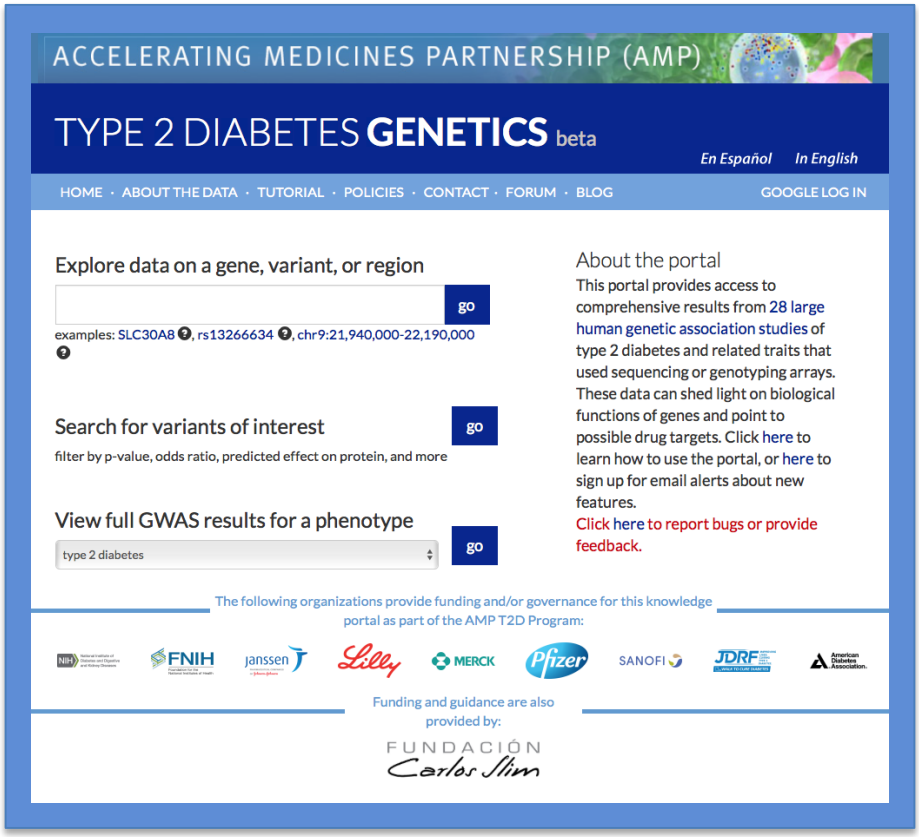
Introduction to the Type 2 Diabetes Knowledge Portal

The portal is an entry point for exploration of data from large studies that have discovered genetic associations between sequence variants and type 2 diabetes (T2D) or related traits. In late 2015 the portal contained data from 28 studies; data from more studies will be added in the future.

This tutorial provides a brief introduction to the major ways the portal can be used:

1) Explore data…
    - by gene
    - by sequence variant
2) Search for variants of interest
3) View genome-wide association results for a phenotype

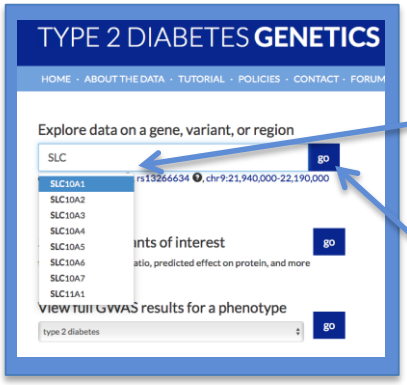
If you have questions or suggestions, please post in our forum or contact us at help@type2diabetesgenetics.org.



Home page: www.type2diabetesgenetics.org

Explore data by gene

Start with a particular gene of interest and retrieve the genetic associations of sequence variants in or near that gene.

