times 10^{-8}$) is shown as a blue dashed line, while the experiment-wide Bonferroni-adjusted significance threshold ($p < 3.0 \times 10^{-6}$ for the discovery analysis, $p < 3.3 \times 10^{-6}$ for the meta-analysis) is shown as a red dashed line. Each analyzed gene is given equal space along the horizontal axis, with all tested genetic variants for a given gene plotted at the midline of the gene block.

Figure 5. Association study and meta-analysis results for glucose-related traits

Hudson plot illustrating the results of the discovery association analysis (top) and replication meta-analysis (bottom) of common variants within 122 ciliary gene with glucose-related traits. Data for A1c is in teal, and data for glucose is in red. The study-wide Bonferroni-adjusted significance threshold ($p < 3.4 \times 10^{-8}$) is shown as a blue dashed line, while the experiment-wide Bonferroni-adjusted significance threshold ($p < 3.0 \times 10^{-6}$ for the discovery analysis, $p < 3.3 \times 10^{-6}$ for the meta-analysis) is shown as a red dashed line. Each analyzed gene is given equal space along the horizontal axis, with all tested genetic variants for a given gene plotted at the midline of the gene block.

Figure 6. Results of tissue-specific expression analysis

Heatmap depicting the results of the tissue-specific expression analysis for each of the 34 genes with significant associations in our discovery analysis. The phenotype domain(s) significantly associated with each gene are indicated by the colored bars at the top of the graph.

Figure 7. Schematics illustrating the results of DiCE/Pathway analysis, and a depiction of the primary cilium

(A) To integrate the multiple lines of evidence supporting each ciliary gene's association with a given phenotype domain, we employed a Diverse Convergent Evidence (DiCE) analysis approach estimating the strength of available corroborating data (detailed in Table S14). The DiCE scores for each gene, illustrating the strength of evidence supporting each ciliary gene's association with a given phenotype domain, are displayed here, with the magnitude of the score indicated by the size of each gene bubble. For each phenotype domain, gene bubbles are displayed at the same location for ease of comparison. Genes are colored by ciliary subcompartment – green for IFT-related genes, pink for the BBSome, orange for the transition zone, and grey for the basal body. A stacked bar graph is also displayed for each phenotype domain, illustrating the summed DiCE scores for all genes, by ciliary subcompartment, as a proportion of the total DiCE score per phenotype domain. (B) A schematic of the primary cilium. The basal body and centrioles are displayed in grey, the ciliary transition zone in orange, the IFT machinery in green, and the BBSome in pink. The core of 9 microtubule doublets and the overlying membrane of the organelle are shown, and receptors being trafficked into/along the cilium are shown in blue.

REFERENCES

1. Satir P, Pedersen LB, Christensen ST. The primary cilium at a glance. *J Cell Sci.* 2010 Feb 15;123(4):499–503. PMID: 20144997
2. Anvarian Z, Mykityn K, Mukhopadhyay S, Pedersen LB, Christensen ST. Cellular signalling by primary cilia in development, organ function and disease. *Nat Rev Nephrol.* 2019;15(4):199–219. PMCID: PMC6426138
3. Liu B, Chen S, Cheng D, Jing W, Helms JA. Primary cilia integrate hedgehog and Wnt signaling during tooth development. *J Dent Res.* 2014 May;93(5):475–482. PMCID: PMC3988628
4. Wei Q, Zhang Y, Li Y, Zhang Q, Ling K, Hu J. The BBSome controls IFT assembly and turnaround in cilia. *Nat Cell Biol.* 2012 Sep;14(9):950–957. PMCID: PMC3434251
5. Wingfield JL, Lechtreck K-F, Lorentzen E. Trafficking of ciliary membrane proteins by the intraflagellar transport/BBSome machinery. *Essays Biochem.* 2018 Dec 7;62(6):753–763. PMCID: PMC6737936
6. Lechtreck KF. IFT-Cargo Interactions and Protein Transport in Cilia. *Trends Biochem Sci.* 2015 Dec;40(12):765–778. PMCID: PMC4661101
7. Grochowsky A, Gunay-Aygun M. Clinical characteristics of individual organ system disease in non-motile ciliopathies. *Transl Sci Rare Dis.* 2019 Jul 4;4(1–2):1–23. PMCID: PMC6864414
8. Tsang SH, Aycinena ARP, Sharma T. Ciliopathy: Bardet-Biedl Syndrome. *Adv Exp Med Biol.* 2018;1085:171–174. PMID: 30578506
9. Parisi M, Glass I. Joubert Syndrome. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJ, Stephens K, Amemiya A, editors. *GeneReviews®* [Internet]. Seattle (WA): University of Washington, Seattle; 1993 [cited 2020 Jun 24]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK1325/> PMID: 20301500
10. Tsang SH, Aycinena ARP, Sharma T. Ciliopathy: Alström Syndrome. *Adv Exp Med Biol.* 2018;1085:179–180. PMID: 30578508
11. Hopp K, Heyer CM, Hommerding CJ, Henke SA, Sundsbak JL, Patel S, Patel P, Consugar MB, Czarnecki PG, Gliem TJ, Torres VE, Rossetti S, Harris PC. B9D1 is revealed as a novel Meckel syndrome (MKS) gene by targeted exon-enriched next-generation sequencing and deletion analysis. *Hum Mol Genet.* 2011 Jul 1;20(13):2524–2534. PMCID: PMC3109998
12. Gilissen C, Arts HH, Hoischen A, Spruijt L, Mans DA, Arts P, van Lier B, Steehouwer M, van Reeuwijk J, Kant SG, Roepman R, Knoers NVAM, Veltman JA, Brunner HG. Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. *Am J Hum Genet.* 2010 Sep 10;87(3):418–423. PMCID: PMC2933349

13. Halbritter J, Porath JD, Diaz KA, Braun DA, Kohl S, Chaki M, Allen SJ, Soliman NA, Hildebrandt F, Otto EA, GPN Study Group. Identification of 99 novel mutations in a worldwide cohort of 1,056 patients with a nephronophthisis-related ciliopathy. *Hum Genet.* 2013 Aug;132(8):865–884. PMID: PMC4643834
14. Otto EA, Loeys B, Khanna H, Hellemans J, Sudbrak R, Fan S, Muerb U, O'Toole JF, Helou J, Attanasio M, Utsch B, Sayer JA, Lillo C, Jimeno D, Coucke P, De Paepe A, Reinhardt R, Klages S, Tsuda M, Kawakami I, Kusakabe T, Omran H, Imm A, Tippens M, Raymond PA, Hill J, Beales P, He S, Kispert A, Margolis B, Williams DS, Swaroop A, Hildebrandt F. Nephrocystin-5, a ciliary IQ domain protein, is mutated in Senior-Loken syndrome and interacts with RPGR and calmodulin. *Nat Genet.* 2005 Mar;37(3):282–288. PMID: 15723066
15. Drivas TG, Holzbaur ELF, Bennett J. Disruption of CEP290 microtubule/membrane-binding domains causes retinal degeneration. *J Clin Invest.* 2013 Oct;123(10):4525–4539. PMID: PMC3784542
16. Drivas TG, Wojno AP, Tucker BA, Stone EM, Bennett J. Basal exon skipping and genetic pleiotropy: A predictive model of disease pathogenesis. *Sci Transl Med.* 2015 Jun 10;7(291):291ra97. PMID: PMC4486480
17. Reiter JF, Leroux MR. Genes and molecular pathways underpinning ciliopathies. *Nat Rev Mol Cell Biol.* 2017 Sep;18(9):533–547. PMID: PMC5851292
18. Gerdes JM, Christou-Savina S, Xiong Y, Moede T, Moruzzi N, Karlsson-Edlund P, Leibiger B, Leibiger IB, Östenson C-G, Beales PL, Berggren P-O. Ciliary dysfunction impairs beta-cell insulin secretion and promotes development of type 2 diabetes in rodents. *Nat Commun.* 2014 Nov 6;5:5308. PMID: 25374274
19. Volta F, Scerbo MJ, Seelig A, Wagner R, O'Brien N, Gerst F, Fritsche A, Häring H-U, Zeigerer A, Ullrich S, Gerdes JM. Glucose homeostasis is regulated by pancreatic β -cell cilia via endosomal EphA-processing. *Nat Commun* [Internet]. 2019 Dec 12 [cited 2020 Feb 2];10. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6908661/> PMID: PMC6908661
20. Gencer S, Oleinik N, Kim J, Selvam SP, De Palma R, Dany M, Nganga R, Thomas RJ, Senkal CE, Howe PH, Ogretmen B. TGF- β receptor I/II trafficking and signaling at primary cilia are inhibited by ceramide to attenuate cell migration and tumor metastasis. *Sci Signal* [Internet]. 2017 Oct 24 [cited 2020 Aug 12];10(502). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5818989/> PMID: PMC5818989
21. Labour M-N, Riffault M, Christensen ST, Hoey DA. TGF β 1 – induced recruitment of human bone mesenchymal stem cells is mediated by the primary cilium in a SMAD3-dependent manner. *Sci Rep* [Internet]. 2016 Oct 17 [cited 2020 Aug 12];6. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5066273/> PMID: PMC5066273
22. Clement CA, Ajbro KD, Koefoed K, Vestergaard ML, Veland IR, Henriques de Jesus MPR, Pedersen LB, Benmerah A, Andersen CY, Larsen LA, Christensen ST. TGF- β

- signaling is associated with endocytosis at the pocket region of the primary cilium. *Cell Rep.* 2013 Jun 27;3(6):1806–1814. PMID: 23746451
23. Schneider L, Clement CA, Teilmann SC, Pazour GJ, Hoffmann EK, Satir P, Christensen ST. PDGFR α signaling is regulated through the primary cilium in fibroblasts. *Curr Biol.* 2005 Oct 25;15(20):1861–1866. PMID: 16243034
 24. Schneider L, Cammer M, Lehman J, Nielsen SK, Guerra CF, Veland IR, Stock C, Hoffmann EK, Yoder BK, Schwab A, Satir P, Christensen ST. Directional cell migration and chemotaxis in wound healing response to PDGF-AA are coordinated by the primary cilium in fibroblasts. *Cell Physiol Biochem.* 2010;25(2–3):279–292. PMCID: PMC2924811
 25. Schmid FM, Schou KB, Vilhelm MJ, Holm MS, Breslin L, Farinelli P, Larsen LA, Andersen JS, Pedersen LB, Christensen ST. IFT20 modulates ciliary PDGFR α signaling by regulating the stability of Cbl E3 ubiquitin ligases. *J Cell Biol.* The Rockefeller University Press; 2018 Jan 2;217(1):151–161.
 26. Umberger NL, Caspary T, Bettencourt-Dias M. Ciliary transport regulates PDGF-AA/ $\alpha\alpha$ signaling via elevated mammalian target of rapamycin signaling and diminished PP2A activity. *MBoC.* American Society for Cell Biology (mboc); 2014 Nov 12;26(2):350–358.
 27. Christensen ST, Morthorst SK, Mogensen JB, Pedersen LB. Primary Cilia and Coordination of Receptor Tyrosine Kinase (RTK) and Transforming Growth Factor β (TGF- β) Signaling. *Cold Spring Harb Perspect Biol* [Internet]. 2017 Jun [cited 2020 Aug 13];9(6). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5453389/> PMCID: PMC5453389
 28. Oh EC, Vasanth S, Katsanis N. Metabolic Regulation and Energy Homeostasis through the Primary Cilium. *Cell Metabolism.* Elsevier; 2015 Jan 6;21(1):21–31. PMID: 25543293
 29. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* 2013 Jun;45(6):580–585. PMCID: PMC4010069
 30. Mitchison HM, Valente EM. Motile and non-motile cilia in human pathology: from function to phenotypes. *The Journal of Pathology.* 2017;241(2):294–309.
 31. RitchieLab/eQTLot [Internet]. Ritchie Lab; 2020 [cited 2020 Aug 17]. Available from: <https://github.com/RitchieLab/eQTLot>
 32. Drivas TG, Lucas A, Ritchie MD. eQTLot: an R package for the visualization and colocalization of eQTL and GWAS signals. *bioRxiv.* Cold Spring Harbor Laboratory; 2020 Aug 26;2020.08.26.268268.
 33. Nishimura DY, Searby CC, Carmi R, Elbedour K, Van Maldergem L, Fulton AB, Lam BL, Powell BR, Swiderski RE, Bugge KE, Haider NB, Kwitek-Black AE, Ying L, Duhl DM, Gorman SW, Heon E, Iannaccone A, Bonneau D, Biesecker LG, Jacobson SG, Stone EM, Sheffield VC. Positional cloning of a novel gene on chromosome 16q causing Bardet-Biedl syndrome (BBS2). *Hum Mol Genet.* 2001 Apr 1;10(8):865–874. PMID: 11285252

