TYPE 2 DIABETES KNOWLEDGE PORTAL NEWSLETTER

Visit us at ADA!

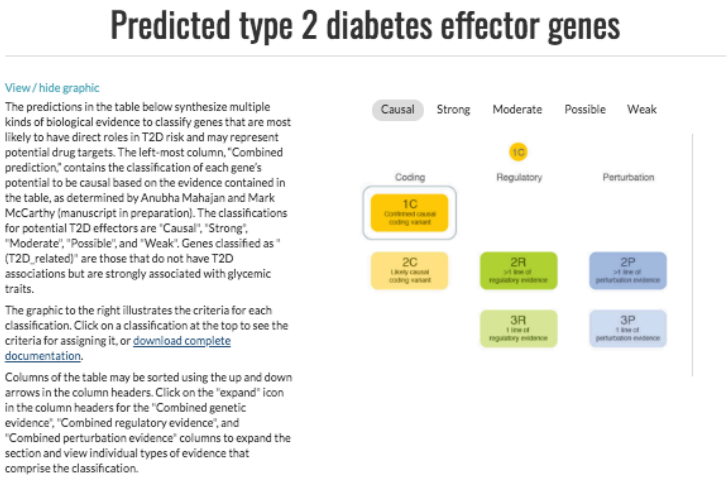
The T2DKP team will once again be presenting an exhibit booth at the 79th Scientific Sessions of the American Diabetes Association. This year we'll be joined at the booth by our colleagues from the Diabetes Epigenome Atlas (DGA). Stop by our booth (#2306) to get a personal, hands-on demonstration of the new tools and features, or just to say hello and let us know what new data and features you'd like to see in the T2DKP and DGA. We'll be there during all the exhibit hall hours:

Saturday 6/8 10am-4pm
Sunday 6/9 10am-4pm
Monday 6/10 10am-2pm

Upcoming webinar

Join us on Thursday July 18 at noon EDT for a webinar featuring gene-specific resources in the T2DKP. We'll cover the new predicted T2D effector gene list and our experimental gene prioritization toolkit, and delve into the Gene Page! Watch our home page for connection information.

Explore a curated list of predicted T2D effector genes and examine the evidence behind it



Combined prediction	Gene and locus	Combined genetic evidence	Combined regulatory evidence	Combined perturbation evidence
STRONG	ABCB9	2C	2R	3P
CAUSAL	ABCC8	1C		2P
WEAK	ABO		3R	
MODERATE	ADCY5		2R	3P

Despite worldwide efforts to understand the etiology of type 2 diabetes, its incidence is growing rapidly and effective treatments are few. The Accelerating Medicines Partnership in Type 2 Diabetes (AMP T2D) was established in 2014 to address this challenge, bringing together stakeholders from government, academia, and industry to speed up the translation of genetic data into insights about disease mechanisms and drug targets. The collaboration aimed to generate genetic association data for T2D and related traits, along with orthogonal data types to help identify and prioritize causal variants and genes, and to present this information to researchers via the T2D Knowledge Portal.

New datasets

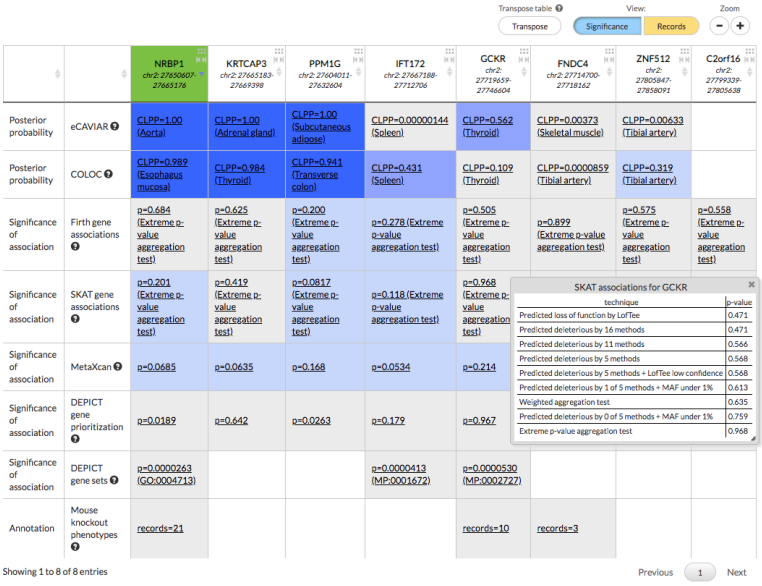
In the past few months, we've added a whole range of new datasets, including the largest disease-specific set of exome sequencing results to date:

