

# Making Sense of Gene Lists using MGI

Short Course Bioinformatics  
Workshops  
2014



- Gene Lists
  - Differentially Expressed Genes (DEGS) from RNA Seq Analysis
  - Genes in regions associated with complex trait mapping
  - Genes identified from exome analysis
- Gene List Analyses
  - What are the functions, phenotypes, and disease associations for all of the genes in a list?
  - Is there a statistically significant functional or phenotypic “theme” in a gene list?
  - Are there mouse models available?

## Bioinformatics Tools for Gene List Annotation and Analysis in Mouse

- Batch Query Tool
  - <http://www.informatics.jax.org/batch>
- Mouse Mine
  - [www.mousemine.org](http://www.mousemine.org)

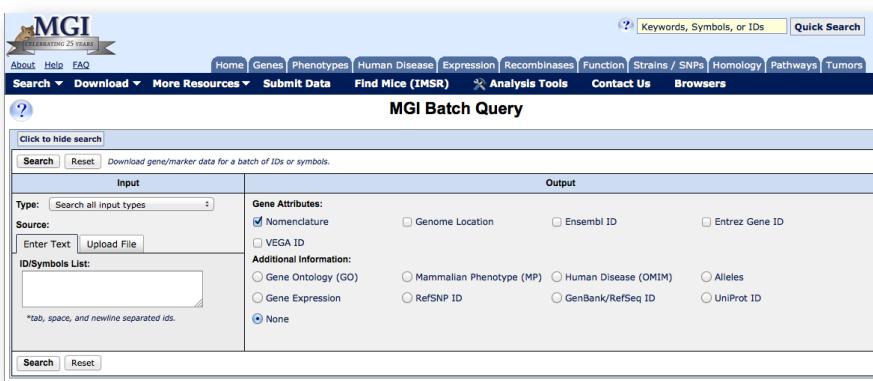


## Batch Query in a Nutshell



- Batch Query – Annotate a list of genes
  - Upload a list of symbols or other acceptable identifiers
  - Generate the annotations you are interested in
  - Download the results
- Batch Query does not provide gene list analysis tools such as gene set enrichment
- *Example gene lists are available on DropBox site*
  - [https://www.dropbox.com/sh/seqishl631f363x/AADoBskzYjPt\\_iJI21dxJqV8a](https://www.dropbox.com/sh/seqishl631f363x/AADoBskzYjPt_iJI21dxJqV8a)

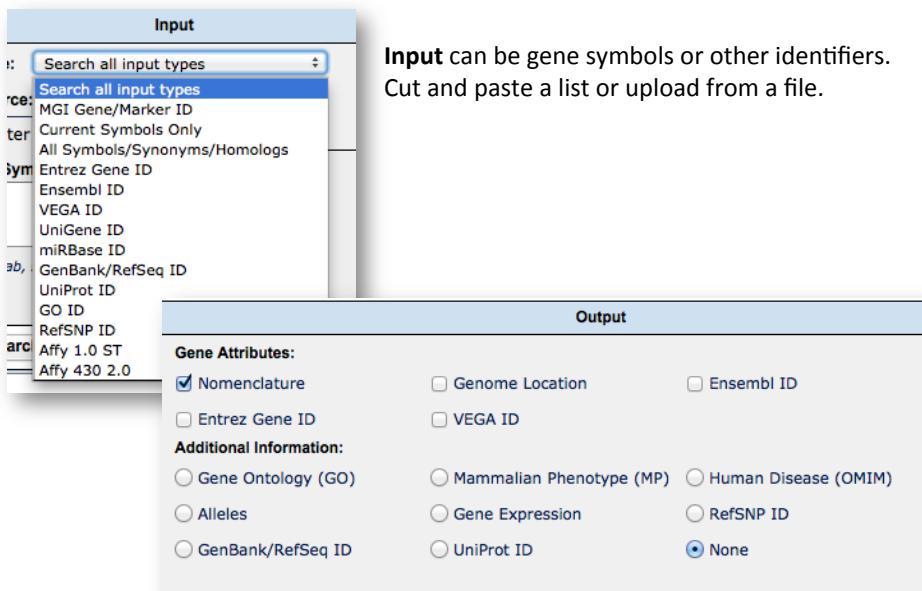
## Batch Query



The screenshot shows the MGI Batch Query interface. The top navigation bar includes links for About, Help, FAQ, Home, Genes, Phenotypes, Human Disease, Expression, Recombinases, Function, Strains / SNPs, Homology, Pathways, and Tumors. Below the navigation is a search bar with a 'Quick Search' button. The main title 'MGI Batch Query' is centered above a form. The form has two main sections: 'Input' and 'Output'. The 'Input' section contains fields for 'Type' (set to 'Search all input types'), 'Source' (with options 'Enter Text' or 'Upload File'), and a 'ID/Symbols List' text area. The 'Output' section contains checkboxes for 'Gene Attributes' (Nomenclature is checked) and 'Additional Information' (including Genome Location, Ensembl ID, Entrez Gene ID, VEGA ID, etc.). Below the form are 'Search' and 'Reset' buttons. A note at the bottom left says 'Click to hide search'.

<http://www.informatics.jax.org/batch>

*Gene lists available from DropBox site*



This screenshot shows the same MGI Batch Query interface as above, but with specific sections highlighted. The 'Input' section on the left is highlighted, showing the dropdown menu for 'Type' set to 'Search all input types' and a list of other identifiers like MGI Gene/Marker ID, Current Symbols Only, All Symbols/Synonyms/Homologs, Entrez Gene ID, Ensembl ID, VEGA ID, UniGene ID, miRBase ID, GenBank/RefSeq ID, UniProt ID, GO ID, RefSNP ID, Affy 1.0 ST, and Affy 430 2.0. The 'Output' section on the right is also highlighted, showing the 'Gene Attributes' and 'Additional Information' checkboxes. A callout box to the right of the input section states: 'Input can be gene symbols or other identifiers. Cut and paste a list or upload from a file.' A callout box below the output section states: 'Output can be genome location, gene function, disease associations, etc.'