34. Benzinou M, Walley A, Lobbens S, Charles M-A, Jouret B, Fumeron F, Balkau B, Meyre D, Froguel P. Bardet-Biedl syndrome gene variants are associated with both childhood and adult common obesity in French Caucasians. *Diabetes*. 2006 Oct;55(10):2876–2882. PMID: 17003356
35. Mujahid S, Hunt KF, Cheah YS, Forsythe E, Hazlehurst JM, Sparks K, Mohammed S, Tomlinson JW, Amiel SA, Carroll PV, Beales PL, Huda MSB, McGowan BM. The Endocrine and Metabolic Characteristics of a Large Bardet-Biedl Syndrome Clinic Population. *J Clin Endocrinol Metab*. 2018 01;103(5):1834–1841. PMID: 29409041
36. Johnston JJ, Lee C, Wentzensen IM, Parisi MA, Crenshaw MM, Sapp JC, Gross JM, Wallingford JB, Biesecker LG. Compound heterozygous alterations in intraflagellar transport protein CLUAP1 in a child with a novel Joubert and oral–facial–digital overlap syndrome. *Cold Spring Harb Mol Case Stud [Internet]*. 2017 Jul [cited 2020 May 18];3(4). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5495032/> PMID: PMC5495032
37. Kichaev G, Bhatia G, Loh P-R, Gazal S, Burch K, Freund MK, Schoech A, Pasaniuc B, Price AL. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet*. 2019 03;104(1):65–75. PMID: PMC6323418
38. Zhu Z, Guo Y, Shi H, Liu C-L, Panganiban RA, Chung W, O'Connor LJ, Himes BE, Gazal S, Hasegawa K, Camargo CA, Qi L, Moffatt MF, Hu FB, Lu Q, Cookson WOC, Liang L. Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. *J Allergy Clin Immunol*. 2020 Feb;145(2):537–549. PMID: PMC7010560
39. Shrine N, Guyatt AL, Erzurumluoglu AM, Jackson VE, Hobbs BD, Melbourne CA, Batini C, Fawcett KA, Song K, Sakornsakolpat P, Li X, Boxall R, Reeve NF, Obeidat M, Zhao JH, Wielscher M, Weiss S, Kentistou KA, Cook JP, Sun BB, Zhou J, Hui J, Karrasch S, Imboden M, Harris SE, Marten J, Enroth S, Kerr SM, Surakka I, Vitart V, Lehtimäki T, Allen RJ, Bakke PS, Beaty TH, Bleecker ER, Bossé Y, Brandsma C-A, Chen Z, Crapo JD, Danesh J, DeMeo DL, Dudbridge F, Ewert R, Gieger C, Gulsvik A, Hansell AL, Hao K, Hoffman JD, Hokanson JE, Homuth G, Joshi PK, Joubert P, Langenberg C, Li X, Li L, Lin K, Lind L, Locantore N, Luan J, Mahajan A, Maranville JC, Murray A, Nickle DC, Packer R, Parker MM, Paynton ML, Porteous DJ, Prokopenko D, Qiao D, Rawal R, Runz H, Sayers I, Sin DD, Smith BH, Soler Artigas M, Sparrow D, Tal-Singer R, Timmers PRHJ, Van den Berge M, Whittaker JC, Woodruff PG, Yerges-Armstrong LM, Troyanskaya OG, Raitakari OT, Kähönen M, Polašek O, Gyllenstein U, Rudan I, Deary IJ, Probst-Hensch NM, Schulz H, James AL, Wilson JF, Stubbe B, Zeggini E, Jarvelin M-R, Wareham N, Silverman EK, Hayward C, Morris AP, Butterworth AS, Scott RA, Walters RG, Meyers DA, Cho MH, Strachan DP, Hall IP, Tobin MD, Wain LV, Understanding Society Scientific Group. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet*. 2019;51(3):481–493. PMID: PMC6397078
40. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, Payne AJ, Steinthorsdottir V, Scott RA, Grarup N, Cook JP, Schmidt EM, Wuttke M, Sarnowski C,

- Mägi R, Nano J, Gieger C, Trompet S, Lecoœur C, Preuss MH, Prins BP, Guo X, Bielak LF, Below JE, Bowden DW, Chambers JC, Kim YJ, Ng MCY, Petty LE, Sim X, Zhang W, Bennett AJ, Bork-Jensen J, Brummett CM, Canouil M, Ec Kardt K-U, Fischer K, Kardia SLR, Kronenberg F, Läll K, Liu C-T, Locke AE, Luan J, Ntalla I, Nylander V, Schönherr S, Schurmann C, Yengo L, Bottinger EP, Brandslund I, Christensen C, Dedoussis G, Florez JC, Ford I, Franco OH, Frayling TM, Giedraitis V, Hackinger S, Hattersley AT, Herder C, Ikram MA, Ingelsson M, Jørgensen ME, Jørgensen T, Kriebel J, Kuusisto J, Ligthart S, Lindgren CM, Linneberg A, Lyssenko V, Mamakou V, Meitinger T, Mohlke KL, Morris AD, Nadkarni G, Pankow JS, Peters A, Sattar N, Stančáková A, Strauch K, Taylor KD, Thorand B, Thorleifsson G, Thorsteinsdottir U, Tuomilehto J, Witte DR, Dupuis J, Peyser PA, Zeggini E, Loos RJJ, Froguel P, Ingelsson E, Lind L, Groop L, Laakso M, Collins FS, Jukema JW, Palmer CNA, Grallert H, Metspalu A, Dehghan A, Köttgen A, Abecasis GR, Meigs JB, Rotter JJ, Marchini J, Pedersen O, Hansen T, Langenberg C, Wareham NJ, Stefansson K, Gloyn AL, Morris AP, Boehnke M, McCarthy MI. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet.* 2018;50(11):1505–1513. PMID: PMC6287706
41. Greer YE, Westlake CJ, Gao B, Bharti K, Shiba Y, Xavier CP, Pazour GJ, Yang Y, Rubin JS. Casein kinase 1δ functions at the centrosome and Golgi to promote ciliogenesis. *Mol Biol Cell.* 2014 May;25(10):1629–1640. PMID: PMC4019494
 42. Gurkaslar HK, Culfa E, Arslanhan MD, Lince-Faria M, Firat-Karalar EN. CCDC57 Cooperates with Microtubules and Microcephaly Protein CEP63 and Regulates Centriole Duplication and Mitotic Progression. *Cell Reports.* 2020 May 12;31(6):107630.
 43. Bader JR, Kusik BW, Besharse JC. Analysis of KIF17 distal tip trafficking in zebrafish cone photoreceptors. *Vision Res.* 2012 Dec 15;75:37–43. PMID: PMC3556995
 44. Kanai M, Akiyama M, Takahashi A, Matoba N, Momozawa Y, Ikeda M, Iwata N, Ikegawa S, Hirata M, Matsuda K, Kubo M, Okada Y, Kamatani Y. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet.* 2018;50(3):390–400. PMID: 29403010
 45. Gonçalves J, Coyaud É, Laurent EMN, Raught B, Pelletier L. LUZP1 and the tumour suppressor EPLIN are negative regulators of primary cilia formation. *bioRxiv.* 2019 Aug 15;736389.
 46. Gupta GD, Coyaud É, Gonçalves J, Mojarad BA, Liu Y, Wu Q, Gheiratmand L, Comartin D, Tkach JM, Cheung SWT, Bashkurov M, Hasegan M, Knight JD, Lin Z-Y, Schueler M, Hildebrandt F, Moffat J, Gingras A-C, Raught B, Pelletier L. A dynamic protein interaction landscape of the human centrosome-cilium interface. *Cell.* 2015 Dec 3;163(6):1484–1499. PMID: PMC5089374
 47. Bozal-Basterra L, Gonzalez-Santamarta M, Bermejo-Arteagabeitia A, Fonseca CD, Pampliega O, Andrade R, Martín-Martín N, Branon TC, Ting AY, Carracedo A, Rodríguez JA, Elortza F, Sutherland JD, Barrio R. LUZP1, a novel regulator of primary cilia and the

actin cytoskeleton, is altered in Townes-Brooks Syndrome. *bioRxiv*. Cold Spring Harbor Laboratory; 2019 Jul 31;721316.

48. Hebbar M, Kanthi A, Shukla A, Bielas S, Girisha KM. A biallelic 36-bp insertion in PIBF1 is associated with Joubert syndrome. *J Hum Genet*. 2018 Jul;63(8):935–939. PMID: PMC6060014
49. Wheway G, Schmidts M, Mans DA, Szymanska K, Nguyen T-MT, Racher H, Phelps IG, Toedt G, Kennedy J, Wunderlich KA, Soroush N, Abdelhamed ZA, Natarajan S, Herridge W, van Reeuwijk J, Horn N, Boldt K, Parry DA, Letteboer SJF, Roosing S, Adams M, Bell SM, Bond J, Higgins J, Morrison EE, Tomlinson DC, Slaats GG, van Dam TJP, Huang L, Kessler K, Giessl A, Logan CV, Boyle EA, Shendure J, Anazi S, Aldahmesh M, Al Hazzaa S, Hegele RA, Ober C, Frosk P, Mhanni AA, Chodirker BN, Chudley AE, Lamont R, Bernier FP, Beaulieu CL, Gordon P, Pon RT, Donahue C, Barkovich AJ, Wolf L, Toomes C, Thiel CT, Boycott KM, McKibbin M, Inglehearn CF, Stewart F, Omran H, Huynen MA, Sergouniotis PI, Alkuraya FS, Parboosingh JS, Innes AM, Willoughby CE, Giles RH, Webster AR, Ueffing M, Blacque O, Gleeson JG, Wolfrum U, Beales PL, Gibson T, Doherty D, Mitchison HM, Roepman R, Johnson CA. An siRNA-based functional genomics screen for the identification of regulators of ciliogenesis and ciliopathy genes. *Nat Cell Biol*. 2015 Aug;17(8):1074–1087. PMID: PMC4536769
50. Kim K, Lee K, Rhee K. CEP90 Is Required for the Assembly and Centrosomal Accumulation of Centriolar Satellites, Which Is Essential for Primary Cilia Formation. *PLoS One* [Internet]. 2012 Oct 24 [cited 2020 Jun 28];7(10). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3480484/> PMID: PMC3480484
51. Hoffmann TJ, Theusch E, Haldar T, Ranatunga DK, Jorgenson E, Medina MW, Kvale MN, Kwok P-Y, Schaefer C, Krauss RM, Iribarren C, Risch N. A large electronic-health-record-based genome-wide study of serum lipids. *Nat Genet*. 2018;50(3):401–413. PMID: PMC5942247
52. Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, Gagnon DR, DuVall SL, Li J, Peloso GM, Chaffin M, Small AM, Huang J, Tang H, Lynch JA, Ho Y-L, Liu DJ, Emdin CA, Li AH, Huffman JE, Lee JS, Natarajan P, Chowdhury R, Saleheen D, Vujkovic M, Baras A, Pyarajan S, Di Angelantonio E, Neale BM, Naheed A, Khera AV, Danesh J, Chang K-M, Abecasis G, Willer C, Dewey FE, Carey DJ, Global Lipids Genetics Consortium, Myocardial Infarction Genetics (MIGen) Consortium, Geisinger-Regeneron DiscovEHR Collaboration, VA Million Veteran Program, Concato J, Gaziano JM, O'Donnell CJ, Tsao PS, Kathiresan S, Rader DJ, Wilson PWF, Assimes TL. Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat Genet*. 2018;50(11):1514–1523. PMID: PMC6521726
53. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, Tin A, Wang L, Chu AY, Hoppmann A, Kirsten H, Giri A, Chai J-F, Sveinbjornsson G, Tayo BO, Nutile T, Fuchsberger C, Marten J, Cocca M, Ghasemi S, Xu Y, Horn K, Noce D, Most PJ van der, Sedaghat S, Yu Z, Akiyama M, Afaq S, Ahluwalia TS, Almgren P, Amin N, Ärnlöv J, Bakker SJL, Bansal N, Baptista D, Bergmann S, Biggs ML, Biino G, Boehnke M,