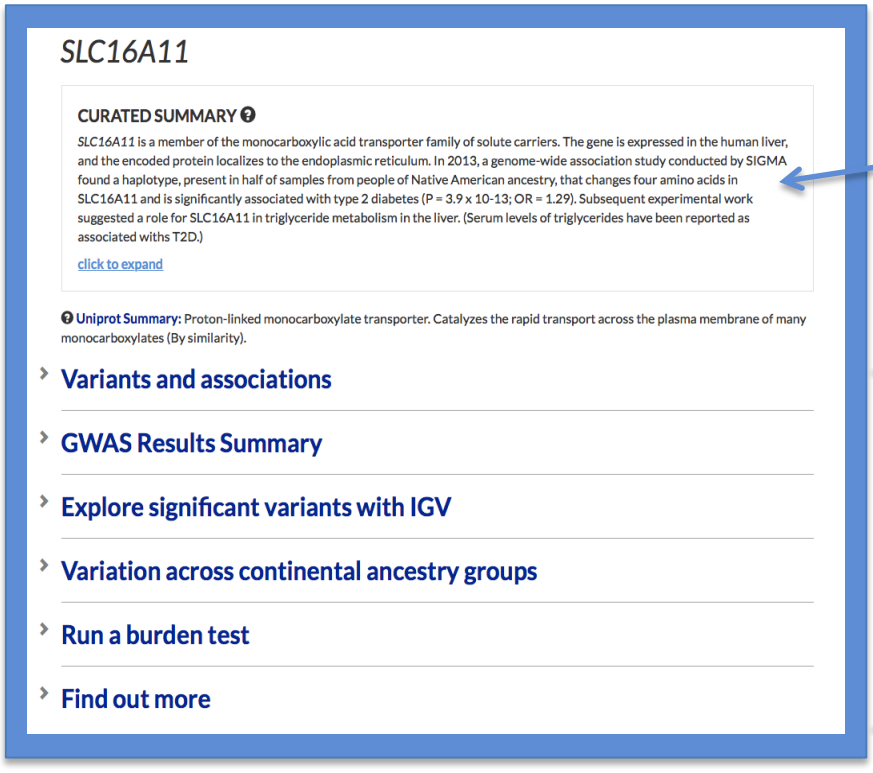


Type the gene name into the search box on the home page. Matching gene names are shown, and you can select one.

Clicking 'go' leads you to a page that displays all the data relevant to that gene.

Search by gene from home page

SLCA11 gene page



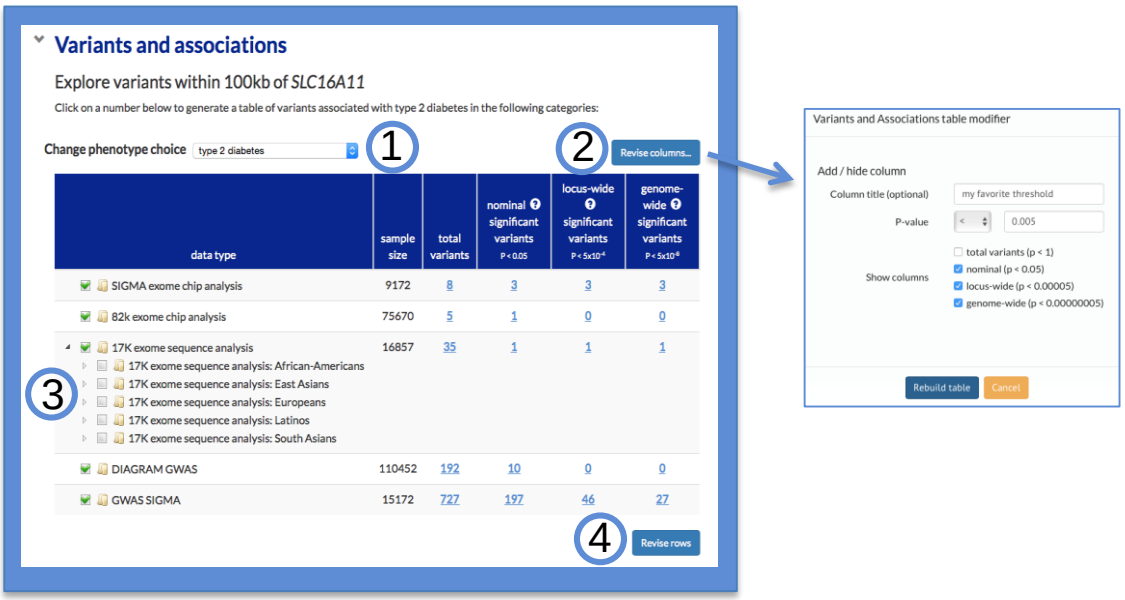
For major T2D-associated genes, a curated summary describes the biological role of the gene's product and the evidence for its association with T2D.

