



An Introduction To Mouse Genome Informatics

Mouse Genome Informatics (MGI) provides free, public access to integrated data on the genetics, genomics and biology of the laboratory mouse. In this self-guided tutorial you will use MGI to:

- learn how to navigate the MGI website efficiently
- locate mouse genotypes associated with specific phenotypes
- find mouse models of human diseases
- find candidate genes in a region of interest
- identify distributors of stocks carrying mutations of interest
- identify strain-specific SNPs and sequences
- compare chromosomal regions between strains
- find gene expression assays and images for specific anatomical structures
- perform batch queries

Phenotypes

Alleles of the human *APOE* gene are implicated in hypercholesterolemia, atherosclerosis and Alzheimer's Disease. How can you find more information on the mouse *Apoe* gene?

1. Go to the Mouse Genome Informatics (MGI) (www.informatics.jax.org) web site, enter *apoe* in the Quick Search field at the top of the web page and click the *Quick Search* button.
2. Click on the *Apoe* gene symbol to see its Gene Detail page, the page with links to all the information we have on this mouse gene, including genome coordinates and links to sequences, alleles, SNPs, Gene Ontology, expression data and references.

How can you find information on the different mouse *Apoe* alleles?

1. In the Mutations, Alleles and Phenotypes section of the *Apoe* Gene Detail page, click on the All Mutations and Alleles link. This will display a summary of the known alleles of the mouse gene and show official and unofficial (synonyms) nomenclature, and any high-level phenotype terms and human disease model information associated with each allele.

This allele summary provides the best overview of the alleles but what if you were more interested in seeing more specific phenotype data for the *Apoe* alleles?

1. Backup to the *Apoe* Gene Detail page and click on the first link for the Phenotype Summary in the Mutations, Alleles and Phenotypes section, the ### phenotypes from ## alleles in ## genetic backgrounds.

This will show the lower level phenotype terms for each single mutant allele genotype. (For multiple-allele genotypes, use the second link for the Phenotype Summary, phenotypes from multigenic genotypes.)

Jan Breslow's lab has knocked out the mouse gene. What percent of these mice on a 129P2/OlaHsd * C57BL/6 background develop cerebral xanthoma on a high cholesterol diet?

1. Scroll down and click on the *Apoe^{tm1Bres}* allele symbol.
2. In the Phenotypes section of this Allele Detail page, click on hm3 (hm for homozygous) for this strain background.

Genotype/ Background:	Genotype	Genetic Background
hm1	<i>Apoe^{tm1Bres}/Apoe^{tm1Bres}</i>	B6.129P2- <i>Apoe^{tm1Bres}</i>
hm2	<i>Apoe^{tm1Bres}/Apoe^{tm1Bres}</i>	either: (Involves: 129P2/OlaHsd * BALB/c) or (Involves: 129P2/OlaHsd * C57BL/6J)
hm3	<i>Apoe^{tm1Bres}/Apoe^{tm1Bres}</i>	involves: 129P2/OlaHsd * C57BL/6

(Note: *Apoe^{tm1Bres}* means this is the first targeted mutation (tm1) in *Apoe*, from the Jan Breslow lab (Bres)).

Find Mice (IMSR)

How can you order these mice?

1. At the top of the page, in the Find Mice section, notice that though there are 10 mouse strains available carrying this mutation and there are 82 strains or lines available carrying any *Apoe* mutation. Click the 10 strains available link to search the IMSR.
2. This will search for stocks with this *Apoe* mutation from several providers. In your IMSR search results, the Strain Name links to the repository's site, the Repository links to an email form so you can ask about the stock and the Allele Symbol links to the MGI Allele Detail page.

Note: The Find Mice (IMSR) section also appears on Allele Detail pages. The IMSR (International Mouse Strain Resource) is a searchable online database of mouse strains and stocks available worldwide. Neither it, nor MGI, house any stocks. You can search the IMSR directly by clicking on the *Find Mice (IMSR)* link in the dark blue banner near the top of most MGI web pages. If a desired allele is unavailable, it usually means the authors didn't contribute the stock to a public repository. Your best option then is to try contacting the authors of the original publication, noted at the bottom of the allele page.

Recombinase (cre) Activity

How you can you find recombinase (cre) alleles detected in muscle?

1. Return to an MGI web page, then select Recombinase (cre) from the Search Menu or click on the blue Recombinases tab on most MGI web pages.
2. In the Recombinase activity assayed in: field, type: *skel* and select skeletal muscle from the list. Note: items are grayed out when we have no cre activity data for that tissue.
3. Click the *Search* button.
4. Scroll down to the Tg(Camk2a-cre/ERT2)1Aibs allele, where *Camk2a* is the driver (promoter) and click on the Matrix View icon in the second column. This allows you to compare endogenous wild-type *Camk2a* gene expression with reported recombinase activity data for this allele (highlighted) and other *Camk2a* driven alleles.
5. Return to the cre search results page and click on the allele symbol, Tg(Camk2a-cre/ERT2)1Aibs, to access the Transgene Details.
6. On the Transgene Detail page, in the Recombinase Activity section, click the blue triangle after muscle to expand the section.
7. Click on the check mark in the Images column for muscle.
8. The pattern of expression is described and the images can be resized and moved.
9. Note that in the top section of the page, 1 strain is available carrying this transgene.