Boerwinkle E, Boissel M, Bottinger EP, Boutin TS, Brenner H, Brumat M, Burkhardt R, Butterworth AS, Campana E, Campbell A, Campbell H, Canouil M, Carroll RJ, Catamo E, Chambers JC, Chee M-L, Chee M-L, Chen X, Cheng C-Y, Cheng Y, Christensen K, Cifkova R, Ciullo M, Concas MP, Cook JP, Coresh J, Corre T, Sala CF, Cusi D, Danesh J, Daw EW, Borst MH de, Grandi AD, Mutsert R de, Vries APJ de, Degenhardt F, Delgado G, Demirkan A, Angelantonio ED, Dittrich K, Divers J, Dorajoo R, Eckardt K-U, Ehret G, Elliott P, Endlich K, Evans MK, Felix JF, Foo VHX, Franco OH, Franke A, Freedman BI, Freitag-Wolf S, Friedlander Y, Froguel P, Gansevoort RT, Gao H, Gasparini P, Gaziano JM, Giedraitis V, Gieger C, Girotto G, Giulianini F, Gögele M, Gordon SD, Gudbjartsson DF, Gudnason V, Haller T, Hamet P, Harris TB, Hartman CA, Hayward C, Hellwege JN, Heng C-K, Hicks AA, Hofer E, Huang W, Hutri-Kähönen N, Hwang S-J, Ikram MA, Indridason OS, Ingelsson E, Ising M, Jaddoe VWV, Jakobsdottir J, Jonas JB, Joshi PK, Josyula NS, Jung B, Kähönen M, Kamatani Y, Kammerer CM, Kanai M, Kastarinen M, Kerr SM, Khor C-C, Kiess W, Kleber ME, Koenig W, Kooner JS, Körner A, Kovacs P, Kraja AT, Krajcoviechova A, Kramer H, Krämer BK, Kronenberg F, Kubo M, Kühnel B, Kuokkanen M, Kuusisto J, Bianca ML, Laakso M, Lange LA, Langefeld CD, Lee JJ-M, Lehne B, Lehtimäki T, Lieb W, Lim S-C, Lind L, Lindgren CM, Liu J, Liu J, Loeffler M, Loos RJJ, Lucae S, Lukas MA, Lyytikäinen L-P, Mägi R, Magnusson PKE, Mahajan A, Martin NG, Martins J, März W, Mascalfzoni D, Matsuda K, Meisinger C, Meitinger T, Melander O, Metspalu A, Mikaelssdottir EK, Milaneschi Y, Miliku K, Mishra PP, Mohlke KL, Mononen N, Montgomery GW, Mook-Kanamori DO, Mychaleckyj JC, Nadkarni GN, Nalls MA, Nauck M, Nikus K, Ning B, Nolte IM, Noordam R, O'Connell J, O'Donoghue ML, Olafsson I, Oldehinkel AJ, Orho-Melander M, Ouwehand WH, Padmanabhan S, Palmer ND, Palsson R, Penninx BWJH, Perls T, Perola M, Pirastu M, Pirastu N, Pistis G, Podgornaia AI, Polasek O, Ponte B, Porteous DJ, Poulain T, Pramstaller PP, Preuss MH, Prins BP, Province MA, Rabelink TJ, Raffield LM, Raitakari OT, Reilly DF, Rettig R, Rheinberger M, Rice KM, Ridker PM, Rivadeneira F, Rizzi F, Roberts DJ, Robino A, Rossing P, Rudan I, Rueedi R, Ruggiero D, Ryan KA, Saba Y, Sabanayagam C, Salomaa V, Salvi E, Saum K-U, Schmidt H, Schmidt R, Schöttker B, Schulz C-A, Schupf N, Shaffer CM, Shi Y, Smith AV, Smith BH, Soranzo N, Spracklen CN, Strauch K, Stringham HM, Stumvoll M, Svensson PO, Szymczak S, Tai E-S, Tajuddin SM, Tan NYQ, Taylor KD, Teren A, Tham Y-C, Thiery J, Thio CHL, Thomsen H, Thorleifsson G, Toniolo D, Tönjes A, Tremblay J, Tzoulaki I, Uitterlinden AG, Vaccargiu S, Dam RM van, Harst P van der, Duijn CM van, Edward DRV, Verweij N, Vogelesang S, Völker U, Vollenweider P, Waeber G, Waldenberger M, Wallentin L, Wang YX, Wang C, Waterworth DM, Wei WB, White H, Whitfield JB, Wild SH, Wilson JF, Wojczynski MK, Wong C, Wong T-Y, Xu L, Yang Q, Yasuda M, Yerges-Armstrong LM, Zhang W, Zonderman AB, Rotter JI, Bochud M, Psaty BM, Vitart V, Wilson JG, Dehghan A, Parsa A, Chasman DI, Ho K, Morris AP, Devuyst O, Akilesh S, Pendergrass SA, Sim X, Böger CA, Okada Y, Edwards TL, Snieder H, Stefansson K, Hung AM, Heid IM, Scholz M, Teumer A, Köttgen A, Pattaro C. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet.* 2019 Jun;51(6):957–972.

54. Hellwege JN, Velez Edwards DR, Giri A, Qiu C, Park J, Torstenson ES, Keaton JM, Wilson OD, Robinson-Cohen C, Chung CP, Roumie CL, Klarin D, Damrauer SM, DuVall SL, Siew E, Akwo EA, Wuttke M, Gorski M, Li M, Li Y, Gaziano JM, Wilson PWF, Tsao PS, O'Donnell CJ, Kovesdy CP, Pattaro C, Köttgen A, Susztak K, Edwards TL, Hung AM.

Mapping eGFR loci to the renal transcriptome and phenome in the VA Million Veteran Program. *Nat Commun.* 2019 26;10(1):3842. PMID: PMC6710266

55. Nagy R, Boutin TS, Marten J, Huffman JE, Kerr SM, Campbell A, Evenden L, Gibson J, Amador C, Howard DM, Navarro P, Morris A, Deary IJ, Hocking LJ, Padmanabhan S, Smith BH, Joshi P, Wilson JF, Hastie ND, Wright AF, McIntosh AM, Porteous DJ, Haley CS, Vitart V, Hayward C. Exploration of haplotype research consortium imputation for genome-wide association studies in 20,032 Generation Scotland participants. *Genome Med.* 2017 07;9(1):23. PMID: PMC5339960
56. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, Highland HM, Patel YM, Sorokin EP, Avery CL, Belbin GM, Bien SA, Cheng I, Cullina S, Hodonsky CJ, Hu Y, Huckins LM, Jeff J, Justice AE, Kocarnik JM, Lim U, Lin BM, Lu Y, Nelson SC, Park S-SL, Poisner H, Preuss MH, Richard MA, Schurmann C, Setiawan VW, Sockell A, Vahi K, Verbanck M, Vishnu A, Walker RW, Young KL, Zubair N, Acuña-Alonso V, Ambite JL, Barnes KC, Boerwinkle E, Bottinger EP, Bustamante CD, Caberto C, Canizales-Quinteros S, Conomos MP, Deelman E, Do R, Doheny K, Fernández-Rhodes L, Fornage M, Hailu B, Heiss G, Henn BM, Hindorff LA, Jackson RD, Laurie CA, Laurie CC, Li Y, Lin D-Y, Moreno-Estrada A, Nadkarni G, Norman PJ, Pooler LC, Reiner AP, Romm J, Sabatti C, Sandoval K, Sheng X, Stahl EA, Stram DO, Thornton TA, Wassel CL, Wilkens LR, Winkler CA, Yoneyama S, Buyske S, Haiman CA, Kooperberg C, Marchand LL, Loos RJJ, Matise TC, North KE, Peters U, Kenny EE, Carlson CS. Genetic analyses of diverse populations improves discovery for complex traits. *Nature.* 2019 Jun;570(7762):514–518.
57. Kulminski AM, Huang J, Loika Y, Arbeev KG, Bagley O, Yashkin A, Duan M, Culminskaya I. Strong impact of natural-selection-free heterogeneity in genetics of age-related phenotypes. *Aging (Albany NY).* 2018 29;10(3):492–514. PMID: PMC5892700
58. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A, Chen J, Buchkovich ML, Mora S, Beckmann JS, Bragg-Gresham JL, Chang H-Y, Demirkan A, Den Hertog HM, Do R, Donnelly LA, Ehret GB, Esko T, Feitosa MF, Ferreira T, Fischer K, Fontanillas P, Fraser RM, Freitag DF, Gurdasani D, Heikkilä K, Hyppönen E, Isaacs A, Jackson AU, Johansson Å, Johnson T, Kaakinen M, Kettunen J, Kleber ME, Li X, Luan J, Lyytikäinen L-P, Magnusson PKE, Mangino M, Mihailov E, Montasser ME, Müller-Nurasyid M, Nolte IM, O'Connell JR, Palmer CD, Perola M, Petersen A-K, Sanna S, Saxena R, Service SK, Shah S, Shungin D, Sidore C, Song C, Strawbridge RJ, Surakka I, Tanaka T, Teslovich TM, Thorleifsson G, Van den Herik EG, Voight BF, Volcik KA, Waite LL, Wong A, Wu Y, Zhang W, Absher D, Asiki G, Barroso I, Been LF, Bolton JL, Bonnycastle LL, Brambilla P, Burnett MS, Cesana G, Dimitriou M, Doney ASF, Döring A, Elliott P, Epstein SE, Ingi Eyjolfsson G, Gigante B, Goodarzi MO, Grallert H, Gravitto ML, Groves CJ, Hallmans G, Hartikainen A-L, Hayward C, Hernandez D, Hicks AA, Holm H, Hung Y-J, Illig T, Jones MR, Kaleebu P, Kastelein JJP, Khaw K-T, Kim E, Klopp N, Komulainen P, Kumari M, Langenberg C, Lehtimäki T, Lin S-Y, Lindström J, Loos RJJ, Mach F, McArdle WL, Meisinger C, Mitchell BD, Müller G, Nagaraja R, Narisu N, Nieminen TVM, Nsubuga RN, Olafsson I, Ong KK, Palotie A, Papamarkou T, Pomilla C, Pouta A, Rader DJ, Reilly MP, Ridker PM, Rivadeneira F, Rudan I, Ruokonen A, Samani N, Scharnagl H, Seeley J, Silander K, Stančáková A, Stirrups K, Swift AJ, Tiet L,

- Uitterlinden AG, van Pelt LJ, Vedantam S, Wainwright N, Wijmenga C, Wild SH, Willemsen G, Wilsgaard T, Wilson JF, Young EH, Zhao JH, Adair LS, Arveiler D, Assimes TL, Bandinelli S, Bennett F, Bochud M, Boehm BO, Boomsma DI, Borecki IB, Bornstein SR, Bovet P, Burnier M, Campbell H, Chakravarti A, Chambers JC, Chen Y-DI, Collins FS, Cooper RS, Danesh J, Dedoussis G, de Faire U, Feranil AB, Ferrières J, Ferrucci L, Freimer NB, Gieger C, Groop LC, Gudnason V, Gyllenstein U, Hamsten A, Harris TB, Hingorani A, Hirschhorn JN, Hofman A, Hovingh GK, Hsiung CA, Humphries SE, Hunt SC, Hveem K, Iribarren C, Järvelin M-R, Jula A, Kähönen M, Kaprio J, Kesäniemi A, Kivimäki M, Kooner JS, Koudstaal PJ, Krauss RM, Kuh D, Kuusisto J, Kyvik KO, Laakso M, Lakka TA, Lind L, Lindgren CM, Martin NG, März W, McCarthy MI, McKenzie CA, Meneton P, Metspalu A, Moilanen L, Morris AD, Munroe PB, Njølstad I, Pedersen NL, Power C, Pramstaller PP, Price JF, Psaty BM, Quertermous T, Rauramaa R, Saleheen D, Salomaa V, Sanghera DK, Saramies J, Schwarz PEH, Sheu WH-H, Shuldiner AR, Siegbahn A, Spector TD, Stefansson K, Strachan DP, Tayo BO, Tremoli E, Tuomilehto J, Uusitupa M, van Duijn CM, Vollenweider P, Wallentin L, Wareham NJ, Whitfield JB, Wolffenbuttel BHR, Ordovas JM, Boerwinkle E, Palmer CNA, Thorsteinsdottir U, Chasman DI, Rotter JI, Franks PW, Ripatti S, Cupples LA, Sandhu MS, Rich SS, Boehnke M, Deloukas P, Kathiresan S, Mohlke KL, Ingelsson E, Abecasis GR, Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013 Nov;45(11):1274–1283. PMID: PMC3838666
59. Kilpeläinen TO, Bentley AR, Noordam R, Sung YJ, Schwander K, Winkler TW, Jakupović H, Chasman DI, Manning A, Ntalla I, Aschard H, Brown MR, de Las Fuentes L, Franceschini N, Guo X, Vojinovic D, Aslibekyan S, Feitosa MF, Kho M, Musani SK, Richard M, Wang H, Wang Z, Bartz TM, Bielak LF, Campbell A, Dorajoo R, Fisher V, Hartwig FP, Horimoto ARVR, Li C, Lohman KK, Marten J, Sim X, Smith AV, Tajuddin SM, Alver M, Amini M, Boissel M, Chai JF, Chen X, Divers J, Evangelou E, Gao C, Graff M, Harris SE, He M, Hsu F-C, Jackson AU, Zhao JH, Kraja AT, Kühnel B, Laguzzi F, Lyytikäinen L-P, Nolte IM, Rauramaa R, Riaz M, Robino A, Rueedi R, Stringham HM, Takeuchi F, van der Most PJ, Varga TV, Verweij N, Ware EB, Wen W, Li X, Yanek LR, Amin N, Arnett DK, Boerwinkle E, Brumat M, Cade B, Canouil M, Chen Y-DI, Concas MP, Connell J, de Mutsert R, de Silva HJ, de Vries PS, Demirkan A, Ding J, Eaton CB, Faul JD, Friedlander Y, Gabriel KP, Ghanbari M, Giulianini F, Gu CC, Gu D, Harris TB, He J, Heikkinen S, Heng C-K, Hunt SC, Ikram MA, Jonas JB, Koh W-P, Komulainen P, Krieger JE, Kritchevsky SB, Kutalik Z, Kuusisto J, Langefeld CD, Langenberg C, Launer LJ, Leander K, Lemaitre RN, Lewis CE, Liang J, Lifelines Cohort Study, Liu J, Mägi R, Manichaikul A, Meitinger T, Metspalu A, Milaneschi Y, Mohlke KL, Mosley TH, Murray AD, Nalls MA, Nang E-EK, Nelson CP, Nona S, Norris JM, Nwuba CV, O'Connell J, Palmer ND, Papanicolaou GJ, Pazoki R, Pedersen NL, Peters A, Peyser PA, Polasek O, Porteous DJ, Poveda A, Raitakari OT, Rich SS, Risch N, Robinson JG, Rose LM, Rudan I, Schreiner PJ, Scott RA, Sidney SS, Sims M, Smith JA, Snieder H, Sofer T, Starr JM, Sternfeld B, Strauch K, Tang H, Taylor KD, Tsai MY, Tuomilehto J, Uitterlinden AG, van der Ende MY, van Heemst D, Voortman T, Waldenberger M, Wennberg P, Wilson G, Xiang Y-B, Yao J, Yu C, Yuan J-M, Zhao W, Zonderman AB, Becker DM, Boehnke M, Bowden DW, de Faire U, Deary IJ, Elliott P, Esko T, Freedman BI, Froguel P, Gasparini P, Gieger C, Kato N, Laakso M, Lakka TA, Lehtimäki T, Magnusson PKE, Oldehinkel AJ, Penninx BWJH, Samani NJ, Shu X-O, van der Harst P, Van Vliet-Ostaptchouk JV,