4. Click

2. Select input type

3. Select output options

1. Paste gene list

1. Download options

Gene Ontology annotations for input list of genes.

Input	Input Type	MGI Gene/Marker ID	Nomenclature		Feature Type	Gene Ontology (GO)	
			Symbol	Name		ID	Term
Crp	current symbol	MGI:88512	Crp	C-reactive protein, pentraxin-related	protein coding gene	GO:0006953 GO:0005509 GO:0071277 GO:0015485 GO:0006958 GO:0001849 GO:0005576 GO:0005615 GO:0070062 GO:0030175 GO:0030424 GO:0030169 GO:0046872 GO:0010888 GO:0010745 GO:1900007 GO:0042803 GO:0051258 GO:0042060	acute-phase response calcium ion binding cellular response to calcium ion cholesterol binding complement activation, classical pathway complement component C1q binding extracellular region extracellular space filopodium growth cone low-density lipoprotein particle binding metal ion binding negative regulation of lipid storage negative regulation of macrophage derived foam cell differentiation positive regulation of dendrite development protein homodimerization activity protein polymerization wound healing
Adss	current symbol	MGI:87948	Adss	adenylosuccinate synthetase, non-muscle	protein coding gene	GO:0004019 GO:0004019 GO:0006167	adenylosuccinate synthase activity adenylosuccinate synthase activity AMP biosynthetic process

**Modify search to select a new output type**

1. Click

You searched for:  
Number of IDs/symbols entered: 347  
Input Type: Current Symbols Only  
Output options: Nomenclature, GO  
347 matching genes/markers found

Export: Text File Excel File

Input	Input Type	MGI Gene/Marker ID	Nomenclature	Gene Ontology (GO)			
			Symbol	Name	Feature Type	ID	Term
Crp	current symbol	MGI:88512	Crp	C-reactive protein, pentraxin-related	protein coding gene	GO:0006953	acute-phase response
						GO:0005509	calcium ion binding
						GO:0071277	cellular response to calcium ion
						GO:0015485	cholesterol binding
						GO:0006958	complement activation, classical pathway
						GO:0001849	complement component C1q binding
						GO:0005576	extracellular region
						GO:0005615	extracellular space
						GO:0070062	extracellular vesicular exosome
						GO:0030175	filopodium
						GO:0030426	growth cone
						GO:0030169	low-density lipoprotein particle binding
						GO:0046872	metal ion binding
						GO:0010888	negative regulation of lipid storage
						GO:0010745	negative regulation of macrophage derived foam cell differentiation
						GO:1900006	positive regulation of dendrite development
						GO:0042803	protein homodimerization activity
						GO:0051258	protein polymerization
						GO:0042060	wound healing
Adss	current symbol	MGI:87948	Adss	adenylosuccinate synthetase, non muscle	protein coding gene	GO:0004019	adenylosuccinate synthase activity
						GO:0004019	adenylosuccinate synthase activity
						GO:0006167	AMP biosynthetic process

**Batch Summary**

1. Click

Input remains the same.  
Select new output option.

Click to hide search

Search Reset Download gene/marker data for a batch of IDs or symbols.

Input		Output							
Type:	Current Symbols Only	<input checked="" type="checkbox"/> Nomenclature <input type="checkbox"/> Ensembl ID <input type="checkbox"/> VEGA ID <input type="checkbox"/> Mammalian Phenotype (MP) <input type="checkbox"/> Gene Expression <input type="checkbox"/> UniProt ID <input checked="" type="radio"/> Human Disease (OMIM) <input type="checkbox"/> RefSNP ID <input type="checkbox"/> None							
Source:		Enter Text	Upload File						
ID/Symbols List:		Crp Adss Dcaf8 *tab, space, comma							
Search Reset									

**Results**

You searched for:  
Number of IDs/symbols entered: 347  
Input Type: Current Symbols Only  
Output options: Nomenclature, GO  
347 matching genes/markers found.

Export: Text File Excel File

Input	Input Type	MGI Gene/Marker ID	Nomenclature	Gene Ontology (GO)			
			Symbol	Name	Feature Type	ID	Term
Crp	current symbol	MGI:88512	Crp	C-reactive protein	protein coding gene	GO:0006953	acute-phase response

**Batch Summary**

Click to modify search

**Results**

You searched for:

Number of IDs/symbols entered: 347  
Input Type: Current Symbols Only  
Output options: Nomenclature, OMIM  
347 matching genes/markers found.

Export:  Text File  Excel File

Input	Input Type	MGI Gene/Marker ID	Nomenclature		Feature Type	Human Disease (OMIM) <small>Caveat &amp; Help</small>	
			Symbol	Name		ID	Term
Crp	current symbol	MGI:88512	<b>Crp</b>	C-reactive protein, pentraxin-related	protein coding gene		
Adss	current symbol	MGI:87948	<b>Adss</b>	adenylosuccinate synthetase, non muscle	protein coding gene		
Dcaf8	current symbol	MGI:91860	<b>Dcaf8</b>	DDB1 and CUL4 associated factor 8	protein coding gene		
Fcer1a	current symbol	MGI:95494	<b>Fcer1a</b>	Fc receptor, IgE, high affinity I, alpha polypeptide	protein coding gene		
Fcer1g	current symbol	MGI:95496	<b>Fcer1g</b>	Fc receptor, IgE, high affinity I, gamma polypeptide	protein coding gene		
Fcgr2b	current symbol	MGI:95499	<b>Fcgr2b</b>	Fc receptor, IgG, low affinity IIb	protein coding gene	611162	Malaria, Susceptibility to Systemic Lupus Erythematosus; SLE
Fcar3	current	MGI:95500	<b>Fcar3</b>	Fc receptor, IgG, low affinity III	protein coding gene	152700	