Different sections of the gene page can be expanded or collapsed by clicking on their headers.

The **Variants and associations** table on the gene page lets you search for sequence variants in the region of the gene that are significantly associated with any of 25 different phenotypes. Click on the number of variants in any category to see details for that set of variants.

1. Choose a phenotype to see the variants associated with it. The table displays the number of variants associated with that phenotype, at several confidence levels. Only the data sets that include data for the chosen phenotype are shown in the table.

2. Optional step: to add a new column showing the number of variants associated with the phenotype at a different confidence level, click the 'Revise columns…' button. This lets you specify a custom p-value for the variants to display in the new column. You can also remove columns by un-checking them. Click 'Rebuild table' to implement your selections.

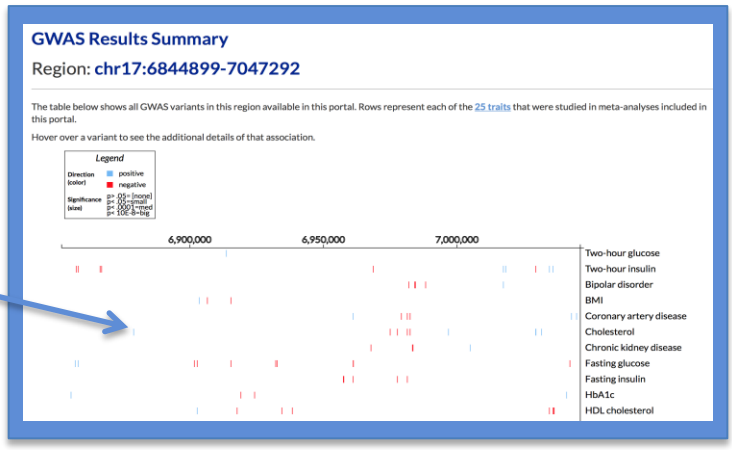


3. A small triangle to the left of a data set name indicates that it is comprised of sub-sets. Click on the triangle to expand or collapse the list. Check or un-check the boxes to select or remove data sets.

4. If you selected or removed data sets in step 3, use the 'Revise rows' button to implement your selections.

The **GWAS Results Summary** section of the gene page graphically displays all the genome-wide association study (GWAS) results across the region surrounding the gene, for variants associated with the 25 traits. You can zoom in or out by scrolling through the graphic.

Bars indicate sequence variant positions; color indicates direction; thickness indicates significance



Mouse over a bar to see the variant name and other details

The **Explore significant variants with IGV** section of the gene page shows significant variants in the region of the gene of interest, mapped onto the genome via the Integrative Genomics Viewer (IGV).