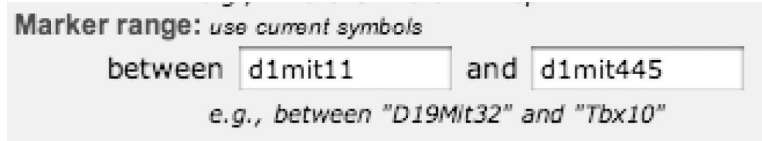


The Gene Ontology (GO) project uses defined vocabularies to describe the biological processes, molecular functions and cellular components associated with a gene's products. How can you find the GO terms associated with the *Lrp4* gene?

1. Go to the MGI *Lrp4* Gene Detail page and scroll down to the Gene Ontology (GO) Classifications section.
2. Click on the All GO Annotations link.
3. Jump down the page by clicking on the Tabular View link.
4. Click on any Classification Term to view its place in the vocabulary hierarchy, see its definition, and to find a link to all the genes annotated to the term.

You are interested in a QTL defined by MIT markers, *D1Mit11* and *D1Mit445*. How can you see all the genes mapped between these markers?

1. Go to Search beneath the MGI logo on the upper left hand side of an MGI web page and select Genes —> Genes & Markers Query.
2. In the Feature type section, check protein coding gene. (Click *Show* to see all the types and the counts for each.)
3. In the Genome location section of the page, enter d1mit11 and d1mit445 in the Marker range fields.
4. Click the *Search* button



Marker range: use current symbols
between and
e.g., between "D19Mit32" and "Tbx10"

Now that you have all the candidate genes in this region, you can search for stronger candidates by adding phenotype or GO terms to your search parameters. Use the *Click to modify search* button above the results. Also notice that you have four *Export* options.

Human - Mouse: Disease Connection

You are interested in regions of two human chromosomes. How can you see phenotypes and diseases associated with genes in these regions?

1. Click on the *Human Disease* tab at the top of an MGI web page.
2. Toggle “Please select a field” to “Genome Location.”
3. Select the species and build: Human (GRCh38).
4. Enter the region: 19:43341000-44910000 (This format is also supported: Chr19:43341000-44910000)
5. Click the *Add* button.
6. Toggle the second “Please select a field” to “Genome Location.”
7. Select the species and build: Human (GRCh38).
8. Enter the region: 21:25001000-26971000
9. Toggle the default Boolean from AND to OR.
10. Click the *Search* button.

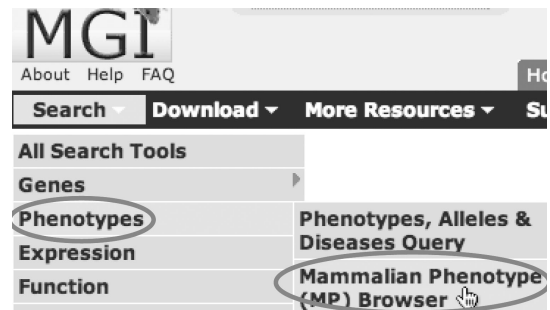
The search finds human genes mapped to those regions and returns them, their mouse homologs, and phenotypes and diseases associated with all the returned genes.

11. Scroll out to the right to see diseases associated with these genes. Blue squares indicate mouse data and tan is for human data.
12. Click on the blue square at the intersection of *PLAUR* and immune system to see the mouse *Plaur* genotypes with immune system abnormalities.

Vocabulary Browsers

Because free text searches can produce spurious results, MGI creates and annotates to defined vocabularies such as the Mammalian Phenotype (MP) Ontology. The Ontology can be searched or browsed. For example, some *Lrp4* alleles display abnormal motor neuron morphology. How can you find other alleles that show this phenotype?