**1. Download options**

**OMIM annotations for orthologous human genes.**

## MouseMine in a Nutshell



- Mouse Mine
  - Retrieve a list of genes in a region
  - Upload a list of genes
  - Generate annotations for a gene list
  - One click access to orthologs in other model organism “mines”
- MouseMine provides widgets for gene set enrichment and list comparison functions
  - Save gene lists
  - Create templated searches you can return to again and again

[www.mousemine.org](http://www.mousemine.org)

## MouseMine Demo

1. Search for genes in a genomic region
2. Save the list of genes (gene list 1)
3. Review the annotations for the genes
4. Find all genes associated with lung development
5. Save the list of genes (gene list 2)
6. Find the subset of genes in gene list 1 that are annotated to "lung development"
7. Save the intersected gene list (gene list 3)

**1. Find all genes in a region**

The screenshot shows the MouseMine homepage. At the top, there's a navigation bar with links for Home, Templates, Lists, QueryBuilder, Regions, Data, API, and MyMine. A search bar at the top right contains the placeholder "e.g. Parkinson, Hox\*, MGI:96677" with a "GO" button next to it. Below the search bar, there are two main sections: "Search" and "Analyze". The "Search" section has a magnifying glass icon and a text input field for "Search MouseMine...". The "Analyze" section has a text input field for "Enter a list of identifiers and/or symbols." with a dropdown menu set to "Gene" and a text entry "e.g. MGI:88388 MGI:96677 Fgf2 Bmp4". Below these are tabs for GENOME, FUNCTION, PHENOTYPE, DISEASE, HOMOLOGY, and LITERATURE. A "Read more" link and a "Query for genome:" section with various options like Features, Lookups, Coordinates, etc., are also present. On the right side, there's a "Welcome Back!" message, a "VIDEO TOUR (@FLYMINE)" button, and a "popular queries" section. At the bottom, there's information about the Perl, Python, Ruby, and Java API, as well as news and updates.

**Genomic Regions Search**

Search for features that overlap a list of genome coordinates you enter or upload, e.g. 17:11334822..11620296

Genome coordinates help

- 1. Click** 1. Select Organism: **M. musculus** genome build: **GRCm38**
- 2. Click** 2. Select Feature Types:
  - Complex Cluster Region
  - DNA Segment
  - Gene
  - Gene Segment
  - Heritable Phenotypic Marker
  - Insertion
  - Other Genome Feature
  - Polymorphic Pseudogene
  - Pseudogene
  - Pseudogenic Gene Segment
  - Syntenic Region
  - Transgene
- 3. Type or paste** 3. Type/Paste in genomic regions in  base coordinate  interbase coordinate (click to see an example) **1:147249520-183092234**
- 4. Click** 4. Extend your regions at both sides:

*For this example, we only want the genes.*

*You can enter multiple regions on the same or different chromosomes!*

or Upload genomic regions from a .txt file...  No file chosen

**2. Create the gene list.**

1. Click 

Selected organism: *M. musculus*  
Selected feature types: Gene

Export for all regions: TAB CSV GF3 BED or Create List by feature type: Gene Go Page size: 10 << First < Prev > Next > Last >>

GENOME REGION	FEATURE	FEATURE TYPE	LOCATION
1:147249520..183092234	Gm9931 MGI:3642402	Gene 	1:147280758..147281936
	Gm22966 MGI:5452743	Gene 	1:148326980..148327108
	Gm23535 MGI:5453312	Gene 	1:149707377..149707482
	n-RS218 MGI:4422083	Gene 	1:149734750..149734858
	Pla2g4a MGI:1195256	Gene 	1:149829618..149961290
	n-RS219 MGI:4422084	Gene 	1:149841530..149841661
	Ptgs2os MGI:2443180	Gene 	1:150076661..150099766
	Ptgs2 MGI:97798	Gene 	1:150100124..150108010
	Ptgs2os2 MGI:5477181	Gene 	1:150159043..150164948
	Pd MGI:98090	Gene 	1:150319417..150333905
	2310030A07RIK MGI:1922834	Gene 	1:150320681..150346771
	BC003331 MGI:2385108	Gene 	1:150361305..150393080
	Tor MGI:1922056	Gene 	1:150392838..150449935

Note that you can save this list of genes in several different formats.

**3. Explore gene annotations**

1. Click 

MouseMine Integrated mouse data. Populated by MGI, powered by InterMine. Log In | Help | FAQ | About | Contact Us MGI update: 2014-07-08

Home Templates Lists QueryBuilder Regions Data API MyMine Upload | View Search: e.g. Parkinson, Hox\*, MGI:96677 GO

List Analysis for all\_regions\_Gene\_list\_3 (561 Genes)

Manage Columns 1 Filters Create / Add to List Get Code Download Rows per page: 10 p. 1

ID	Symbol	Name	Type	Chromosome
MGI:101776	Lhx4	LIM homeobox protein 4	protein_coding_gene	1
MGI:101847	Ifi205	interferon activated gene 205	protein_coding_gene	1
MGI:101899	Pou2f1	POU domain, class 2, transcription factor 1	protein_coding_gene	1
MGI:101913	Chm1	choroideremia-like	protein_coding_gene	1
MGI:103177	Mpz	myelin protein zero	protein_coding_gene	1
MGI:104511	Tnfsf4	tumor necrosis factor (		
MGI:104593	Xcl1	chemokine (C motif) lig		
MGI:104665	Sot1	sterol O-acvtransferase		

Orthologues  
C. elegans (398) D. melanogaster (313) D. rerio (355) H. sapiens (629)  
M. musculus (52) R. norvegicus (373) S. cerevisiae (126)  
View homologues in other Mines:  
RatMine  
R. norvegicus x  
ZFINMine  
D. rerio x  
FlyMine  
D. melanogaster x C. elegans x  
modMine  
D. melanogaster x C. elegans x  
metabolicMine  
H. sapiens x M. musculus x  
YeastMine  
S. cerevisiae x

External Links  
MGI Batch Query

Explore the information on this page...there are quick links to orthologs in other organisms, gene list enrichment results, and summaries of gene annotations.