- Vollenweider P, Wagenknecht LE, Wang YX, Wareham NJ, Weir DR, Wu T, Zheng W, Zhu X, Evans MK, Franks PW, Gudnason V, Hayward C, Horta BL, Kelly TN, Liu Y, North KE, Pereira AC, Ridker PM, Tai ES, van Dam RM, Fox ER, Kardina SLR, Liu C-T, Mook-Kanamori DO, Province MA, Redline S, van Duijn CM, Rotter JI, Kooperberg CB, Gauderman WJ, Psaty BM, Rice K, Munroe PB, Fornage M, Cupples LA, Rotimi CN, Morrison AC, Rao DC, Loos RJF. Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. *Nat Commun.* 2019 22;10(1):376. PMID: PMC6342931
60. Morris AP, Le TH, Wu H, Akbarov A, van der Most PJ, Hemani G, Smith GD, Mahajan A, Gaulton KJ, Nadkarni GN, Valladares-Salgado A, Wachter-Rodarte N, Mychaleckyj JC, Dueker ND, Guo X, Hai Y, Haessler J, Kamatani Y, Stilp AM, Zhu G, Cook JP, Ärnlöv J, Blanton SH, de Borst MH, Bottinger EP, Buchanan TA, Cechova S, Charchar FJ, Chu P-L, Damman J, Eales J, Gharavi AG, Giedraitis V, Heath AC, Ipp E, Kiryluk K, Kramer HJ, Kubo M, Larsson A, Lindgren CM, Lu Y, Madden PAF, Montgomery GW, Papanicolaou GJ, Raffel LJ, Sacco RL, Sanchez E, Stark H, Sundstrom J, Taylor KD, Xiang AH, Zivkovic A, Lind L, Ingelsson E, Martin NG, Whitfield JB, Cai J, Laurie CC, Okada Y, Matsuda K, Kooperberg C, Chen Y-DI, Rundek T, Rich SS, Loos RJF, Parra EJ, Cruz M, Rotter JI, Snieder H, Tomaszewski M, Humphreys BD, Franceschini N. Trans-ethnic kidney function association study reveals putative causal genes and effects on kidney-specific disease aetiologies. *Nat Commun.* 2019 03;10(1):29. PMID: PMC6318312
61. Bujakowska KM, Zhang Q, Siemiatkowska AM, Liu Q, Place E, Falk MJ, Consugar M, Lancelot M-E, Antonio A, Lonjou C, Carpentier W, Mohand-Saïd S, den Hollander AI, Cremers FPM, Leroy BP, Gai X, Sahel J-A, van den Born LI, Collin RWJ, Zeitz C, Audo I, Pierce EA. Mutations in IFT172 cause isolated retinal degeneration and Bardet-Biedl syndrome. *Hum Mol Genet.* 2015 Jan 1;24(1):230–242. PMID: PMC4326328
62. Taylor SP, Dantas TJ, Duran I, Wu S, Lachman RS, University of Washington Center for Mendelian Genomics Consortium, Nelson SF, Cohn DH, Vallee RB, Krakow D. Mutations in DYNC2LI1 disrupt cilia function and cause short rib polydactyly syndrome. *Nat Commun.* 2015 Jun 16;6:7092. PMID: PMC4470332
63. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A, Peloso GM, Barbalic M, Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, Chambers J, Orho-Melander M, Melander O, Johnson T, Li X, Guo X, Li M, Shin Cho Y, Jin Go M, Jin Kim Y, Lee J-Y, Park T, Kim K, Sim X, Tzee-Hee Ong R, Croteau-Chonka DC, Lange LA, Smith JD, Song K, Hua Zhao J, Yuan X, Luan J, Lamina C, Ziegler A, Zhang W, Zee RYL, Wright AF, Witteman JCM, Wilson JF, Willemssen G, Wichmann H-E, Whitfield JB, Waterworth DM, Wareham NJ, Waeber G, Vollenweider P, Voight BF, Vitart V, Uitterlinden AG, Uda M, Tuomilehto J, Thompson JR, Tanaka T, Surakka I, Stringham HM, Spector TD, Soranzo N, Smit JH, Sinisalo J, Silander K, Sijbrands EJG, Scuteri A, Scott J, Schlessinger D, Sanna S, Salomaa V, Saharinen J, Sabatti C, Ruukonen A, Rudan I, Rose LM, Roberts R, Rieder M, Psaty BM, Pramstaller PP, Pichler I, Perola M, Penninx BWJH, Pedersen NL, Pattaro C, Parker AN, Pare G, Oostra BA, O'Donnell CJ, Nieminen MS, Nickerson DA, Montgomery GW, Meitinger T, McPherson R, McCarthy

- MI, McArdle W, Masson D, Martin NG, Marroni F, Mangino M, Magnusson PKE, Lucas G, Luben R, Loos RJF, Lokki M-L, Lettre G, Langenberg C, Launer LJ, Lakatta EG, Laaksonen R, Kyvik KO, Kronenberg F, König IR, Khaw K-T, Kaprio J, Kaplan LM, Johansson A, Jarvelin M-R, Janssens ACJW, Ingelsson E, Igl W, Kees Hovingh G, Hottenga J-J, Hofman A, Hicks AA, Hengstenberg C, Heid IM, Hayward C, Havulinna AS, Hastie ND, Harris TB, Haritunians T, Hall AS, Gyllenstein U, Guiducci C, Groop LC, Gonzalez E, Gieger C, Freimer NB, Ferrucci L, Erdmann J, Elliott P, Ejebe KG, Döring A, Dominiczak AF, Demissie S, Deloukas P, de Geus EJC, de Faire U, Crawford G, Collins FS, Chen YI, Caulfield MJ, Campbell H, Burt NP, Bonnycastle LL, Boomsma DI, Boekholdt SM, Bergman RN, Barroso I, Bandinelli S, Ballantyne CM, Assimes TL, Quertermous T, Altshuler D, Seielstad M, Wong TY, Tai E-S, Feranil AB, Kuzawa CW, Adair LS, Taylor HA, Borecki IB, Gabriel SB, Wilson JG, Holm H, Thorsteinsdottir U, Gudnason V, Krauss RM, Mohlke KL, Ordovas JM, Munroe PB, Kooner JS, Tall AR, Hegele RA, Kastelein JJP, Schadt EE, Rotter JI, Boerwinkle E, Strachan DP, Mooser V, Stefansson K, Reilly MP, Samani NJ, Schunkert H, Cupples LA, Sandhu MS, Ridker PM, Rader DJ, van Duijn CM, Peltonen L, Abecasis GR, Boehnke M, Kathiresan S. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010 Aug 5;466(7307):707–713. PMID: PMC3039276
64. Manning AK, Hivert M-F, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, Rybin D, Liu C-T, Bielak LF, Prokopenko I, Amin N, Barnes D, Cadby G, Hottenga J-J, Ingelsson E, Jackson AU, Johnson T, Kanoni S, Ladenvall C, Lagou V, Lahti J, Lecoeur C, Liu Y, Martinez-Larrad MT, Montasser ME, Navarro P, Perry JRB, Rasmussen-Torvik LJ, Salo P, Sattar N, Shungin D, Strawbridge RJ, Tanaka T, van Duijn CM, An P, de Andrade M, Andrews JS, Aspelund T, Atalay M, Aulchenko Y, Balkau B, Bandinelli S, Beckmann JS, Beilby JP, Bellis C, Bergman RN, Blangero J, Boban M, Boehnke M, Boerwinkle E, Bonnycastle LL, Boomsma DI, Borecki IB, Böttcher Y, Bouchard C, Brunner E, Budimir D, Campbell H, Carlson O, Chines PS, Clarke R, Collins FS, Corbatón-Anchuelo A, Couper D, de Faire U, Dedoussis GV, Deloukas P, Dimitriou M, Egan JM, Eiriksdottir G, Erdos MR, Eriksson JG, Eury E, Ferrucci L, Ford I, Forouhi NG, Fox CS, Franzosi MG, Franks PW, Frayling TM, Froguel P, Galan P, de Geus E, Gigante B, Glazer NL, Goel A, Groop L, Gudnason V, Hallmans G, Hamsten A, Hansson O, Harris TB, Hayward C, Heath S, Herberg S, Hicks AA, Hingorani A, Hofman A, Hui J, Hung J, Jarvelin M-R, Jhun MA, Johnson PCD, Jukema JW, Jula A, Kao WH, Kaprio J, Kardia SLR, Keinanen-Kiukkaanniemi S, Kivimäki M, Kolcic I, Kovacs P, Kumari M, Kuusisto J, Kyvik KO, Laakso M, Lakka T, Lannfelt L, Lathrop GM, Launer LJ, Leander K, Li G, Lind L, Lindstrom J, Lobbens S, Loos RJF, Luan J, Lyssenko V, Mägi R, Magnusson PKE, Marmot M, Meneton P, Mohlke KL, Mooser V, Morken MA, Miljkovic I, Narisu N, O'Connell J, Ong KK, Oostra BA, Palmer LJ, Palotie A, Pankow JS, Peden JF, Pedersen NL, Pehlic M, Peltonen L, Penninx B, Pericic M, Perola M, Perusse L, Peyser PA, Polasek O, Pramstaller PP, Province MA, Rääkkönen K, Rauramaa R, Rehnberg E, Rice K, Rotter JI, Rudan I, Ruukonen A, Saaristo T, Sabater-Lleal M, Salomaa V, Savage DB, Saxena R, Schwarz P, Seedorf U, Sennblad B, Serrano-Rios M, Shuldiner AR, Sijbrands EJG, Siscovick DS, Smit JH, Small KS, Smith NL, Smith AV, Stančáková A, Stirrups K, Stumvoll M, Sun YV, Swift AJ, Tönjes A, Tuomilehto J, Trompet S, Uitterlinden AG, Uusitupa M, Vikström M, Vitart V, Vohl M-C, Voight BF, Vollenweider P, Waechter G, Waterworth DM, Watkins H, Wheeler E, Widen E, Wild SH, Willems SM, Willemsen G,

- Wilson JF, Witteman JCM, Wright AF, Yaghootkar H, Zelenika D, Zemunik T, Zgaga L, DIABetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Multiple Tissue Human Expression Resource (MUTHER) Consortium, Wareham NJ, McCarthy MI, Barroso I, Watanabe RM, Florez JC, Dupuis J, Meigs JB, Langenberg C. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet.* 2012 May 13;44(6):659–669. PMID: PMC3613127
65. Wessel J, Chu AY, Willems SM, Wang S, Yaghootkar H, Brody JA, Dauriz M, Hivert M-F, Raghavan S, Lipovich L, Hidalgo B, Fox K, Huffman JE, An P, Lu Y, Rasmussen-Torvik LJ, Grarup N, Ehm MG, Li L, Baldrige AS, Stančáková A, Abrol R, Besse C, Boland A, Bork-Jensen J, Fornage M, Freitag DF, Garcia ME, Guo X, Hara K, Isaacs A, Jakobsdottir J, Lange LA, Layton JC, Li M, Hua Zhao J, Meidtner K, Morrison AC, Nalls MA, Peters MJ, Sabater-Lleal M, Schurmann C, Silveira A, Smith AV, Southam L, Stoiber MH, Strawbridge RJ, Taylor KD, Varga TV, Allin KH, Amin N, Aponte JL, Aung T, Barbieri C, Bihlmeyer NA, Boehnke M, Bombieri C, Bowden DW, Burns SM, Chen Y, Chen Y-D, Cheng C-Y, Correa A, Czajkowski J, Dehghan A, Ehret GB, Eiriksdottir G, Escher SA, Farmaki A-E, Fränberg M, Gambaro G, Giulianini F, Goddard WA, Goel A, Gottesman O, Grove ML, Gustafsson S, Hai Y, Hallmans G, Heo J, Hoffmann P, Ikram MK, Jensen RA, Jørgensen ME, Jørgensen T, Karaleftheri M, Khor CC, Kirkpatrick A, Kraja AT, Kuusisto J, Lange EM, Lee IT, Lee W-J, Leong A, Liao J, Liu C, Liu Y, Lindgren CM, Linneberg A, Malerba G, Mamakou V, Marouli E, Maruthur NM, Matchan A, McKean-Cowdin R, McLeod O, Metcalf GA, Mohlke KL, Muzny DM, Ntalla I, Palmer ND, Pasko D, Peter A, Rayner NW, Renström F, Rice K, Sala CF, Sennblad B, Serafetinidis I, Smith JA, Soranzo N, Speliotes EK, Stahl EA, Stirrups K, Tentolouris N, Thanopoulou A, Torres M, Traglia M, Tsaftakakis E, Javad S, Yanek LR, Zengini E, Becker DM, Bis JC, Brown JB, Cupples LA, Hansen T, Ingelsson E, Karter AJ, Lorenzo C, Mathias RA, Norris JM, Peloso GM, Sheu WH-H, Toniolo D, Vaidya D, Varma R, Wagenknecht LE, Boeing H, Bottinger EP, Dedoussis G, Deloukas P, Ferrannini E, Franco OH, Franks PW, Gibbs RA, Gudnason V, Hamsten A, Harris TB, Hattersley AT, Hayward C, Hofman A, Jansson J-H, Langenberg C, Launer LJ, Levy D, Oostra BA, O'Donnell CJ, O'Rahilly S, Padmanabhan S, Pankow JS, Polasek O, Province MA, Rich SS, Ridker PM, Rudan I, Schulze MB, Smith BH, Uitterlinden AG, Walker M, Watkins H, Wong TY, Zeggini E, EPIC-InterAct Consortium, Laakso M, Borecki IB, Chasman DI, Pedersen O, Psaty BM, Tai ES, van Duijn CM, Wareham NJ, Waterworth DM, Boerwinkle E, Kao WHL, Florez JC, Loos RJF, Wilson JG, Frayling TM, Siscovick DS, Dupuis J, Rotter JJ, Meigs JB, Scott RA, Goodarzi MO. Low-frequency and rare exome chip variants associate with fasting glucose and type 2 diabetes susceptibility. *Nat Commun.* 2015 Jan 29;6:5897. PMID: PMC4311266
66. Mahajan A, Rodan AR, Le TH, Gaulton KJ, Haessler J, Stilp AM, Kamatani Y, Zhu G, Sofer T, Puri S, Schellinger JN, Chu P-L, Cechova S, van Zuydam N, SUMMIT Consortium, BioBank Japan Project, Arnlov J, Flessner MF, Giedraitis V, Heath AC, Kubo M, Larsson A, Lindgren CM, Madden PAF, Montgomery GW, Papanicolaou GJ, Reiner AP, Sundström J, Thornton TA, Lind L, Ingelsson E, Cai J, Martin NG, Kooperberg C, Matsuda K, Whitfield JB, Okada Y, Laurie CC, Morris AP, Franceschini N. Trans-ethnic Fine Mapping Highlights Kidney-Function Genes Linked to Salt Sensitivity. *Am J Hum Genet.* 2016 Sep 1;99(3):636–646. PMID: PMC5011075