Add tracks to show variants associated with different phenotypes

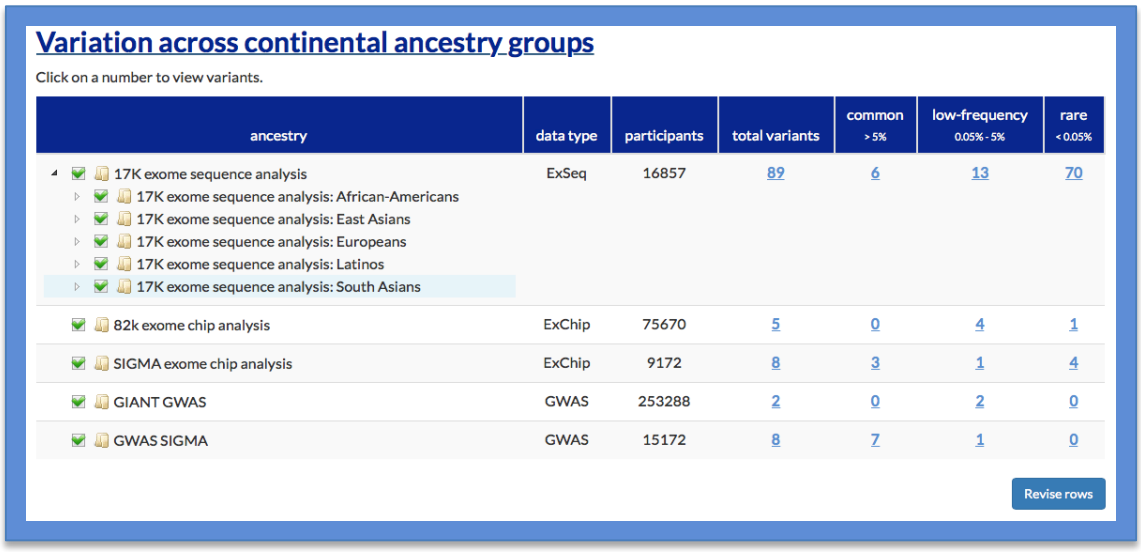
To view a different part of the genome, enter a gene or variant name, or genomic coordinates

Zoom in or out

Remove or customize tracks

Click on variant to view details

The **Variation across continental ancestry groups** table displays the number of variants found in at different frequencies in different ancestry groups, in or near your gene of interest.



Variation across continental ancestry groups

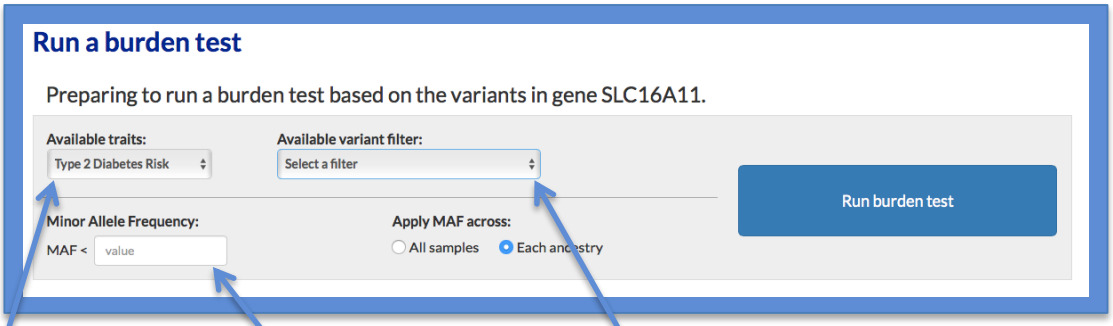
Click on a number to view variants.

ancestry	data type	participants	total variants	common > 5%	low-frequency 0.05% - 5%	rare < 0.05%
17K exome sequence analysis	ExSeq	16857	89	6	13	70
17K exome sequence analysis: African-Americans						
17K exome sequence analysis: East Asians						
17K exome sequence analysis: Europeans						
17K exome sequence analysis: Latinos						
17K exome sequence analysis: South Asians						
82k exome chip analysis	ExChip	75670	5	0	4	1
SIGMA exome chip analysis	ExChip	9172	8	3	1	4
GIANT GWAS	GWAS	253288	2	0	2	0
GWAS SIGMA	GWAS	15172	8	7	1	0

Revise rows

The functionality of this table is similar to that of the **Variants and associations** table. Select data sets to view by checking or un-checking them. Small triangles indicate that data sets contain subsets that can be expanded or collapsed. After selecting data sets to view or remove, click the 'Revise rows' button to update the table. The number of variants in each column is hyperlinked to a table containing more details about those variants.