You can also forward this gene list to Batch Query!

**List Details**

Name: Gene List 1 |

Description: an optional description

Create Cancel Reset

1. Click

4. Find the genes annotated to “lung development”.

List A: All\_regions\_Gene\_list\_3 (561 Genes)

1. Click

ID	Symbol	Name	Type	Chromosome
MGI:101776	Lhx4	LIM homeobox protein 4	protein_coding_gene	1
MGI:101847	Ifi205	interferon activated gene 205	protein_coding_gene	1
MGI:101898	Pou2f1	POU domain, class 2, transcription factor 1	protein_coding_gene	1
MGI:101913	Chm1	choroideremia-like	protein_coding_gene	1
MGI:103177	Mpz	myelin protein zero	protein_coding_gene	1
MGI:104511	Tnfrsf4	tumor necrosis factor (ligand) superfamily, member 4	protein_coding_gene	1
MGI:104593	Xcl1	chemokine (C motif) ligand 1	protein_coding_gene	1
MGI:104665	Socat1	sterol O-acetyltransferase 1	protein_coding_gene	1

The screenshot shows the MouseMine homepage. At the top, there's a navigation bar with links for Home, Templates, Lists, QueryBuilder, Regions, Data, API, and MyMine. A search bar at the top right contains the text "e.g. Parkinson, Hox\*, MGI:96677" and a "GO" button. Below the navigation, there are two main sections: "Search" and "Analyze". The "Search" section has a search input field with placeholder text "Search MouseMine. Enter name, identifier, or keyword for genes, strains, phenotypes, diseases, etc." and a "SEARCH" button. The "Analyze" section has an input field for "Gene" with placeholder text "e.g. MGI:88388 MGI:96677 Fgf2 Bmp4" and a "ANALYSE" button. Below these sections are tabs for GENOME, FUNCTION, PHENOTYPE, DISEASE, HOMOLOGY, and LITERATURE. A "Welcome Back!" message on the right side mentions the InterMine framework. A yellow arrow labeled "1. Click" points to the "SEARCH" button in the search section. Another yellow arrow labeled "2. Click" points to the "FUNCTION" tab. A green diagonal banner on the right says "popular queries".

This screenshot shows the "Functions (GO terms and children) ➔ Mouse features" section of the QueryBuilder. It displays a list of mouse features associated with the specified GO terms. The interface includes a "GO Term > Parents" section with a "LOOKUP" input field containing "Lung development", a "constrain to be" dropdown set to "IN", and a "saved GOTerm list" dropdown showing "Pax GO Terms 2013-03-14". A yellow arrow labeled "1. Type" points to the "Lung development" text in the lookup field. Another yellow arrow labeled "2. Click" points to the "Show Results" button below the input fields. At the bottom, there are links for "web service URL", "Perl | Python | Ruby | Java [help]", and "export XML".

5. Save the list of lung development genes.

1. Click

Trail: Query  
**Functions (GO terms and children) ➔ Mouse features**  
 Returns mouse features associated with specified GO terms (and any children of this GO term).

Manage Columns 5 Filters Create / Add to List Get Code Download

Showing 1 to 25 of 392 rows Rows per page: 25 p. 1

Ontology Term Namespace	Ontology Annotations Ontology Term . Identifier	Ontology Annotations Term Name	Ontology Annotations Qualifier	Code Code	Evidence With Text	ID	Symbol	Name	Mgi Type
biological_process	GO:0048286	lung alveolus development	NO VALUE	IMP	MGI:3842511	MGI:1202886	Meg3	maternally expressed 3	lincRNA gene
biological_process	GO:0030324	lung development	NO VALUE	IGI	MGI:2676844 MGI:3618743	MGI:3619065	Mir17	microRNA 17	miRNA gene

1. Click

2. Click

Create New List Add to Existing List

- All 55 GOTerms
- All 356 OntologyAnnotations
- All 6 OntologyAnnotationEvidence Codes
- All 392 OntologyAnnotationEvidences
- All 197 SequenceFeatures**
- All 189 Publications
- Choose individual items from the table

**List Details**

MINIMISE CLOSE

1. Type

Name: Gene List 2

2. Type

Description: All lung development genes

3. Click

Create Cancel Reset

6. Intersect the gene lists

1. Click

MouseMine Integrated mouse data. Populated by MGI, powered by InterMine.

Log In | Help | FAQ | About | Contact Us  
MGI update: 2014-07-08

Home Templates Lists QueryBuilder Regions Data API MyMine

Search: e.g. Parkinson, Hox\*, MGI:96677 GO Hide

Success: Created a new list CytoskeletonGenesinMouse( 1839 SequenceFeatures)