67. Moon S, Kim YJ, Han S, Hwang MY, Shin DM, Park MY, Lu Y, Yoon K, Jang H-M, Kim YK, Park T-J, Song DS, Park JK, Lee J-E, Kim B-J. The Korea Biobank Array: Design and Identification of Coding Variants Associated with Blood Biochemical Traits. *Sci Rep*. 2019 04;9(1):1382. PMID: PMC6361960
68. Olafsson S, Alexandersson KF, Gizurarson JGK, Hauksdottir K, Gunnarsson O, Olafsson K, Gudmundsson J, Stacey SN, Sveinbjornsson G, Saemundsdottir J, Bjornsson ES, Olafsson S, Bjornsson S, Orvar KB, Vikingsson A, Geirsson AJ, Arinbjarnarson S, Bjornsdottir G, Thorgeirsson TE, Sigurdsson S, Halldorsson GH, Magnusson OT, Masson G, Holm H, Jonsdottir I, Sigurdardottir O, Eyjolfsson GI, Olafsson I, Sulem P, Thorsteinsdottir U, Jonsson T, Rafnar T, Gudbjartsson DF, Stefansson K. Common and Rare Sequence Variants Influencing Tumor Biomarkers in Blood. *Cancer Epidemiol Biomarkers Prev*. 2020 Jan;29(1):225–235. PMID: PMC6954334
69. Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, Van der Harst P, Holm H, Sanna S, Kavousi M, Baumeister SE, Coin LJ, Deng G, Gieger C, Heard-Costa NL, Hottenga J-J, Kühnel B, Kumar V, Lagou V, Liang L, Luan J, Vidal PM, Mateo Leach I, O'Reilly PF, Peden JF, Rahmioglu N, Soininen P, Speliotes EK, Yuan X, Thorleifsson G, Alizadeh BZ, Atwood LD, Borecki IB, Brown MJ, Charoen P, Cucca F, Das D, de Geus EJC, Dixon AL, Döring A, Ehret G, Eyjolfsson GI, Farrall M, Forouhi NG, Friedrich N, Goessling W, Gudbjartsson DF, Harris TB, Hartikainen A-L, Heath S, Hirschfield GM, Hofman A, Homuth G, Hyppönen E, Janssen HLA, Johnson T, Kangas AJ, Kema IP, Kühn JP, Lai S, Lathrop M, Lerch MM, Li Y, Liang TJ, Lin J-P, Loos RJF, Martin NG, Moffatt MF, Montgomery GW, Munroe PB, Musunuru K, Nakamura Y, O'Donnell CJ, Olafsson I, Penninx BW, Pouta A, Prins BP, Prokopenko I, Puls R, Ruokonen A, Savolainen MJ, Schlessinger D, Schouten JNL, Seedorf U, Sen-Chowdhry S, Siminovitch KA, Smit JH, Spector TD, Tan W, Teslovich TM, Tukiainen T, Uitterlinden AG, Van der Klauw MM, Vasani RS, Wallace C, Wallaschofski H, Wichmann H-E, Willemsen G, Würtz P, Xu C, Yerges-Armstrong LM, Alcohol Genome-wide Association (AlcGen) Consortium, Diabetes Genetics Replication and Meta-analyses (DIAGRAM+) Study, Genetic Investigation of Anthropometric Traits (GIANT) Consortium, Global Lipids Genetics Consortium, Genetics of Liver Disease (GOLD) Consortium, International Consortium for Blood Pressure (ICBP-GWAS), Meta-analyses of Glucose and Insulin-Related Traits Consortium (MAGIC), Abecasis GR, Ahmadi KR, Boomsma DI, Caulfield M, Cookson WO, van Duijn CM, Froguel P, Matsuda K, McCarthy MI, Meisinger C, Mooser V, Pietiläinen KH, Schumann G, Snieder H, Sternberg MJE, Stolk RP, Thomas HC, Thorsteinsdottir U, Uda M, Waeber G, Wareham NJ, Waterworth DM, Watkins H, Whitfield JB, Witteman JCM, Wolffenbuttel BHR, Fox CS, Ala-Korpela M, Stefansson K, Vollenweider P, Völzke H, Schadt EE, Scott J, Jarvelin M-R, Elliott P, Kooner JS. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nat Genet*. 2011 Oct 16;43(11):1131–1138. PMID: PMC3482372
70. Niceta M, Margiotti K, Digilio MC, Guida V, Bruselles A, Pizzi S, Ferraris A, Memo L, Laforgia N, Dentici ML, Consoli F, Torrente I, Ruiz-Perez VL, Dallapiccola B, Marino B, De Luca A, Tartaglia M. Biallelic mutations in *DYNC2LI1* are a rare cause of Ellis-van Creveld syndrome. *Clin Genet*. 2018;93(3):632–639. PMID: 28857138

71. Yamada M, Uehara T, Suzuki H, Takenouchi T, Fukushima H, Morisada N, Tominaga K, Onoda M, Kosaki K. IFT172 as the 19th gene causative of oral-facial-digital syndrome. *Am J Med Genet A*. 2019;179(12):2510–2513. PMID: 31587445
72. Halbritter J, Bizet AA, Schmidts M, Porath JD, Braun DA, Gee HY, McInerney-Leo AM, Krug P, Filhol E, Davis EE, Airik R, Czarnecki PG, Lehman AM, Trnka P, Nitschké P, Bole-Feysot C, Schueler M, Knebelmann B, Burtey S, Szabó AJ, Tory K, Leo PJ, Gardiner B, McKenzie FA, Zankl A, Brown MA, Hartley JL, Maher ER, Li C, Leroux MR, Scambler PJ, Zhan SH, Jones SJ, Kayserili H, Tuysuz B, Moorani KN, Constantinescu A, Krantz ID, Kaplan BS, Shah JV, UK10K Consortium, Hurd TW, Doherty D, Katsanis N, Duncan EL, Otto EA, Beales PL, Mitchison HM, Saunier S, Hildebrandt F. Defects in the IFT-B component IFT172 cause Jeune and Mainzer-Saldino syndromes in humans. *Am J Hum Genet*. 2013 Nov 7;93(5):915–925. PMCID: PMC3824130
73. Ciesielski TH, Pendergrass SA, White MJ, Kodaman N, Sobota RS, Huang M, Bartlett J, Li J, Pan Q, Gui J, Selleck SB, Amos CI, Ritchie MD, Moore JH, Williams SM. Diverse convergent evidence in the genetic analysis of complex disease: coordinating omic, informatic, and experimental evidence to better identify and validate risk factors. *BioData Mining*. 2014 Jun 30;7(1):10.
74. Filges I, Bruder E, Brandal K, Meier S, Undlien DE, Waage TR, Hoesli I, Schubach M, de Beer T, Sheng Y, Hoeller S, Schulzke S, Røsbj O, Miny P, Tercanli S, Oppedal T, Meyer P, Selmer KK, Strømme P. Strømme Syndrome Is a Ciliary Disorder Caused by Mutations in CENPF. *Hum Mutat*. 2016 Apr;37(4):359–363. PMID: 26820108
75. Waters AM, Asfahani R, Carroll P, Bicknell L, Lescai F, Bright A, Chanudet E, Brooks A, Christou-Savina S, Osman G, Walsh P, Bacchelli C, Chapgier A, Vernay B, Bader DM, Deshpande C, O' Sullivan M, Ocaka L, Stanescu H, Stewart HS, Hildebrandt F, Otto E, Johnson CA, Szymanska K, Katsanis N, Davis E, Kleta R, Hubank M, Doxsey S, Jackson A, Stupka E, Winey M, Beales PL. The kinetochore protein, CENPF, is mutated in human ciliopathy and microcephaly phenotypes. *J Med Genet*. 2015 Mar;52(3):147–156. PMCID: PMC4345935
76. Saari J, Lovell MA, Yu H-C, Bellus GA. Compound heterozygosity for a frame shift mutation and a likely pathogenic sequence variant in the planar cell polarity—ciliogenesis gene WPCP in a girl with polysyndactyly, coarctation of the aorta, and tongue hamartomas. *Am J Med Genet A*. 2015 Feb;167A(2):421–427. PMID: 25427950
77. Davidson AE, Sergouniotis PI, Mackay DS, Wright GA, Waseem NH, Michaelides M, Holder GE, Robson AG, Moore AT, Plagnol V, Webster AR. RP1L1 variants are associated with a spectrum of inherited retinal diseases including retinitis pigmentosa and occult macular dystrophy. *Hum Mutat*. 2013 Mar;34(3):506–514. PMID: 23281133
78. Yamashita T, Liu J, Gao J, LeNoue S, Wang C, Kaminoh J, Bowne SJ, Sullivan LS, Daiger SP, Zhang K, Fitzgerald MEC, Kefalov VJ, Zuo J. Essential and Synergistic Roles of RP1 and RP1L1 in Rod Photoreceptor Axoneme and Retinitis Pigmentosa. *J Neurosci*. Society for Neuroscience; 2009 Aug 5;29(31):9748–9760. PMID: 19657028

79. Rees MG, Ng D, Ruppert S, Turner C, Beer NL, Swift AJ, Morken MA, Below JE, Blech I, NISC Comparative Sequencing Program, Mullikin JC, McCarthy MI, Biesecker LG, Gloyne AL, Collins FS. Correlation of rare coding variants in the gene encoding human glucokinase regulatory protein with phenotypic, cellular, and kinetic outcomes. *J Clin Invest.* 2012 Jan;122(1):205–217. PMID: PMC3248284
80. Orho-Melander M, Melander O, Guiducci C, Perez-Martinez P, Corella D, Roos C, Tewhey R, Rieder MJ, Hall J, Abecasis G, Tai ES, Welch C, Arnett DK, Lyssenko V, Lindholm E, Saxena R, de Bakker PIW, Burt N, Voight BF, Hirschhorn JN, Tucker KL, Hedner T, Tuomi T, Isomaa B, Eriksson K-F, Taskinen M-R, Wahlstrand B, Hughes TE, Parnell LD, Lai C-Q, Berglund G, Peltonen L, Vartiainen E, Jousilahti P, Havulinna AS, Salomaa V, Nilsson P, Groop L, Altshuler D, Ordovas JM, Kathiresan S. Common Missense Variant in the Glucokinase Regulatory Protein Gene Is Associated With Increased Plasma Triglyceride and C-Reactive Protein but Lower Fasting Glucose Concentrations. *Diabetes.* 2008 Nov;57(11):3112–3121. PMID: PMC2570409
81. Grimsby J, Coffey JW, Dvornicki MT, Magram J, Li G, Matschinsky FM, Shiota C, Kaur S, Magnuson MA, Grippo JF. Characterization of Glucokinase Regulatory Protein-deficient Mice. *J Biol Chem. American Society for Biochemistry and Molecular Biology;* 2000 Mar 17;275(11):7826–7831. PMID: 10713097
82. Farrelly D, Brown KS, Tieman A, Ren J, Lira SA, Hagan D, Gregg R, Mookhtiar KA, Hariharan N. Mice mutant for glucokinase regulatory protein exhibit decreased liver glucokinase: a sequestration mechanism in metabolic regulation. *Proc Natl Acad Sci USA.* 1999 Dec 7;96(25):14511–14516. PMID: PMC24467
83. Hiskett EK, Suwitheechon O-U, Lindbloom-Hawley S, Boyle DL, Schermerhorn T. Lack of glucokinase regulatory protein expression may contribute to low glucokinase activity in feline liver. *Vet Res Commun.* 2009 Mar;33(3):227–240. PMID: 18780155
84. Brown KS, Kalinowski SS, Megill JR, Durham SK, Mookhtiar KA. Glucokinase regulatory protein may interact with glucokinase in the hepatocyte nucleus. *Diabetes.* 1997 Feb;46(2):179–186. PMID: 9000692
85. Carli JFM, LeDuc CA, Zhang Y, Stratigopoulos G, Leibel RL. The role of Rpgrip11, a component of the primary cilium, in adipocyte development and function. *FASEB J.* 2018;32(7):3946–3956. PMID: PMC5998974
86. Stratigopoulos G, Burnett LC, Rausch R, Gill R, Penn DB, Skowronski AA, LeDuc CA, Lanzano AJ, Zhang P, Storm DR, Egli D, Leibel RL. Hypomorphism of Fto and Rpgrip11 causes obesity in mice. *J Clin Invest.* 2016 02;126(5):1897–1910. PMID: PMC4855930
87. Vogan K. RPGRIP1L , FTO and obesity. *Nature Genetics.* Nature Publishing Group; 2014 Jun;46(6):532–532.
88. Arnaiz O, Malinowska A, Klotz C, Sperling L, Dadlez M, Koll F, Cohen J. Cildb: a knowledgebase for centrosomes and cilia. *Database (Oxford).* 2009;2009:bap022. PMID: PMC2860946

89. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). Online Mendelian Inheritance in Man, OMIM®. [Internet]. Online Mendelian Inheritance in Man, OMIM®. Available from: World Wide Web URL: <https://omim.org/>
90. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203–209. PMID: PMC6786975
91. Box GEP, Cox DR. An analysis of transformations. *Journal of the Royal Statistical Society Series B (Methodological)*. 1964;211–252.
92. Grady BJ, Torstenson E, Dudek SM, Giles J, Sexton D, Ritchie MD. FINDING UNIQUE FILTER SETS IN PLATO: A PRECURSOR TO EFFICIENT INTERACTION ANALYSIS IN GWAS DATA. *Pac Symp Biocomput*. 2010;315–326. PMID: PMC2903053
93. Hall MA, Wallace J, Lucas A, Kim D, Basile AO, Verma SS, McCarty CA, Brilliant MH, Peissig PL, Kitchner TE, Verma A, Pendergrass SA, Dudek SM, Moore JH, Ritchie MD. PLATO software provides analytic framework for investigating complexity beyond genome-wide association studies. *Nat Commun* [Internet]. 2017 Oct 27 [cited 2020 Feb 7];8. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5660079/> PMID: PMC5660079
94. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010 Sep 1;26(17):2190–2191. PMID: PMC2922887
95. Verma SS, Lucas AM, Lavage DR, Leader JB, Metpally R, Krishnamurthy S, Dewey F, Borecki I, Lopez A, Overton J, Penn J, Reid J, Pendergrass SA, Breitwieser G, Ritchie MD. IDENTIFYING GENETIC ASSOCIATIONS WITH VARIABILITY IN METABOLIC HEALTH AND BLOOD COUNT LABORATORY VALUES: DIVING INTO THE QUANTITATIVE TRAITS BY LEVERAGING LONGITUDINAL DATA FROM AN EHR. *Pac Symp Biocomput*. 2017;22:533–544. PMID: 27897004
96. Prins BP, Kuchenbaecker KB, Bao Y, Smart M, Zabaneh D, Fatemifar G, Luan J, Wareham NJ, Scott RA, Perry JRB, Langenberg C, Benzeval M, Kumari M, Zeggini E. Genome-wide analysis of health-related biomarkers in the UK Household Longitudinal Study reveals novel associations. *Sci Rep*. 2017 Sep 8;7(1):1–9.
97. Dennis JK, Sealock JM, Straub P, Hucks D, Actkins K, Faucon A, Goleva SB, Nirachou M, Singh K, Morley T, Ruderfer DM, Mosley JD, Chen G, Davis LK. Lab-wide association scan of polygenic scores identifies biomarkers of complex disease [Internet]. *Genetic and Genomic Medicine*; 2020 Jan. Available from: <http://medrxiv.org/lookup/doi/10.1101/2020.01.24.20018713>
98. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Mägi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H,

Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JRB, Egan JM, Lajunen T, Grarup N, Sparso T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proença C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O'Connell J, Luan J, Elliott A, McCarroll SA, Payne F, Roccascaccia RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Böttcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen Y-DI, Chines P, Clarke R, Coin LJM, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day INM, de Geus EJC, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllenstein U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen A-L, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PRV, Jørgensen T, Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimäki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley R, Mangino M, Manning AK, Martínez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orrù M, Pakyz R, Palmer CNA, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AFH, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S, Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurdsson G, Sijbrands EJG, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvänen A-C, Tanaka T, Thorand B, Tichet J, Tönjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemssen G, Witteman JCM, Yarnell JWG, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, Borecki IB, Loos RJF, Meneton P, Magnusson PKE, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis GV, Serrano-Ríos M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WHL, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature Genetics*. Nature Publishing Group; 2010 Feb;42(2):105–116.

99. Gurdasani D, Carstensen T, Fatumo S, Chen G, Franklin CS, Prado-Martinez J, Bouman H, Abascal F, Haber M, Tachmazidou I, Mathieson I, Ekoru K, DeGorter MK, Nsubuga RN, Finan C, Wheeler E, Chen L, Cooper DN, Schiffels S, Chen Y, Ritchie GRS, Pollard MO, Fortune MD, Mentzer AJ, Garrison E, Bergström A, Hatzikotoulas K, Adeyemo A, Doumatey A, Elding H, Wain LV, Ehret G, Auer PL, Kooperberg CL, Reiner AP, Franceschini N, Maher D, Montgomery SB, Kadie C, Widmer C, Xue Y, Seeley J, Asiki G, Kamali A, Young EH, Pomilla C, Soranzo N, Zeggini E, Pirie F, Morris AP, Heckerman D, Tyler-Smith C, Motala AA, Rotimi C, Kaleebu P, Barroso I, Sandhu MS. Uganda Genome Resource Enables Insights into Population History and Genomic Discovery in Africa. *Cell*. 2019 Oct 31;179(4):984-1002.e36.

100. Wheeler E, Leong A, Liu C-T, Hivert M-F, Strawbridge RJ, Podmore C, Li M, Yao J, Sim X, Hong J, Chu AY, Zhang W, Wang X, Chen P, Maruthur NM, Porneala BC, Sharp SJ, Jia Y, Kabagambe EK, Chang L-C, Chen W-M, Elks CE, Evans DS, Fan Q, Giulianini F, Go MJ, Hottenga J-J, Hu Y, Jackson AU, Kanoni S, Kim YJ, Kleber ME, Ladenvall C, Lecoeur C, Lim S-H, Lu Y, Mahajan A, Marzi C, Nalls MA, Navarro P, Nolte IM, Rose LM, Rybin DV, Sanna S, Shi Y, Stram DO, Takeuchi F, Tan SP, van der Most PJ, Van Vliet-Ostaptchouk JV, Wong A, Yengo L, Zhao W, Goel A, Martinez Larrad MT, Radke D, Salo P, Tanaka T, van Iperen EPA, Abecasis G, Afaq S, Alizadeh BZ, Bertoni AG, Bonnefond A, Böttcher Y, Bottinger EP, Campbell H, Carlson OD, Chen C-H, Cho YS, Garvey WT, Gieger C, Goodarzi MO, Grallert H, Hamsten A, Hartman CA, Herder C, Hsiung CA, Huang J, Igase M, Isona M, Katsuya T, Khor C-C, Kiess W, Kohara K, Kovacs P, Lee J, Lee W-J, Lehne B, Li H, Liu J, Lobbens S, Luan J, Lyssenko V, Meitinger T, Miki T, Miljkovic I, Moon S, Mulas A, Müller G, Müller-Nurasyid M, Nagaraja R, Nauck M, Pankow JS, Polasek O, Prokopenko I, Ramos PS, Rasmussen-Torvik L, Rathmann W, Rich SS, Robertson NR, Roden M, Roussel R, Rudan I, Scott RA, Scott WR, Sennblad B, Siscovick DS, Strauch K, Sun L, Swertz M, Tajuddin SM, Taylor KD, Teo Y-Y, Tham YC, Tönjes A, Wareham NJ, Willemsen G, Wilsgaard T, Hingorani AD, EPIC-CVD Consortium, EPIC-InterAct Consortium, Lifelines Cohort Study, Egan J, Ferrucci L, Hovingh GK, Jula A, Kivimaki M, Kumari M, Njølstad I, Palmer CNA, Serrano Ríos M, Stumvoll M, Watkins H, Aung T, Blüher M, Boehnke M, Boomsma DI, Bornstein SR, Chambers JC, Chasman DI, Chen Y-DI, Chen Y-T, Cheng C-Y, Cucca F, de Geus EJC, Deloukas P, Evans MK, Fornage M, Friedlander Y, Froguel P, Groop L, Gross MD, Harris TB, Hayward C, Heng C-K, Ingelsson E, Kato N, Kim B-J, Koh W-P, Kooner JS, Körner A, Kuh D, Kuusisto J, Laakso M, Lin X, Liu Y, Loos RJF, Magnusson PKE, März W, McCarthy MI, Oldehinkel AJ, Ong KK, Pedersen NL, Pereira MA, Peters A, Ridker PM, Sabanayagam C, Sale M, Saleheen D, Saltevo J, Schwarz PE, Sheu WHH, Snieder H, Spector TD, Tabara Y, Tuomilehto J, van Dam RM, Wilson JG, Wilson JF, Wolffenbuttel BHR, Wong TY, Wu J-Y, Yuan J-M, Zonderman AB, Soranzo N, Guo X, Roberts DJ, Florez JC, Sladek R, Dupuis J, Morris AP, Tai E-S, Selvin E, Rotter JJ, Langenberg C, Barroso I, Meigs JB. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS Med.* 2017 Sep;14(9):e1002383. PMID: PMC5595282
101. Lucas A. anastasia-lucas/hudson [Internet]. 2020 [cited 2020 Jun 16]. Available from: <https://github.com/anastasia-lucas/hudson>
102. Wickham H. ggplot2: Elegant Graphics for Data Analysis [Internet]. 2nd ed. Springer International Publishing; 2016 [cited 2020 Jun 16]. Available from: <https://www.springer.com/gp/book/9783319242750>
103. Hahne F, Ivanek R. Visualizing Genomic Data Using Gviz and Bioconductor. In: Mathé E, Davis S, editors. Statistical Genomics: Methods and Protocols [Internet]. New York, NY: Springer; 2016 [cited 2020 Jun 17]. p. 335–351. Available from: https://doi.org/10.1007/978-1-4939-3578-9_16

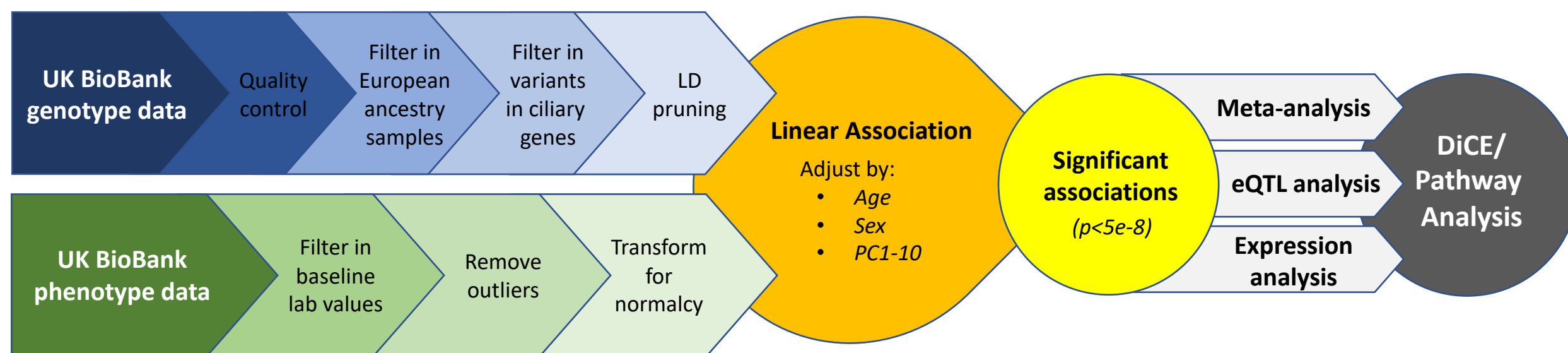


Figure 1. Schematic of analysis workflow for the data presented

Phenotype and genotype data from the UKBB release version 2 was extracted and processed as indicated. Linear association studies between ciliary gene variants and patient phenotypes was performed, with genome-wide significant associations ($p < 5e-8$) being studied further by tissue-specific expression, eQTL, and replication meta-analysis. Using this data, a DiCE/Pathway analysis was performed to identify ciliary subcompartments associated with specific traits.