The **Run a burden test** tool allows you to evaluate the significance of sets of variants within your gene of choice. Pull-down menus let you select particular traits and limit the search to particular kinds of variants, such as those predicted to be protein-truncating. You can also specify an upper limit for minor allele frequency, and choose whether to apply this limit across all samples or across each ancestry.



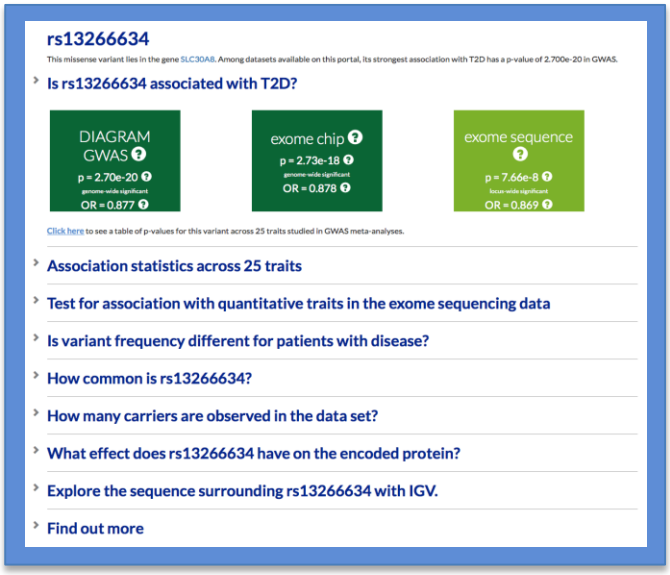
Select a diabetes-related trait

Set MAF

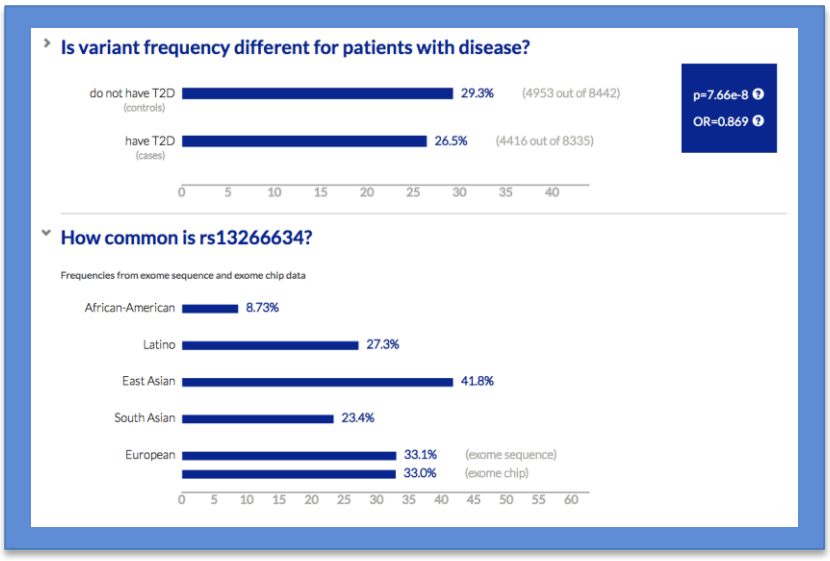
Limit test to variants with specific effects on the encoded protein

Explore data by sequence variant

Enter the identifier for a variant of interest into the home page search box, or click on a hyperlinked variant identifier anywhere it appears, to view a variant page.