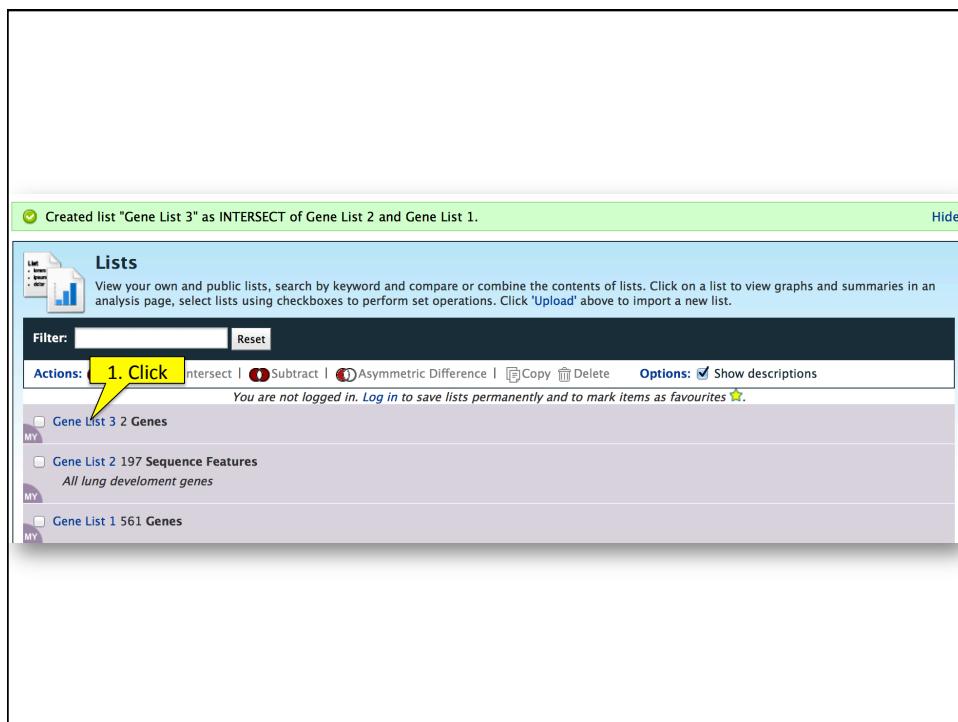
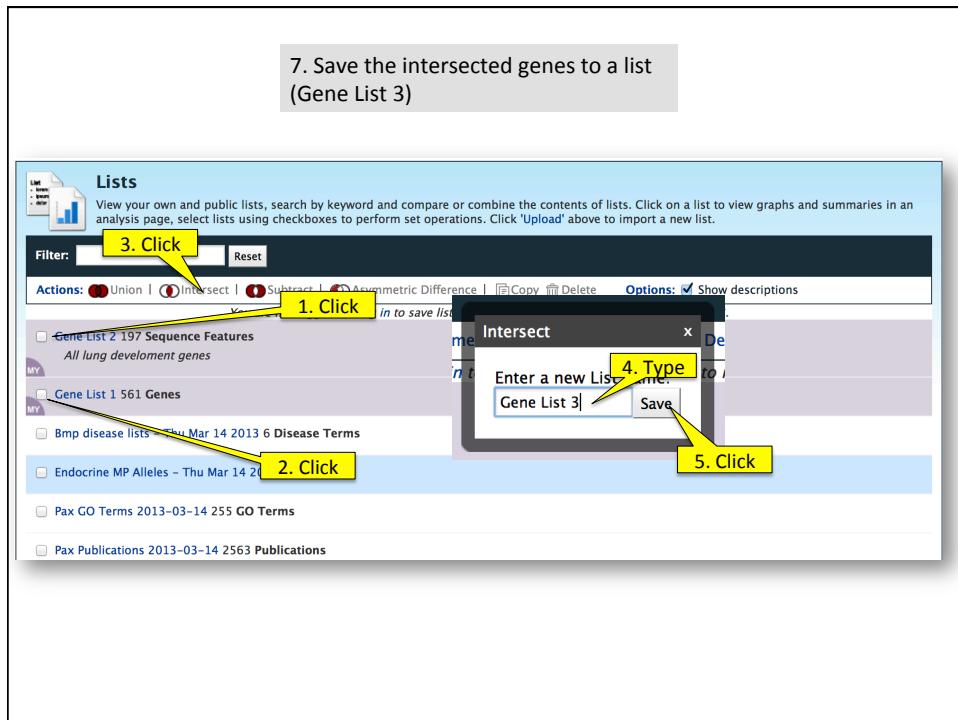
Trail: Query Functions (GO terms and children) → Mouse features

Returns mouse features associated with specified GO terms (and any children of this GO term).

Manage Columns 5 Filters Create / Add to List Get Code Download

Showing 1 to 25 of 3,963 rows Rows per page: 25 p. 1

Ontology Term Namespace	Ontology Annotations Ontology Term Identifier	Ontology Annotations Term Name	Ontology Annotations Qualifier	Code	Evidence With Text	ID	Symbol	Name	Mgi Type
cellular_component	GO:0000235	astral microtubule	colocalizes_with	IDA	NO VALUE	MGI:2442208	Map9	microtubule-associated protein 9	protein coding g
cellular_component	GO:0000235	astral microtubule	NO VALUE	ISO	UniProtKB:P43034 UniProtKB:P43034	MGI:109520	Pafah1b1	platelet-activating factor acetylhydrolase, isoform 1b, subunit 1	protein coding g
cellular_component	GO:0000235	astral microtubule	NO VALUE	ISO	UniProtKB:Q86Y91 UniProtKB:Q86Y91	MGI:2446979	Kif18b	kinesin family member 18B	protein coding g
cellular_component	GO:0000242	pericentriolar material	NO VALUE	IDA	NO VALUE	MGI:101834	Tubg1	tubulin, gamma 1	protein coding g



**List Analysis for Gene List 3 (2 Genes)**

1. Click

ID	Symbol	Name	Type	Chromosome
MGI:109284	Psen2	presenilin 2	protein_coding_gene	1
MGI:2135272	Vangl2	vang-like 2 (van gogh, Drosophila)	protein_coding_gene	1

**Description** Click here to enter a description for your list  
Date Created: 2014-07-19 13:49

**Orthologues**

- RatMine R. norvegicus
- ZFINMine D. rerio
- FlyMine D. melanogaster
- modMine D. melanogaster C. elegans
- metabolicMine H. sapiens M. musculus
- YeastMine No results found.