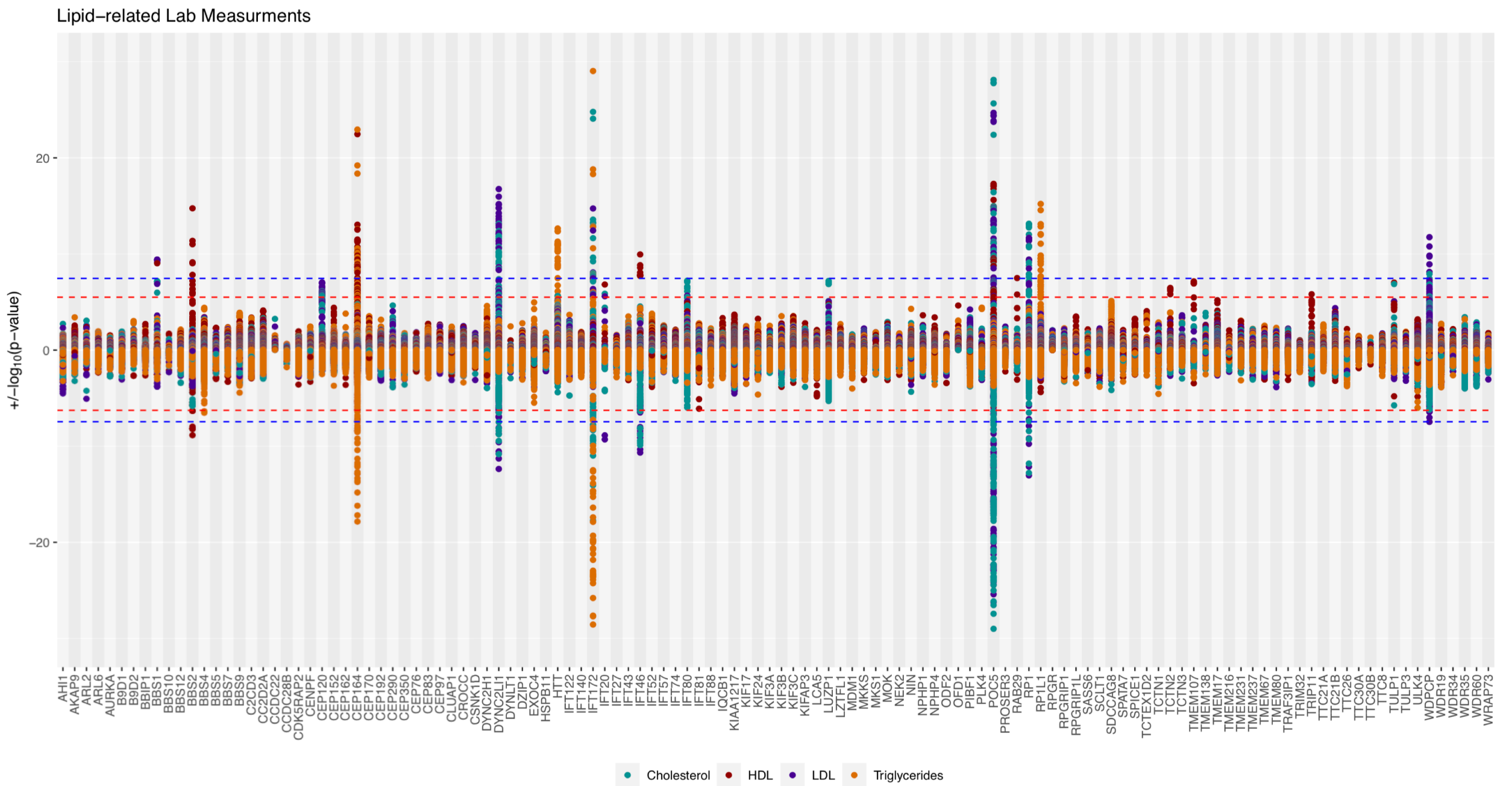


Figure 2. Association study and meta-analysis results for lipid-related traits

Hudson plot illustrating the results of the discovery association analysis (top) and replication meta-analysis (bottom) of common variants within 122 ciliary gene with lipid-related traits. Data for cholesterol is in teal, HDL in red, LDL in purple, and triglycerides in orange. The study-wide Bonferroni-adjusted significance threshold ($p < 3.4 \times 10^{-8}$) is shown as a blue dashed line, while the experiment-wide Bonferroni-adjusted significance threshold ($p < 3.0 \times 10^{-6}$ for the discovery analysis, $p < 3.3 \times 10^{-6}$ for the meta-analysis) is shown as a red dashed line. Each analyzed gene is given equal space along the horizontal axis, with all tested genetic variants for a given gene plotted at the midline of the gene block.

Kidney-related Lab Measurements

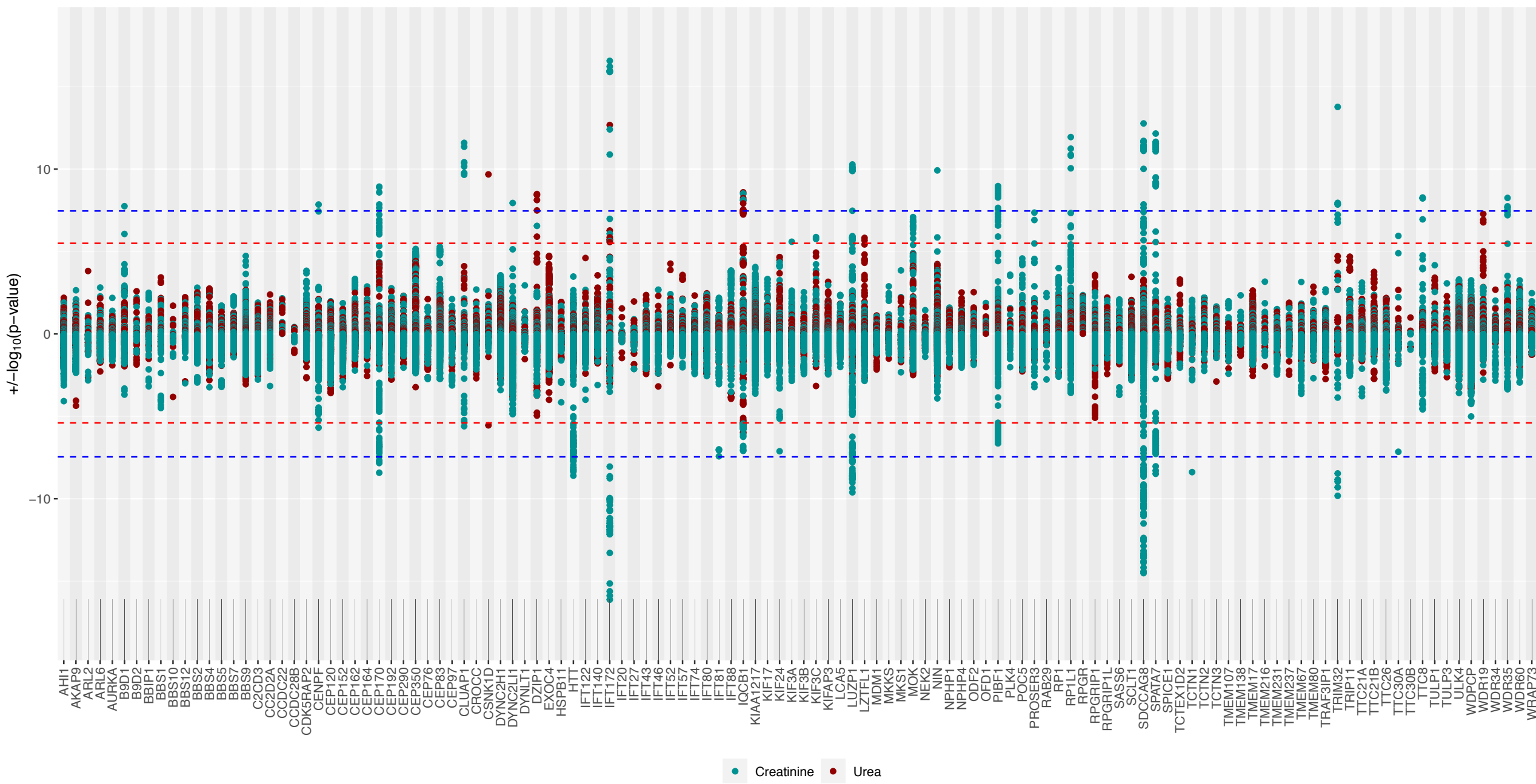


Figure 3. Association study and meta-analysis results for kidney-related traits

Hudson plot illustrating the results of the discovery association analysis (top) and replication meta-analysis (bottom) of common variants within 122 ciliary gene with kidney-related traits. Data for creatinine is in teal, and data for urea is in red. The study-wide Bonferroni-adjusted significance threshold ($p < 3.4 \times 10^{-8}$) is shown as a blue dashed line, while the experiment-wide Bonferroni-adjusted significance threshold ($p < 3.0 \times 10^{-6}$ for the discovery analysis, $p < 3.8 \times 10^{-6}$ for the meta-analysis) is shown as a red dashed line. Each analyzed gene is given equal space along the horizontal axis, with all tested genetic variants for a given gene plotted at the midline of the gene block.

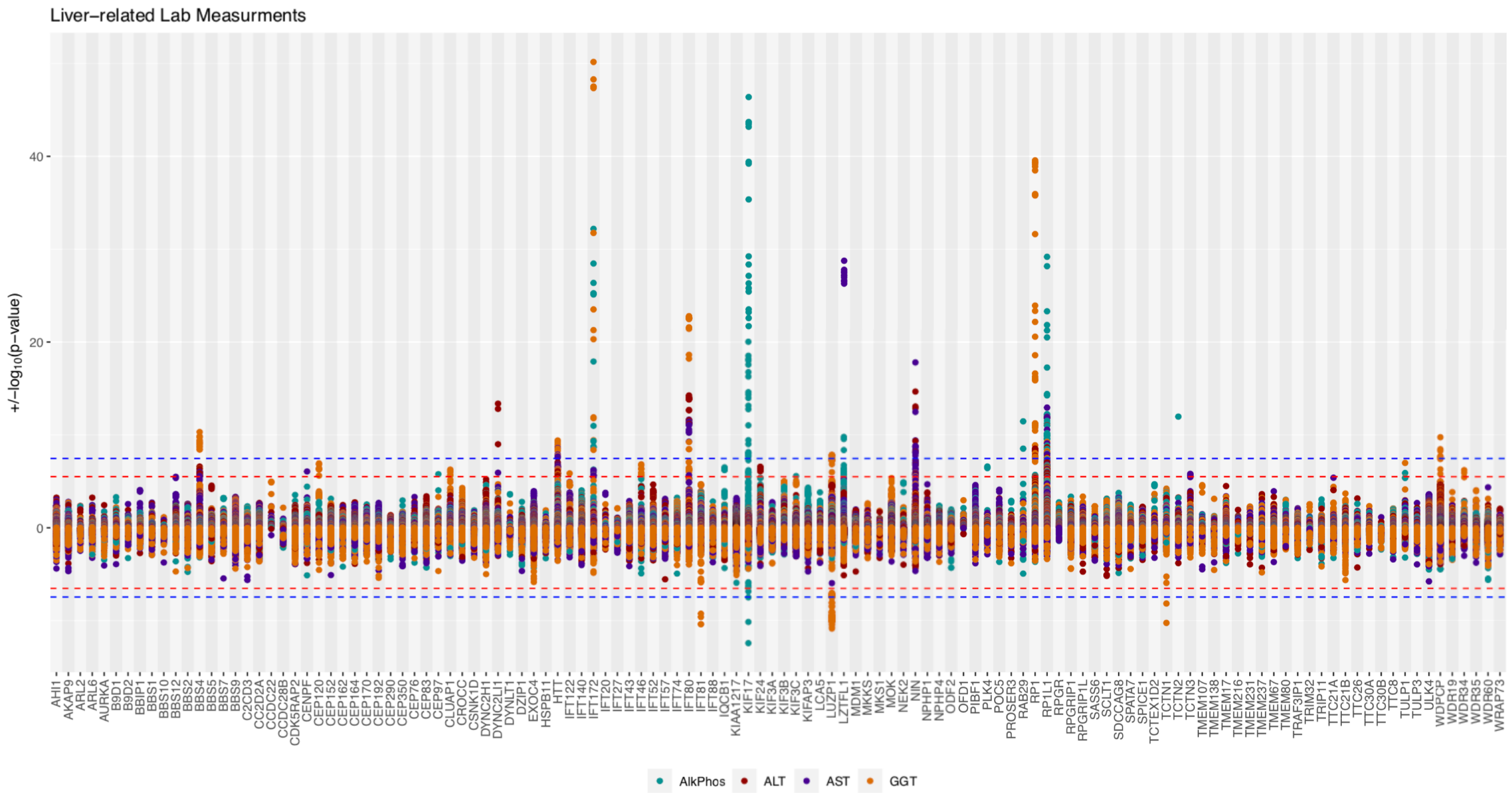


Figure 4. Association study and meta-analysis results for liver-related traits

Hudson plot illustrating the results of the discovery association analysis (top) and replication meta-analysis (bottom) of common variants within 122 ciliary gene with liver-related traits. Data for AlkPhos is in teal, ALT in red, AST in purple, and GGT in orange. The study-wide Bonferroni-adjusted significance threshold ($p < 3.4e-8$) is shown as a blue dashed line, while the experiment-wide Bonferroni-adjusted significance threshold ($p < 3.0e-6$ for the discovery analysis, $p < 3.3e-6$ for the meta-analysis) is shown as a red dashed line. Each analyzed gene is given equal space along the horizontal axis, with all tested genetic variants for a given gene plotted at the midline of the gene block.