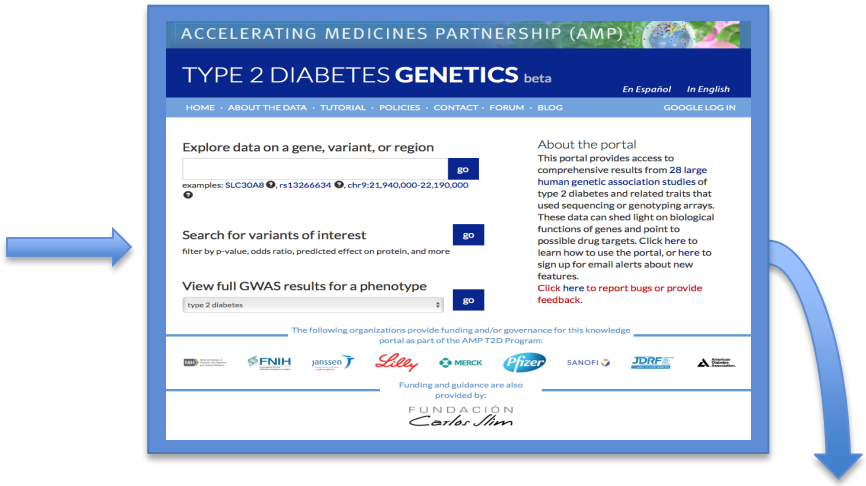


On the variant page, you can see at a glance whether a variant is found at a different frequency in people with T2D compared to people without it. You can also view its frequency in different ancestries. Other sections of the variant page show its association with various traits and its predicted effects on the encoded protein, if applicable.

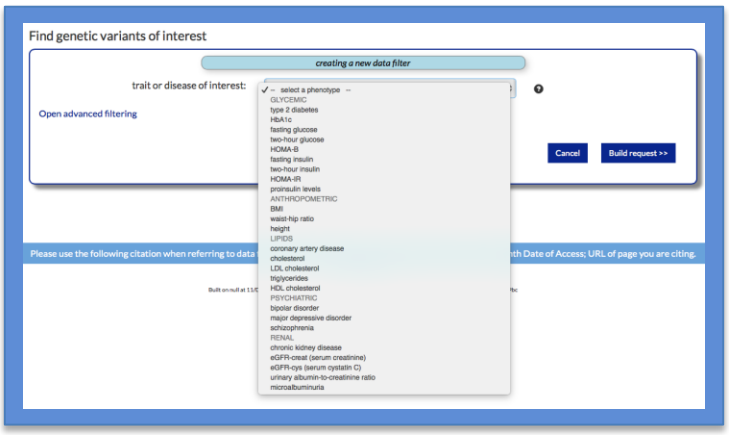


Advanced search for variants

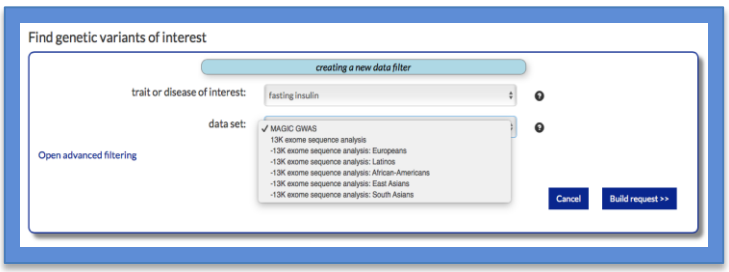
This versatile query builder allows you to specify multiple filters to retrieve a set of sequence variants of interest.

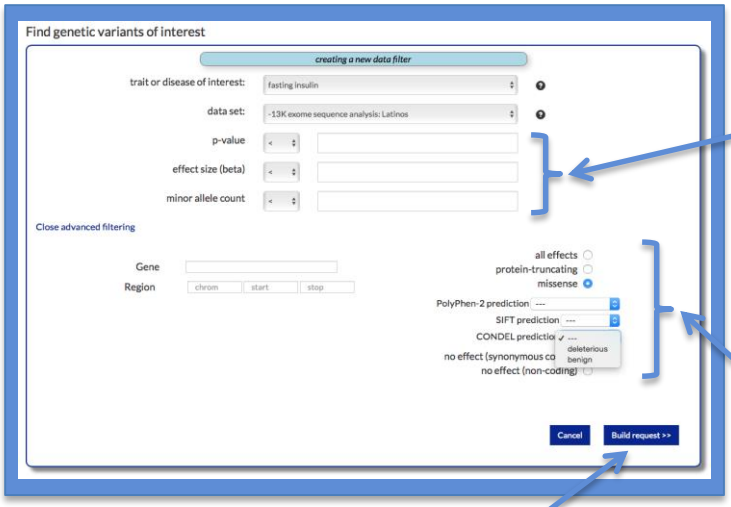


Start by selecting a phenotype from the pull-down menu.