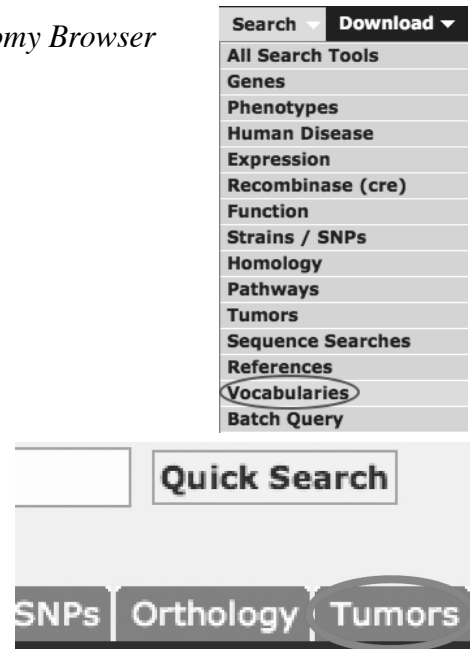
1. You can search for this phrase in MGI's "Genes and Markers," "Phenotypes and Alleles," and "Mammalian Phenotype Browser" forms but today, just browse the Phenotype Browser. Go to Search beneath the MGI logo on the upper left hand side of an MGI web page and select Phenotypes → Mammalian Phenotype (MP) Browser. (Note: you can always return to the MGI Home Page by clicking on the MGI logo.)
2. You can search or browse the ontology. Today, browse by clicking on *nervous system phenotype*.
3. Click on *abnormal nervous system morphology* (a black triangle preceding a term means the term has subterms).
4. Click on *abnormal neuron morphology* and then *abnormal motor neuron morphology*.
5. Click on the number of genotypes.



The MGI *GO Browser* and the *Mouse Developmental Anatomy Browser* work in a similar manner.

Mouse Tumor Biology

MGI also provides a database (MTB) of mouse tumor biology information. To access MTB, click on the light blue *Tumors* tab below the Quick Search button on the upper right hand side of most MGI web pages.



Or, on the MGI Home Page, click on *Mouse Models of Human Cancer*.

For now, just take a peek at the *Tumor Frequency Grid* under Additional Resources in the left hand column of the web page.

1. The grid displays published tumor incidence for inbred strains. Click the blue triangles to view substrain and sub-tissue information.

Patient Derived Xenograft (PDX) models for cancer research are created by the implantation of human cells and tumor tissue into immune compromised NOD-SCID-Gamma2 mouse hosts. How can you find SNPs/Variants for a given tumor?

1. In the Additional Resources section on the left hand side of the page, click on the new *PDX Finder*.

You can think of the PDX Finder as an IMSR for PDX Models, and model providers worldwide contribute to the web resource.

2. In the Search field, start typing breast and select *Breast Cancer* from the list that appears.
3. Currently more data are available for JAX mice, so in the left hand side of the page, for PDX Model, click on *DATASOURCE* and check *JAX*.
4. In your results, click on the Model: *TM00091*.
5. Though some Genomic Data are available through this site, more information is available at MTB, so click on the button to *VIEW DATA AT JAX*.
6. Click to expand the Variant Summary section.
7. You can also expand the Gene Expression section or see Copy Number Variants by expanding the Gene CNV section. Dosing studies are available at the bottom of the page.

Sequences

How can you get the *Lrp4* C57BL/6J genomic sequence plus 5 Kb of flanking sequence?

1. Go to the MGI *Lrp4* Gene Detail page, scroll down to the Sequences & Gene Models section of the page and select *genomic* and add 5 to the Flank box to the right of the genomic sequence.
2. Click the *Go* button to download the FASTA file.

You are also interested in the 129S1/SvImJ and BALB/cJ genomic sequences for the gene. How can you get those?

1. In the Strain Comparison section of the Gene Detail page, click the more> button.

2. In the table that appears, use the last column to select the strains.
3. Above the table, toggle “Send to Multiple Genome Viewer” to “Get FASTA” and click the *Go* button.

How can you find exon information for the *Lrp4* gene in these 3 strains, C57BL/6J, 129S1/SvImJ and BALB/cJ?

1. Go to Search beneath the MGI logo on the upper left hand side of an MGI web page and down near the bottom, select *MouseMine*.
2. On the MouseMine page, search for *Lrp4*.
3. Click on *MGI:2442252 | Lrp4* to access the mouse data.
4. Click on Canonical gene --> Exons in specific strains.
5. This will show you information on the exons for multiple transcripts for 19 mouse inbred strains. Click on the last icon in the Strain Name column header, the graph icon, and filter for the 3 strains.
6. You can use the Export button to get all the exons, but for now, just get the first exon for the first 129S1/SvImJ transcript by clicking MGP_129S1SvImJ_E0056125_1 in the Exon Primary Identifier column.

Note: the MGP (Sanger Mouse Genomes Project) IDs can be decoded:

- MGP_129S1SvImJ_G0026150 The **G** tells you this is the genomic sequence.
- MGP_129S1SvImJ_T0056125 The **T** tells you this a transcript, number T0056125.
- MGP_129S1SvImJ_E0056125_1 the **E** tells you it's an exon, the **_1** gives you the exon number and 0056125 tells you it is associated with transcript 0056125.