Figure 5
Glucose-related Lab Measurements

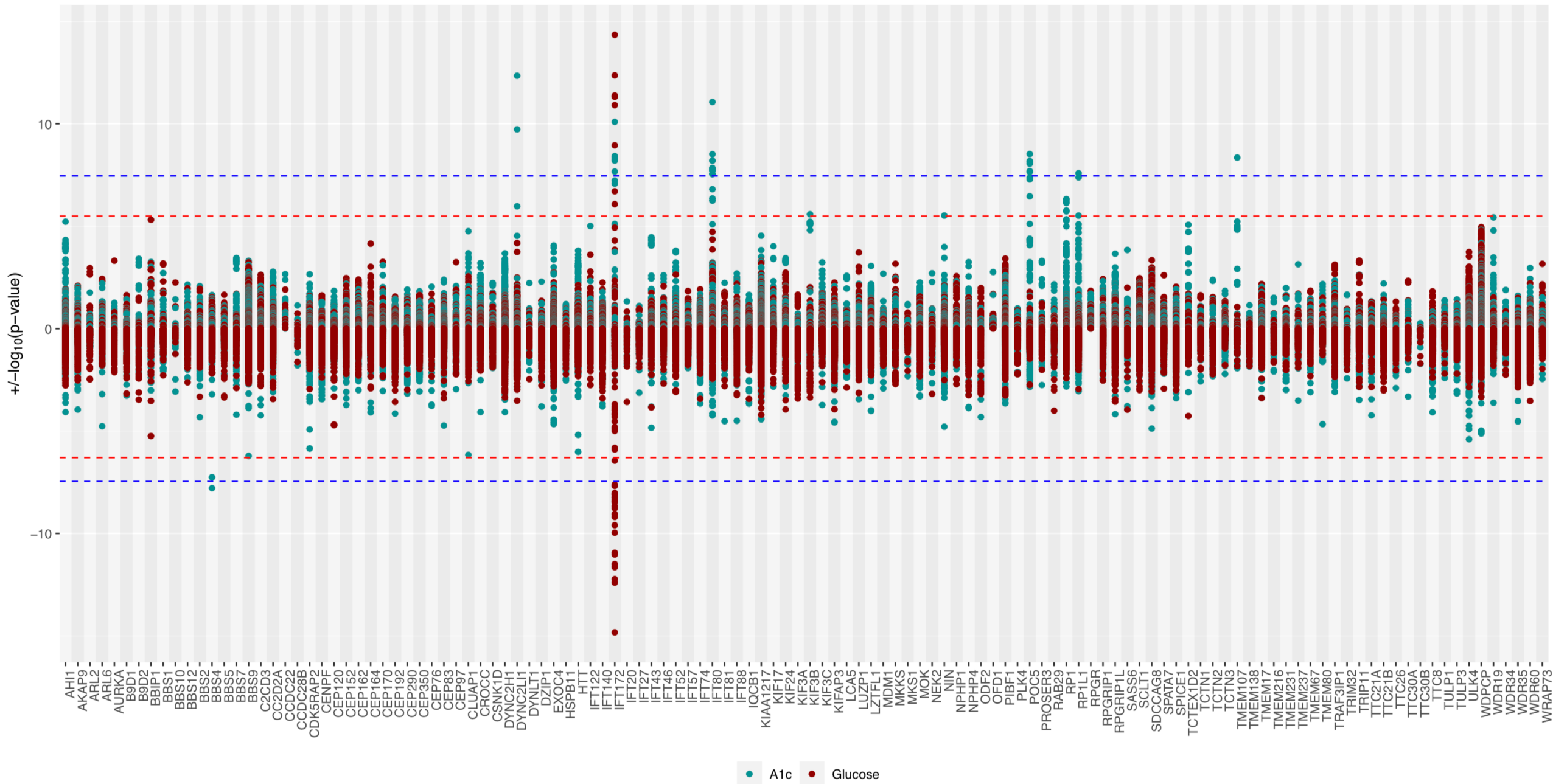


Figure 5. Association study and meta-analysis results for glucose-related traits

Hudson plot illustrating the results of the discovery association analysis (top) and replication meta-analysis (bottom) of common variants within 122 ciliary gene with glucose-related traits. Data for A1c is in teal, and data for glucose is in red. The study-wide Bonferroni-adjusted significance threshold ($p < 3.4 \times 10^{-8}$) is shown as a blue dashed line, while the experiment-wide Bonferroni-adjusted significance threshold ($p < 3.0 \times 10^{-6}$ for the discovery analysis, $p < 3.3 \times 10^{-6}$ for the meta-analysis) is shown as a red dashed line. Each analyzed gene is given equal space along the horizontal axis, with all tested genetic variants for a given gene plotted at the midline of the gene block.

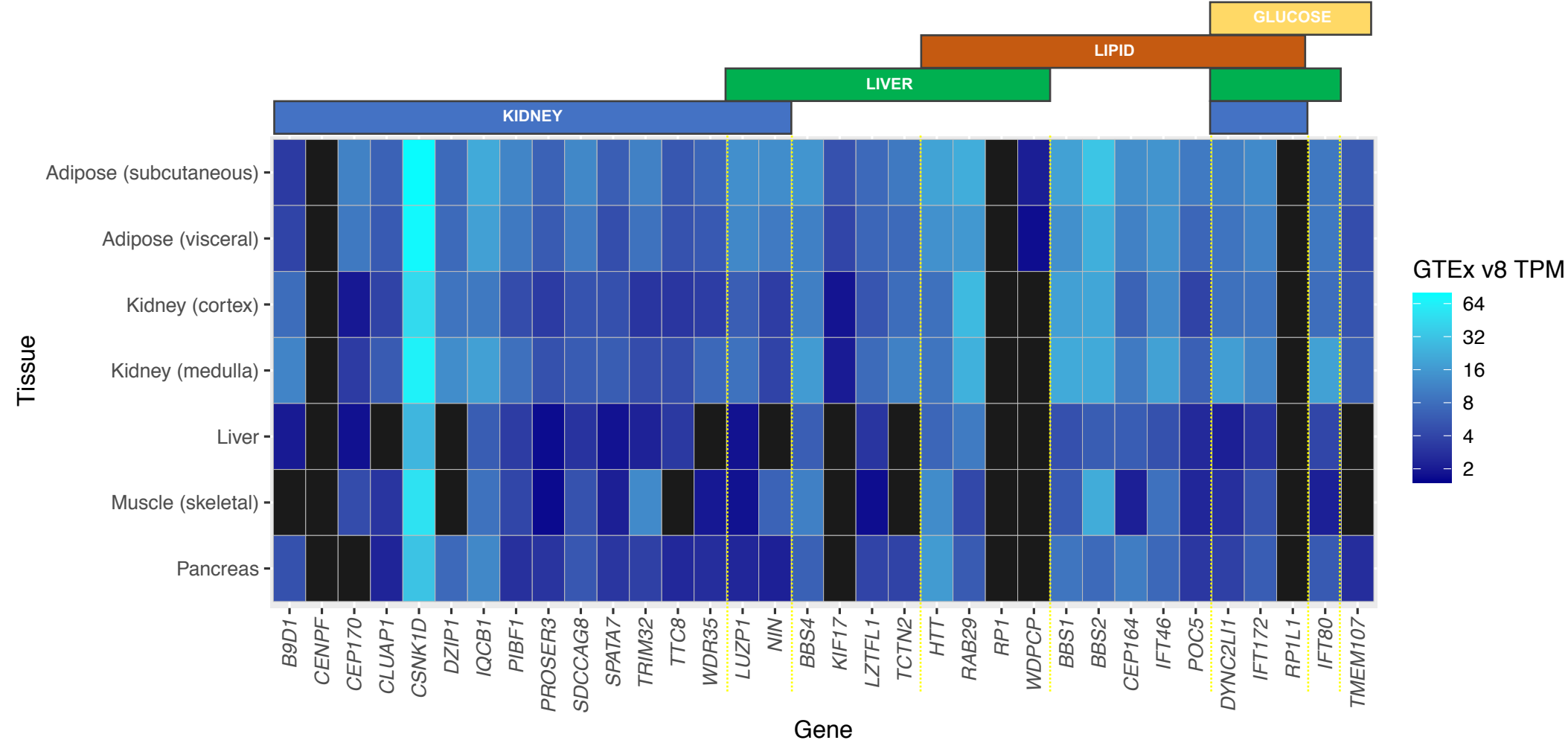


Figure 6. Results of tissue-specific expression analysis

Heatmap depicting the results of the tissue-specific expression analysis for each of the 34 genes with significant associations in our discovery analysis. The phenotype domain(s) significantly associated with each gene are indicated by the colored bars at the top of the graph.

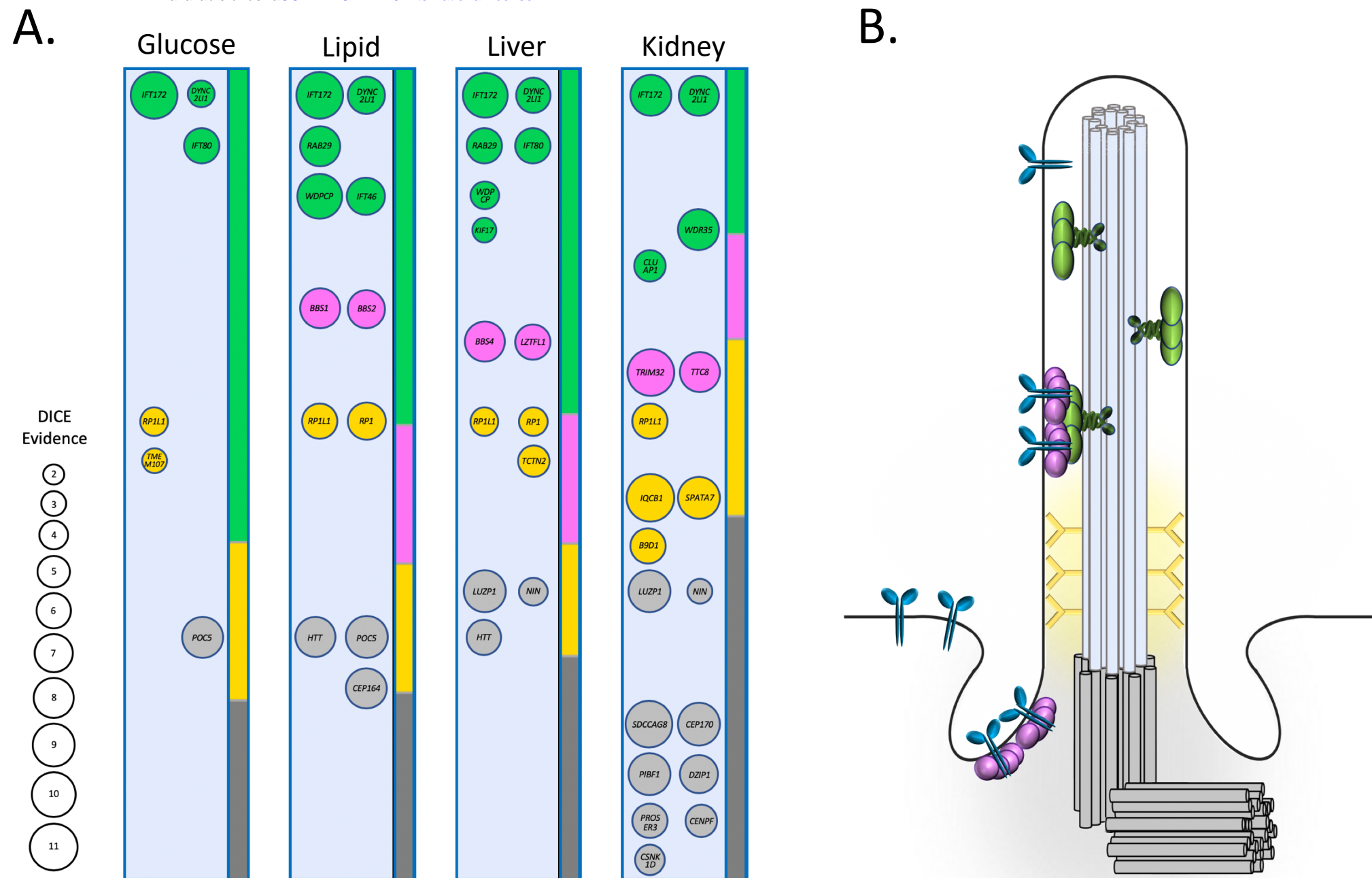


Figure 7. Schematics illustrating the results of DiCE/Pathway analysis, and a depiction of the primary cilium

- A. To integrate the multiple lines of evidence supporting each ciliary gene's association with a given phenotype domain, we employed a Diverse Convergent Evidence (DiCE) analysis approach estimating the strength of available corroborating data (detailed in Table S14). The DiCE scores for each gene, illustrating the strength of evidence supporting each ciliary gene's association with a given phenotype domain, are displayed here, with the magnitude of the score indicated by the size of each gene bubble. For each phenotype domain, gene bubbles are displayed at the same location for ease of comparison. Genes are colored by ciliary subcompartment – green for IFT-related genes, pink for the BBSome, orange for the transition zone, and grey for the basal body. A stacked bar graph is also displayed for each phenotype domain, illustrating the summed DiCE scores for all genes, by ciliary subcompartment, as a proportion of the total DiCE score per phenotype domain.
- B. A schematic of the primary cilium. The basal body and centrioles are displayed in grey, the ciliary transition zone in orange, the IFT machinery in green, and the BBSome in pink. The core of 9 microtubule doublets and the overlying membrane of the organelle are shown, and receptors being trafficked into/along the cilium are shown in blue.