**External Links**  
MGI Batch Query

**Download Results**

CLOSE

Format: tsv  
5 columns  
All rows  
Output  
Destination:  
File  
Configure the export options in these categories

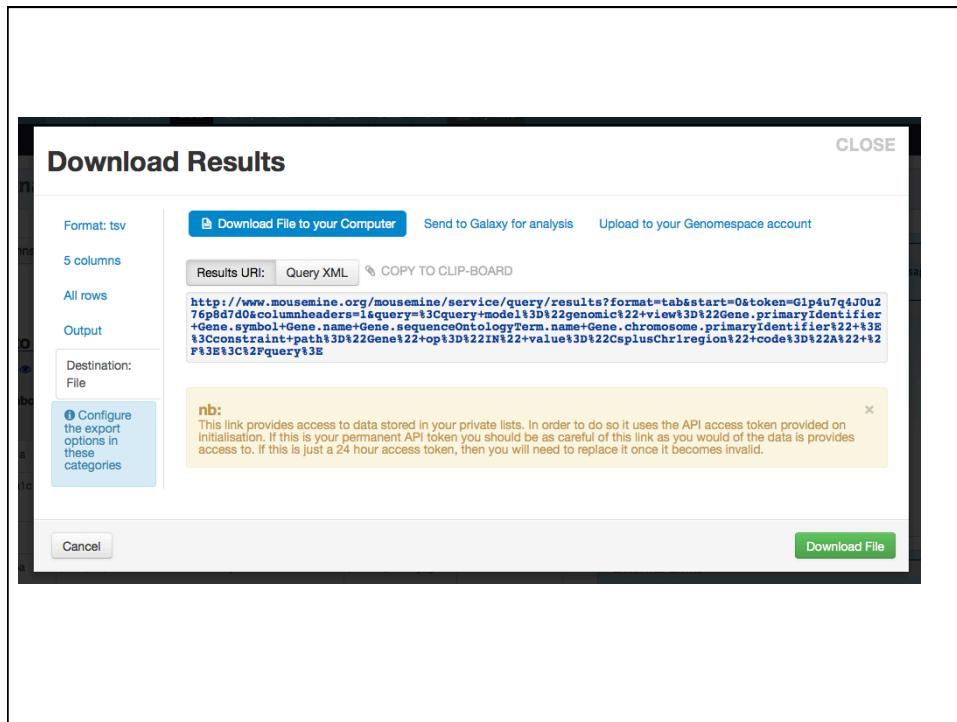
1. Click

**Format**

Spreadsheet (tab separated values)  
 Spreadsheet (comma separated values)  
 XML  
 JSON  
 GFF3 (General Feature Format)  
 UCSC-BED (Browser Extensible Display Format)  
 FASTA sequence

**Download results or forward results to another tool such as Galaxy or Genome Space using Destination**

Cancel Download File



BONUS 1: What are **all** the phenotypes associated with those 561 genes in the Chromosome 1 region?

- Use the template search/query tools in MouseMine

**MouseMine** Integrated mouse data. Populated by MGI, powered by InterMine.

Log In | Help | FAQ | About | Contact  
MGI update: 2013-06-10

Home Templates Lists QueryBuilder Regions Data API MyMine

Search: e.g. Parkinson, Hox\*, MGI:96677 GO

**Search**  
Search MouseMine. Enter name, identifier, or keyword for genes, strains, phenotypes, diseases, etc.  
e.g. Parkinson, Hox\*, MGI:96677

**Analyze**  
Enter a list of identifiers and/or symbols.  
Gene: e.g. MGI:88388 MGI:96677 Fgf2 Bmp4

**Welcome Back!**  
MouseMine provides integrated mouse data from MGI, via the uniquely powerful InterMine framework. Please see this [description](#) of MouseMine.

**VIDEO TOUR (@FLYLINE)**

**GENOME** **FUNCTION** **PHENOTYPE** **DISEASE** **HOMOLOGY** **LITERATURE**

[Read more](#)

Query for phenotype:

- Mammalian phenotypes (MP terms) → Mouse genotypes
- Human genes → Phenotypes of mouse orthologs
- Mouse features → Mammalian phenotypes (MP terms)
- Lookup → Mammalian phenotypes (MP terms)
- Mouse alleles → Genotypes (models)
- Mouse alleles → Human diseases (MEDIC terms)
- Publication → Mammalian phenotypes (MP terms)
- Mouse alleles → Mammalian phenotypes (MP terms)

» [More queries](#)

**popular queries**

**Mouse features → Phenotypes (MP terms), with genotypes**  
Returns phenotypes (MP terms) associated with specified mouse features along with the genotypes (background strain+mutations) in which those phenotypes were observed.

**1. Click** **Sequence Feature**  
LOOKUP: Fgf8

**2. Select** **constraint to be**: IN saved SequenceFeature list: Gene List 1

**3. Click** **Show Results**

[Edit Query](#)

[web service URL](#) | Perl | Python | Ruby | Java [help] | export XML

The screenshot shows the MouseMine web application interface. At the top, there is a navigation bar with links for Home, Templates, Lists, QueryBuilder, Regions, Data, API, and MyMine. A search bar contains the query "e.g. Parkinson, Hox\*, MGI:96677". Below the search bar, a message states "Showing 1 to 25 of 3,687 rows". The main content area displays a table with columns: ID, Symbol, Base Annotations, Subject . Symbol, Background, Ontology Annotations, Ontology Term . Identifier, and Term Name. The table contains three rows of data, each corresponding to the gene Lhx4.