- AMP T2D-GENES exome sequence analysis: variant-level and gene-level T2D associations from nearly 50,000 exomes (see Spotlight below)
- FinnMetSeq exome sequence analysis: associations for over 60 phenotypes from nearly 20,000 exomes
- At the AMP DCC, we've further analyzed the BioMe AMP T2D GWAS data to generate associations for three diabetic complications
- Hoorn DCS 2019: associations for 16 phenotypes from a cohort of nearly 2,000 type 2 diabetics (hosted at the T2DKP Federated node at EBI)
- Singapore Living Biobank GWAS; Singapore Chinese, Malay, and Indian Eye Studies GWAS: T2D and metabolic associations from East and South Asian populations
- BioBank Japan GWAS: associations for 40 phenotypes in over 190,000 East Asian individuals

Now, heading into the final year of this 5-year project, the generation, aggregation, and analysis of unprecedented amounts of data has allowed AMP T2D researchers to create a list of predicted T2D effector genes that may represent possible drug targets. Anubha Mahajan and Mark McCarthy have developed a heuristic for classifying genes according to the likelihood that they are causal for the development of T2D. The heuristic prioritizes and integrates multiple sources of evidence: genetic association data; functional genomics, for example eQTLs and chromatin conformation data; mutant phenotype evidence from model organisms and knockdown screens in human cells; and other evidence gathered from the literature.

We now present the results of their analysis in an interactive table that displays classifications of these genes according to the likelihood that they are causal for T2D, and allows you to view and explore all of the evidence underlying those classifications. While it should be remembered that these are only predictions, we hope that this list will be a valuable resource that can help suggest or support experimental directions for T2D researchers. Over the next year we plan to develop software to facilitate the generation of these results and to update them dynamically. We are pleased to provide to all researchers this distillation of many years of work from the global T2D research community.

Experimental Gene Prioritization Toolkit adds value to GWAS results



New datasets, continued

- Genetic Factors for Osteoporosis Consortium GWAS: bone mineral density and fracture associations for over 426,000 UK Biobank participants
- Liver function GWAS: associations for four liver enzymes in over 61,000 individuals
- MEDIA T2D GWAS: a meta-analysis of T2D associations in over 23,000 individuals of African American ancestry
- VATGen GWAS, with associations for fat distribution phenotypes in over 18,000 participants, has been updated with results from sex-stratified cohorts

All of these datasets are described in detail on the T2DKP Data page.

Stay up to date with the T2DKP

Contact us:
help@type2diabetesgenetics.org

Follow us on Twitter:
@T2DKP

Watch our videos:
Broad Institute Channel

Join our LinkedIn group:
T2D Knowledge Portal

Experimental Gene Prioritization Toolkit adds value to GWAS results, *continued from p.2*

One of the biggest challenges in interpreting genetic association results is to identify the genes that are responsible for the associations between variants and disease risk. Making connections between variants and effector genes is essential for a better understanding of disease genes and pathways, and for the development of new therapeutics.

To help researchers make these connections, we are assembling a toolkit of cutting-edge computational methods and applying them to all of the genetic association data in the T2DKP. The methods integrate GWAS data with transcriptomic data, tissue-specific gene expression results, eQTL data, and more to predict the probability of associations between variants, genes, phenotypes, and tissues. Methods currently in use include MetaXcan, DEPICT, eCAVIAR, and COLOC; we are actively working to include other methods.

