

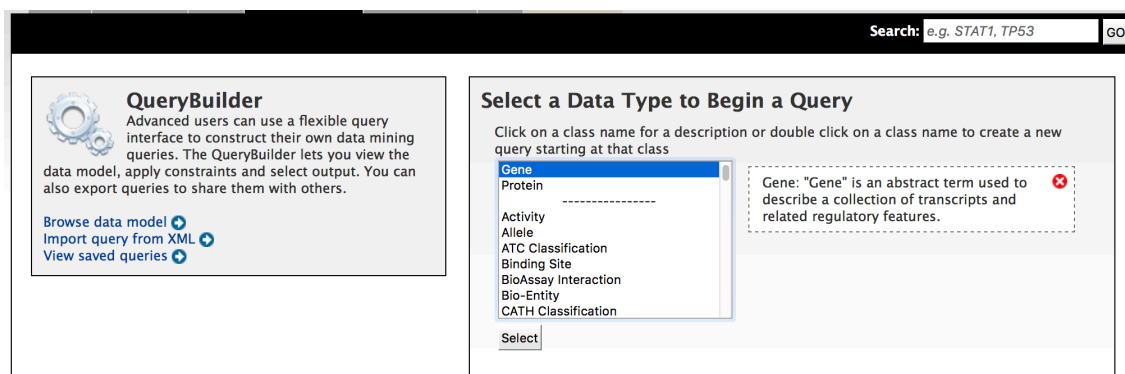
Supplementary File S2: Extraction of gene-disease associations from literature using MeSH descriptors

Note: The users can click the “+” symbol to expand the class/node of interest and view the available nodes/classes within.

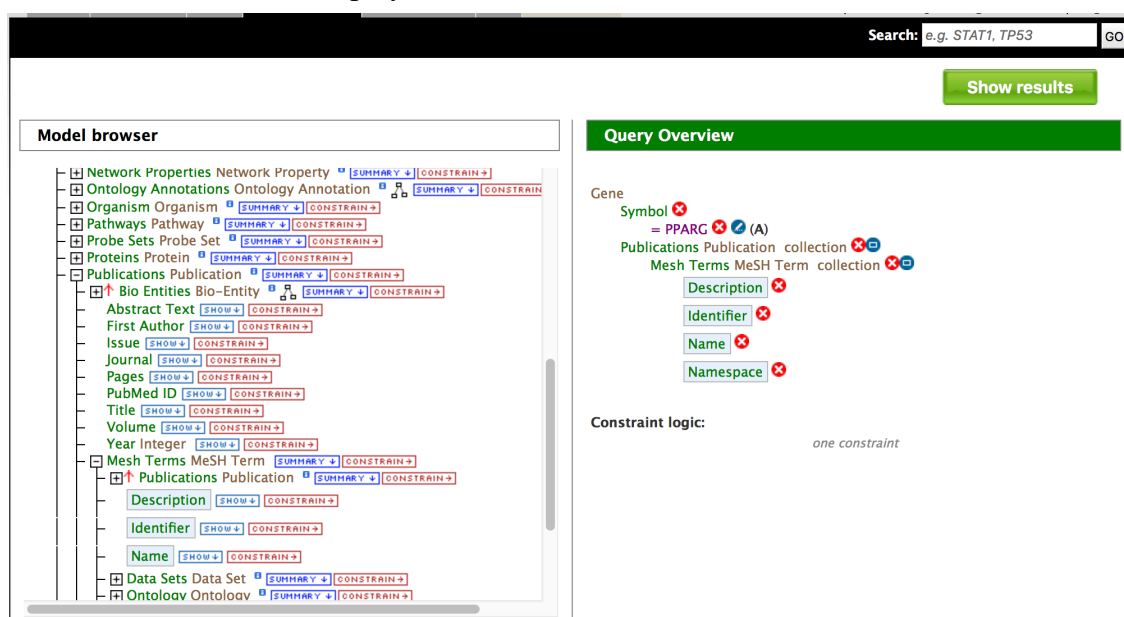
Step 1: Click the “QueryBuilder” tab in the tab panel



Step 2: Select “Gene” from the drop-down menu of data types and click the “Select” button.



Step 3-A: In model browser, select “Publication>MeSH Terms” and click the “SUMMARY” label to display the default MeSH term attributes.