ID	Symbol	Base Annotations	Subject . Symbol	Background	Ontology Annotations	Ontology Term . Identifier	Term Name
MGI:101776	Lhx4	Lhx4<tm1Ssp>/Lhx4<tm1Ssp>		involves: 129S2/SvPas * CF1	MP:0001175		abnormal lung morphology
MGI:101776	Lhx4	Lhx4<tm1Ssp>/Lhx4<tm1Ssp>		involves: 129S2/SvPas * CF1	MP:0002329		abnormal blood gas level
MGI:101776	Lhx4	Lhx4<tm1Ssp>/Lhx4<tm1Ssp>		involves: 129S2/SvPas * CF1	MP:0003817		abnormal pituitary diverticulum morphology

**BONUS 2:** What if you wanted to add genes associated with lung *phenotypes* to your list of lung *development* genes?

- Search for genes associated with lung phenotypes in mouse.
- Add those genes to Gene List 2 (the lung development genes).
- Re-run the list intersection function.

1. Click

MouseMine Integrated mouse data. Populated by MGI, powered by InterMine.

Log In | Help | FAQ | About | MGI update

Home Templates Lists QueryBuilder Regions Data API MyMine

Search: e.g. Parkinson, Hox\*, MGI:96677 GO

**Search**  
Search MouseMine. Enter name, identifier, or keyword for genes, strains, phenotypes, diseases, etc.  
e.g. Parkinson, Hox\*, MGI:96677

**Analyze**  
Enter a list of identifiers and/or symbols.  
Gene  
e.g. MGI:88388 MGI:96677 Fgf2 Bmp4

**Welcome Back!**  
MouseMine provides integrated mouse data from MGI via the uniquely powerful InterMine framework. Please see this [description](#) of MouseMine.

**VIDEO TOUR (@FLYMINE)**

GENOME FUNCTION EXPRESSION PHENOTYPE DISEASE HOMOLOGY LITERATURE

Read more

Query for genome:

1. Click

MouseMine Integrated mouse data. Populated by MGI, powered by InterMine.

Log In | Help | FAQ | About | MGI update

Home Templates Lists QueryBuilder Regions Data API MyMine

Search: e.g. Parkinson, Hox\*, MGI:96677 GO

**Templates**  
Templates are predefined queries, each having a name and a description. Click on a template to run it, you can search for templates by keyword and filter them by category.

Filter: -- all categories -- Reset

Actions: Export selected Options:  Show descriptions

You are not logged in. Log in to mark items as favourites ★.

<input type="checkbox"/> Genes --> Homologs Returns homologs of specified genes.	Filter: -- all categories -- Reset
<input type="checkbox"/> Publication --> Functions (GO terms) Returns gene functions (GO terms) reported in specified publication(s).	Options: <input checked="" type="checkbox"/>
<input type="checkbox"/> Diseases (MEDIC terms) --> Mouse models Returns mouse genes associated with (modeling) specified human diseases.	2. Click
<input type="checkbox"/> Human genes --> Phenotypes of mouse orthologs Returns the phenotypes of mice with mutations in the orthologs of specified genes.	Disease Function Expression Phenotype Disease Homology Literature

in. Log in to mark items as favourites ★.

Options:

Filter: -- all categories -- Reset

Disease  
Function  
Expression  
Phenotype  
Disease  
Homology  
Literature

**Templates**

Templates are predefined queries, each has a simple form and a description. Click on a template to run it, you can search for templates by keyword and filter them by category.

Filter:  Filter: **Phenotype** Reset

Actions: [Export selected](#) Options:  Show descriptions

[Log in](#) to mark items as favourites ★.

1. Click

- Human genes --> Phenotypes of mouse orthologs Returns the phenotypes of mice with mutations in the orthologs of the specified human genes.
- Mammalian phenotypes (MP terms) --> Mouse genotypes Returns mouse genotypes (models) associated with the specified phenotypes (MP terms).
- Mouse features --> Phenotypes (MP terms), with genotypes Returns phenotypes (MP terms) associated with specified mouse features along with the genotypes (background strain+mutations) in which those phenotypes were observed.
- Lookup --> Mammalian phenotypes (MP terms)

**Mammalian phenotypes (MP terms) ➔ Mouse genotypes**

Returns mouse genotypes (models) associated with the specified phenotypes (MP terms).

MP Term > Parents

LOOKUP: \*circulating glucose\* ⓘ

Show Results Not sure what term to type in? See next two slides! Edit Query

web service URL Perl | Python | Ruby | Java [help] export XML

MGI uses standard vocabularies for annotating gene function, phenotypes, etc.

Use the vocabulary browsers at MGI to get ideas for what terms to use in the MouseMine searches.

[www.informatics.jax.org](http://www.informatics.jax.org)

mammalian\_phenotype [MP:0000001] (56201 genotypes, 269369 annotations)

- | adipose tissue phenotype +
- | behavior/neurological phenotype +
- | cardiovascular system phenotype +
- | cellular phenotype +
- | craniofacial phenotype +
- | digestive/alimentary phenotype +
- | embryogenesis phenotype +
- | endocrine/exocrine gland phenotype +
- | growth/size/body phenotype +
- | hearing/vestibular/ear phenotype +
- | hematopoietic system phenotype +
- | homeostasis/metabolism phenotype +
- | immune system phenotype +
- | integument phenotype +
- | limbs/digits/tail phenotype +
- | liver/biliary system phenotype +
- | mortality/aging +
- | muscle phenotype +
- | nervous system phenotype +
- | no phenotypic analysis
- | normal phenotype +
- | other phenotype +
- | pigmentation phenotype +
- | renal/urinary system phenotype +
- | reproductive system phenotype +
- | respiratory system phenotype +
- | skeleton phenotype +
- | taste/olfaction phenotype +
- | tumorigenesis +
- | vision/eye phenotype +

These are the high level Mammalian Phenotype (MP) terms used to annotated mouse phenotypes in MGI.