The results of these analyses are presented, along with biological annotations such as gene-level T2D associations from the AMP T2D-GENES exome sequence analysis (see p. 4) and mouse knockout mutant phenotypes for the homologous gene, in an interactive table that is displayed on the "Genes in region" tab of T2DKP Gene pages. P-values or posterior probability scores are shown in an interactive table; each cell is color-coded by significance and contains full details that may be displayed in a pop-up window. For easy comparison of genes of interest, the table may be transposed and its columns re-ordered. Find more details about how the table works in this blog post.

Our hope is that this interface will help researchers to determine which genes in a region are most likely to be effector genes responsible for nearby variant associations, and to prioritize those genes for further investigation. It is under active development and we welcome your feedback.

Check out the Knowledge Portal Network (KPN)

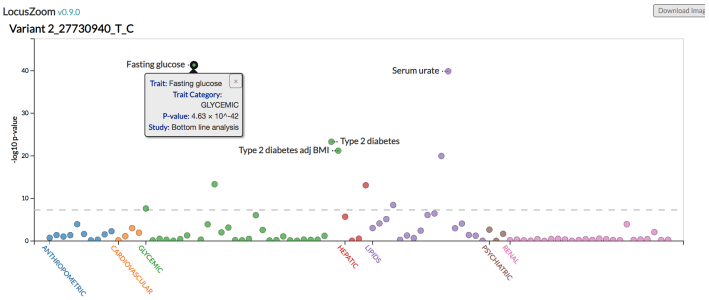
Did you know that the T2DKP has three sister Knowledge Portals that aggregate, analyze, and display genetic and genomic information for other cardiometabolic diseases and traits?

The Cardiovascular Disease Knowledge Portal (CVDKP; broadcvdi.org) includes genetic association results for coronary artery disease, atrial fibrillation, electrocardiogram traits, and lipid traits. Genetic association summary statistics and polygenic risk scores are available for download.

The Cerebrovascular Disease Knowledge Portal (CDKP; cerebrovascularportal.org) includes results for ischemic and hemorrhagic stroke and for MRI traits. Summary statistics from many datasets are also available for download.

The newest member of the KPN, the Sleep Disorder Knowledge Portal (SDKP; sleepdisordergenetics.org) includes genetic associations for multiple traits related to sleep and circadian rhythms and also provides download files.

What's the bottom line?



With so many datasets aggregated in the T2DKP, some variants are seen to be significantly associated with one phenotype in dozens of studies, at many different levels of significance. How can we integrate all these results to come up with a best estimate "bottom line" p-value for a phenotype? Meta-analysis can do this integration, but may get complicated when sample sets overlap between studies. We have implemented the METAL method, developed by our colleagues at the University of Michigan, to perform meta-analysis while taking into account sample overlap between datasets and generate a single bottom line p-value for each variant-phenotype association. Results are available for 150 phenotypes in the T2DKP and its three sister Portals of the Knowledge Portal Network. Bottom line p-values are currently displayed in the PheWAS module of the LocusZoom visualization tool, on T2DKP Variant pages.

Spotlight on the AMP T2D-GENES exome sequence analysis dataset

Exome sequences from 20,791 T2D cases and 24,440 controls are now available for viewing, browsing, and interactive analysis in the T2DKP. This project (Flannick et al., 2019), the result of the AMP T2D-GENES collaboration, brought together researchers across the globe who contributed exome sequences from individuals of multiple ancestries. As well as generating T2D association statistics for individual variants, the analysis produced gene-level T2D association scores that can help prioritize research into new drug targets. The individual-level data are available for secure analysis in the Custom burden test (accessible on T2DKP Gene pages) and the Genetic Association Interactive Tool (GAIT; available on T2DKP Variant pages). See our recent blog posts for details about this study and how the results are represented in the T2DKP.