You will then be offered a choice of data sets from those that contain data for the phenotype you selected. Click on the “data set” pull-down and choose a data set from the menu.



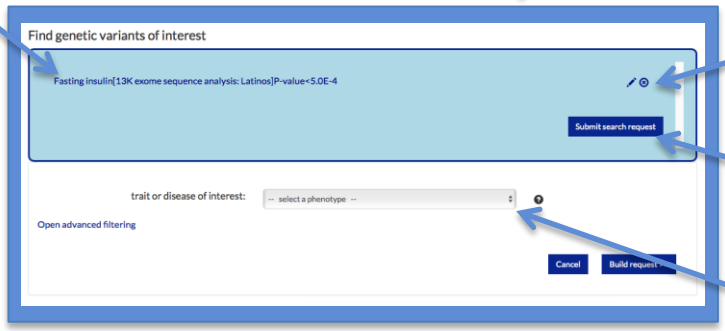


Once you choose a data set, options will appear to specify values relevant to the data set you selected. A p-value is required, and may be entered in the format “0.0005” or “5.0E-4”. Other values are optional.

The Advanced filtering options allow you to specify a gene or a chromosomal region and to choose the predicted effects of the variant on the encoded protein(s). If you select “missense”, additional options will appear allowing you to filter by the predictions of three different algorithms.

After specifying filters, click the ‘Build request’ button.

Details of the filters you specified



Edit or delete filters

Submit the search

Add filters to the search

Search results

Edit your query

‘+’ sign allows you to select more data types to display in the table

Showing variants that meet the following criteria:

- Fasting insulin[13K exome sequence analysis: Latinos]P-value<5.0E-4

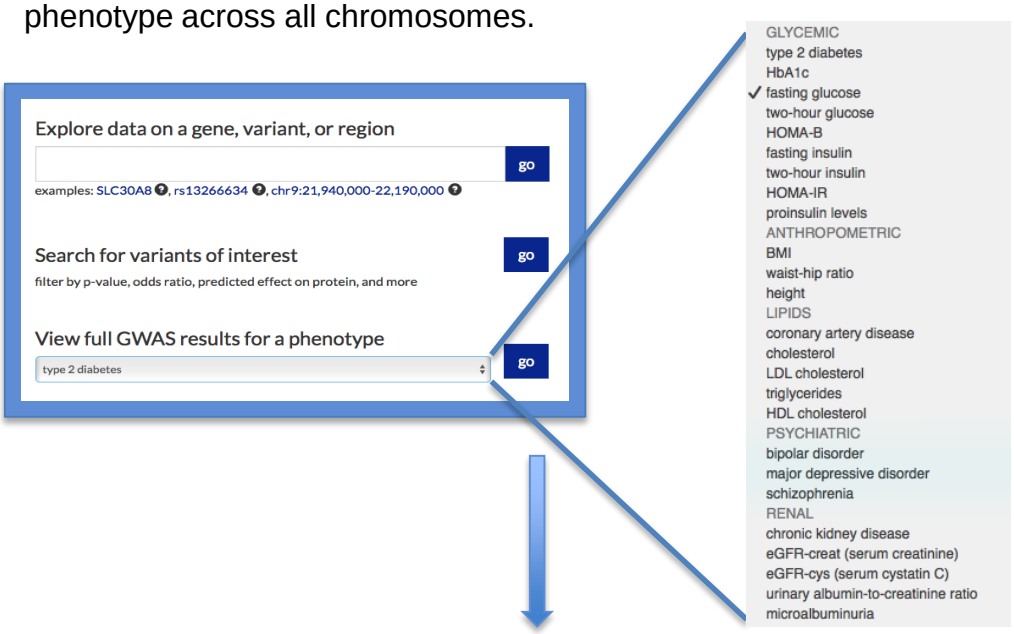
Click here to refine your results

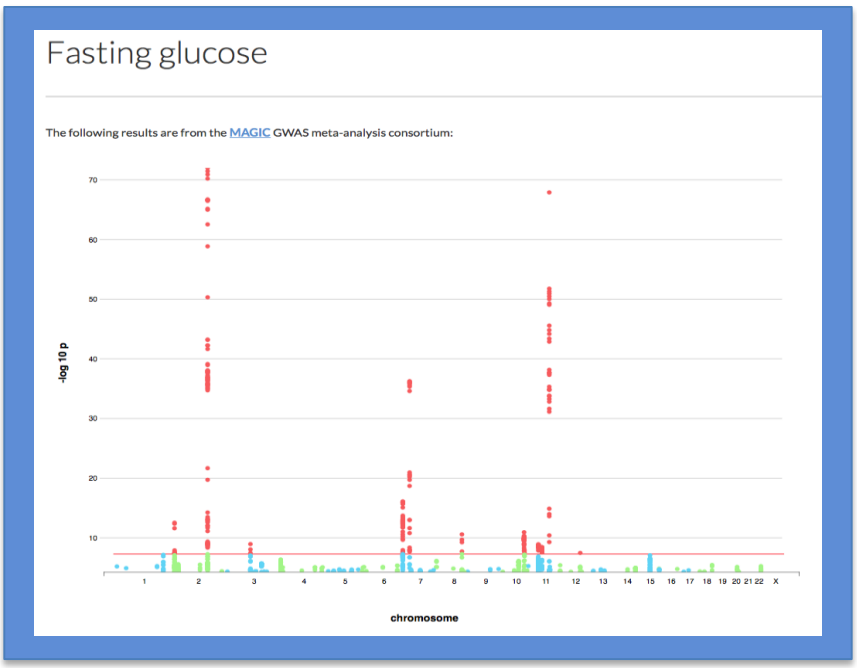
For variants that do not reach significance, odds ratios may be unreliable.

							fasting insulin +	
							+	13K exome sequence analysis: Latinos +
variant ID	nearest gene	dbSNP ID	protein change	consequence	chromosome	position	p-value	effect size (beta)
14:105410471	AHNAK2	rs146965608	p.V3773I	missense variant	14	105410471	0.00000453	-1.43
14:105409903	AHNAK2	rs115776887	p.T3962M	missense variant	14	105409903	0.00000453	-1.43
6:26392917	BTN2A2	rs114760306	p.S432T	missense variant	6	26392917	0.00000558	-1.16
11:93141472	CCDC67	rs76382603	p.R468*	stop gained	11	93141472	0.0000149	1.75
6:43414234	ABCC10	rs1214747		intron variant	6	43414234	0.0000192	0.171
15:86697675	AGBL1	rs150261781	p.R47W	missense variant	15	86697675	0.0000203	-0.339
2:217279768	SMARCAL1	rs11555797	p.R114H	missense variant	2	217279768	0.0000225	-0.324
1:203186093	CHIT1	rs1065761	p.A442V	missense variant	1	203186093	0.0000274	-0.201
2:36970212	VIT	rs146426374		intron variant	2	36970212	0.0000329	0.566

View Full GWAS results for a phenotype

This tool allows you to select one of 25 phenotypes from the list and view a graphical display (a “Manhattan plot”) of variants associated with the phenotype across all chromosomes.





Scroll across the graphic to zoom in or out, Mouse over an individual variant see more details, or click to view a page of information about it. Variants are listed in a table below the plot.