1. Type \*respiratory system phenotype\*

**Mammalian phenotypes (MP terms)** ➔ **Mouse genotypes**

Returns mouse genotypes (models) associated with the specified phenotypes (MP terms).

**MP Term > Parents**

LOOKUP:  ⓘ

**Show Results** **Edit Query**

[web service URL](#) [Perl](#) [Python](#) [Ruby](#) [Java \[help\]](#) [export XML](#)

1. Click

2. Click

3. Click

**Mammalian phenotypes (MP terms)** ➔ **Mouse genotypes**

Returns mouse genotypes (models) associated with the specified phenotypes (MP terms).

**Show Results** **Create / Add to List** **Add to Existing List**

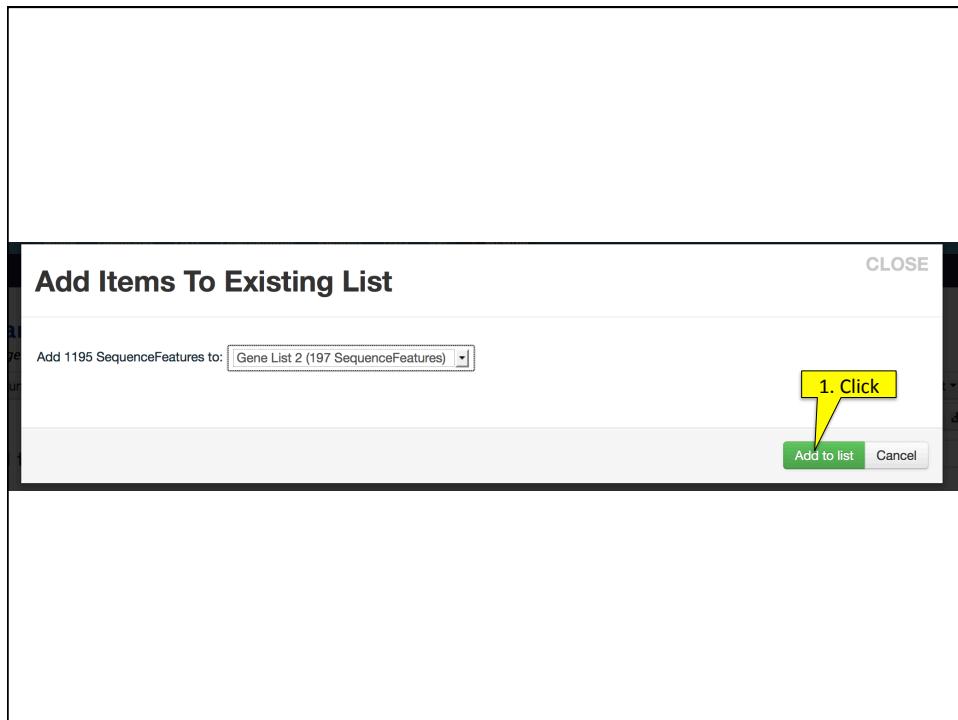
Showing 1 to 25 of 4,093 rows

Subject ID	Subject . Symbol	Base Annotations - Subject . Symbol
MGID:1920942	2410089E03Rik	2410089E03Rik<cb2b012Clo>2410089E03R
MGID:99607	Abca1	Abca1<tm1Jdm>/Abca1<>
MGID:99607	Abca1	Abca1<tm1Jdm>/Abca1<tm1Jdm>
MGID:99607	Abca1	Abca1<tm1Jdm>/Abca1<tm1Jdm>
MGID:99607	Abca1	Abca1<tm1Wpfl>/Abca1<>
MGID:99607	Abca1	Abca1<tm1Wpfl>/Abca1<>
MGID:99607	Abca1	Abca1<tm1Wpfl>/Abca1<tm1Wpfl>

**Create New List** **Add to Existing List**

- All 1195 SequenceFeatures
- All 1893 Genotypes
- All 835 Strains
- All 276 MPTerms
- Choose individual items from the table

MP:0001178



**Lists**  
View your own and public lists, search by keyword and compare or combine the contents of lists. Click on a list to view graphs and summaries in an analysis page, select lists using checkboxes to perform set operations. Click 'Upload' above to import a new list.

Filter:  Reset

Actions:  Union |  Intersect |  Subtract |  Asymmetric Difference |  Copy  Delete Options:  Show descriptions  
You are not logged in. Log in to save lists permanently and to mark items as favourites.

- Gene List 3 2 Genes
- Gene List 2 1265 Sequence Features  
All lung development genes from GO and lung phenotype genes from MP
- Gene List 1 561 Genes

You can see that Gene List 2 has more genes in it now. Click to see the list and update the list description.

Gene List 4 27 Genes

Gene List 3 2 Genes

Gene List 2 1265 Sequence Features  
All lung development genes from GO and lung phenotype genes from MP

Gene List 1 561 Genes

After intersecting Gene List 1 with the updated Gene List 2 you get 27 lung related genes instead of 2.

## Your Turn!

- There are many heritable phenotypes in the MGI database where the phenotype has been mapped but the causative gene is not known.
- Eccp (eccentric pupil) has been mapped to chromosome 12 (Chr12:35901444-82619165; GRCm38)
  - Look up Eccp locus in MGI
    - How was this phenotype identified?
    - What strains were used for mapping?
    - Is the mouse strain available?
  - Use MouseMine to find genes in the mapped region that are associated with eye phenotypes
  - Use the list of genes you find using MouseMine as input into the SNP query form @ MGI to see if there are any SNPs in these genes between the strains used to map the phenotype

MGI: [www.informatics.jax.org](http://www.informatics.jax.org)

MouseMine: [www.mousemine.org](http://www.mousemine.org)