Step 3-B: To restrict the search to a specific gene (e.g. PPARG; Organism- H. sapiens), click the “CONSTRAIN” label next to “Symbol” tab and enter the gene symbol and click “Add to query”.

The screenshot shows the 'Model browser' on the left and the 'Query Overview' on the right. In the 'Model browser', the 'Symbol' field is highlighted with a red 'CONSTRAIN' button. In the 'Query Overview', the 'Symbol' field is listed with a red 'x' icon, and the value 'PPARG' is entered. The 'Constraint logic' section shows 'one constraint'.

Constrain for gene (Gene Symbol; PPARG):

The screenshot shows the 'Model browser' on the left and the 'Query Overview' on the right. A 'Constrain for Gene > Symbol' dialog box is open in the center. The dialog box has a red header 'Constrain' and a section 'Choose a filter'. Under 'Filter query results on this field having a specific value', the 'Gene > Symbol' field is selected, and the value 'PPARG' is entered in the input field. The 'Add to query' button is visible. Under 'Filter query results on this field having any value or not', the 'Has no value' radio button is selected, and the 'Add to query' button is also visible.

Constrain for organism (by Short Name- H. sapiens)

The screenshot shows a web interface with a search bar at the top right containing "e.g. STAT1, TP53" and a "GO" button. Below the search bar is a "Show results" button. The main area is divided into a "Model browser" on the left and a "Query Overview" on the right. The "Model browser" lists various biological entities with expandable/collapsible icons and "SUMMARY" and "CONSTRAIN" buttons. A "Constrain" dialog box is open in the center, titled "Constrain for Gene > Organism > Short Name". It has a red header and contains the following elements:

- Choose a filter**
- Section: "Filter query results on this field having a specific value"
- Field: "Gene > Organism > Short Name"
- Operator: "="
- Value: "H. sapiens" (selected in a dropdown)
- Button: "Add to query"
- Section: "Filter query results on this field having any value or not"
- Options: "Has no value" (selected) and "Has a value"
- Button: "Add to query"

The "Query Overview" on the right shows a "Constraint logic" section with the text "A and B and C and C and C and C and C and C" and a button to edit the logic.

Step 4: To restrict the output to gene-disease associations, select “Publication->MeSH Terms->Tree Numbers->Category” and click the “CONSTRAIN” label next to “Name”. Select “Diseases” from the list of items in the drop-down menu and click “Add to query”.

The screenshot shows the same web interface as above, but with the "Model browser" expanded to show a hierarchical path: "Publication Types > MeSH Term > Tree Numbers > Category > Name". The "Constrain" dialog box is open, titled "Constrain for Gene > Publications > Mesh Terms > Tree Numbers > Category > Name". It contains the following elements:

- Choose a filter**
- Section: "Filter query results on this field having a specific value"
- Field: "Gene > Publications > Mesh Terms > Tree Numbers > Category > Name"
- Operator: "="
- Value: "Diseases" (selected in a dropdown)
- Button: "Add to query"
- Section: "Filter query results on this field having any value or not"
- Options: "Has no value" (selected) and "Has a value"
- Button: "Add to query"

The "Query Overview" on the right shows the "Constraint logic" section with the text "A and B and C and C and C and C and C and C" and a button to edit the logic.

Step 5: Click the “Show results” button on upper right of the panel to view the output.

Model browser

Browse through the classes and attributes. Click on [SUMMARY](#) links to add summary of fields to the results table or on [SHOW](#) links to add individual fields to the results. Use [CONSTRAIN](#) links to constrain a value in the query.

- Gene [SUMMARY](#) [CONSTRAIN](#)
 - Brief Description [SHOW](#) [CONSTRAIN](#)
 - Description [SHOW](#) [CONSTRAIN](#)
 - Name [SHOW](#) [CONSTRAIN](#)
 - NCBI Gene ID [SHOW](#) [CONSTRAIN](#)
 - DB identifier [SHOW](#) [CONSTRAIN](#)
 - Secondary Identifier [SHOW](#) [CONSTRAIN](#)
 - Symbol [SHOW](#) [CONSTRAIN](#)
 - Type [SHOW](#) [CONSTRAIN](#)
 - Alleles Allele [SUMMARY](#) [CONSTRAIN](#)
 - Chromosome Chromosome [SUMMARY](#) [CONSTRAIN](#)
 - Chromosome Location Location [SUMMARY](#) [CONSTRAIN](#)
 - Cross References Cross Reference [SUMMARY](#) [CONSTRAIN](#)
 - Data Sets Data Set [SUMMARY](#) [CONSTRAIN](#)
 - Diseases Disease [SUMMARY](#) [CONSTRAIN](#)
 - Gene Disease Pairs Gene Disease Pair [SUMMARY](#) [CONSTRAIN](#)
 - GO Annotation GO Annotation [SUMMARY](#) [CONSTRAIN](#)
 - Integrated Pathway Clusters Integrated Pathway Cluster [SUMMARY](#) [CONSTRAIN](#)
 - Interactions Interaction [SUMMARY](#) [CONSTRAIN](#)
 - Kegg Orthology Kegg Orthology [SUMMARY](#) [CONSTRAIN](#)
 - Locations Location [SUMMARY](#) [CONSTRAIN](#)

Show empty fields

Query Overview

Gene

- Symbol [✕](#)
- = PPARG [✕](#) [🔍](#) (A)
- Organism Organism [✕](#) [🔍](#)
- Short Name [✕](#)
- = H. sapiens [✕](#) [🔍](#) (B)
- Publications Publication collection [✕](#) [🔍](#)
- Mesh Terms MeSH Term collection [✕](#) [🔍](#)
- Description [✕](#)
- Identifier [✕](#)
- Name [✕](#)
- Namespace [✕](#)
- Tree Numbers MeSH Tree collection [✕](#) [🔍](#)
- Category Mesh Category [✕](#) [🔍](#)
- Name [✕](#)
- = Diseases [✕](#) [🔍](#) (C)

Constraint logic: A and B and C and C and C and C and C and C

A and B and C and C and C and C and C and C [🔄](#)

Results table:

Trail: [Query](#) > Results

Manage Columns
 Manage Filters
 Manage Relationships

Showing 26 to 50 of 397 rows Rows per page: 25 [⏪](#) [⏩](#) [⏴](#) [⏵](#) page 2

Gene Symbol	Mesh Terms Identifier	Mesh Terms Name	Mesh Terms Namespace	Mesh Terms Description
PPARG	D017544	Aortic Aneurysm, Abdominal	NO VALUE	An abnormal balloon- or sac-like dilatation in the wall of the ABDOMINAL AORTA which gives rise to the visceral, the parietal, and the terminal (iliac) branches below the aortic hiatus at the diaphragm...
PPARG	D001024	Aortic Valve Stenosis	NO VALUE	A pathological constriction that can occur above (supraaortic stenosis), below (subaortic stenosis), or at the AORTIC VALVE. It is characterized by restricted outflow from the LEFT VENTRICLE into ...
PPARG	D019571	Arrhythmogenic Right Ventricular Dysplasia	NO VALUE	A congenital cardiomyopathy that is characterized by infiltration of adipose and fibrous tissue into the RIGHT VENTRICLE wall and loss of myocardial cells. Primary injuries usually are at the free wall...
PPARG	D001161	Arteriosclerosis	NO VALUE	Thickening and loss of elasticity of the walls of ARTERIES of all sizes. There are many forms classified by the types of lesions and arteries involved, such as ATHEROSCLEROSIS with fatty lesions in th...
PPARG	D015535	Arthritis, Psoriatic	NO VALUE	A type of inflammatory arthritis associated with PSORIASIS, often involving the axial joints and the peripheral terminal interphalangeal joints. It is characterized by the presence of HLA-B27-associat...
PPARG	D001172	Arthritis, Rheumatoid	NO VALUE	A chronic systemic disease, primarily of the joints, marked by inflammatory changes in the synovial membranes and articular structures, widespread fibrinoid degeneration of the collagen fibers in mese...
PPARG	D001249	Asthma	NO VALUE	A form of bronchial disorder with three distinct components: airway hyper-responsiveness (RESPIRATORY HYPERSENSITIVITY), airway INFLAMMATION, and intermittent AIRWAY OBSTRUCTION. It is characterized b...

Step 6: The users can export the results by “Export” button on the top-right.