7. Click the *FASTA* button to get the sequence.

Strains and Polymorphisms

Four variants of the aryl-hydrocarbon receptor (*Ahr*) gene are known. C57BL/6 carries the responsive *Ahr^b* allele; DBA/2 carries nonresponsive *Ahr^d*. Later work identified two other alleles conferring response. Thus, the allele in C57, C58, and MA/My strains is now *Ahr^{b-1}*; *Ahr^{b-2}* is carried by BALB/cBy, A, and C3H; and *Ahr^{b-3}* by *Mus spretus*, *M. caroli*, and MOLF/Ei. The nonresponsive strains, AKR, DBA/2, and 129, carry *Ahr^d*.

MGI currently gets its SNP data from NCBI's dbSNP. How can you compare SNPs for nonresponsive and b-2 strains?

1. Go to Search beneath the MGI logo on the upper left hand side of an MGI web page and select Strains / SNPs —> SNP Query.
2. In the Associated Genes field, enter *Ahr*.
3. In the Strains and Strain Comparison section, select *Yes* to compare to one or more Reference strains.

4. Check AKR/J and DBA/2J (nonresponsive strains) in the orange R (for Reference strain) column.
5. For the Comparison strains, click the Clear All button, then check A/J, BALB/cByJ and C3H/HeJ (*Ahr^{b-2}* strains) in the blue C (for Comparison strain) column.
6. Click the *Search* button.

The reference strains are listed first in your results, though you can drag strain columns to reposition them. Now you can filter your results.

7. Above the results table, click the button to Filter SNPs by Allele Agreement.
8. Select the middle option: “All reference strains agree and all comparison strains differ from reference” and click the Filter button.

Near the top of your results are three Coding-NonSynonymous SNPs where the b-2 variant strains all differ from the two nonresponsive strains. One SNP has a ? as an allele call for DBA/2J. What does that mean?

8. Click the MGI SNP Detail for this SNP.

You will see that one assay has the *C* allele whereas another makes a *T* call for DBA/2J.

How can you find out more about the DBA/2J strain?

1. Search for the strain in the MGI Quick Search field.
2. In the Vocabulary Terms section of your results, click on *DBA/2J* to access the Strain Detail page.

The Strain Detail page shows mutations and QTL associated with the strain as well as phenotypes and diseases using the strain background. For strains with SNP data, it has a SNP Profile Heat Map.

3. Click the more> button in the SNPs section of the page.
4. Mouse over a cell to see the number of SNPs shared between the page’s strain and a Comparison Strain.
5. Click on a cell to see the SNPs.

Multiple Genome Viewer

In the top, Summary, section of the DBA/2J Strain Detail page, there are links to strain trait information at the Mouse Phenome Database (MPD) and to the MGI Multiple Genome Viewer (MGV). In this example, we will use another way to access the MGV.

1. Click on *Analysis Tools* in the dark blue banner at the top of an MGI web page.
2. Click on *Multiple Genome Viewer (MGV)*.

3. Leave C57BL/6J as the Reference Genome and in the Genomes box, Command (⌘)-click on a Mac, or control-click on a PC, to select CAST/EiJ, DBA/2J, and FVB/NJ.
4. Below the Genomes box is the Find genes box. Toggle “by phenotype or disease” to “by pathway.”
5. In the field, enter: *Signaling by EGFR* and hit your return/enter key.
6. This search will appear in the My lists box and will indicate the number of genes annotated to the pathway in Reactome. The genes are also represented by pink dots on the chromosomes.
7. Click on the upper pink dot on Chromosome 7.
8. This will highlight the gene in the sections below. Mouse over features for their symbols.
9. Play with the Zoom icons.
10. Pink shaded regions are inverted compared to your selected reference genome. Orange shaded regions are translocations. Yellow shaded regions are both inverted and translocated. (You will probably have to zoom out to see any of these.)

Alliance of Genome Resources

The Alliance of Genome Resources was created to provide a consistent user interface with an integrated view of data from 6 Model Organism Databases and the Gene Ontology (GO) Consortium. The founding database members of the Alliance are Flybase, Mouse Genome Database (MGI), Rat Genome Database (RGD), Saccharomyces Genome Database (SGD), Wormbase, and the Zebrafish Information Network (ZFIN). In their 1.0 release, AGR allowed users to search for, and view, gene, functional (GO) data, and disease associations from the above seven sources. A recent release now includes alleles. Orthology data for human and Alliance organisms are built on the DRSC Integrative Ortholog Prediction Tool (DIOPT), which integrates several prediction approaches.

How can you find functional information for orthologs of a binding protein whose mouse gene symbol is *Fkbp5*?

1. Go to The Alliance home page at www.alliancegenome.org, or click on their logo at the top right of an MGI web page.
2. Enter *Fkbp5* in the search field and toggle “All” to “Gene.”

Matching results are highlighted in yellow. For example, the FKBP59 *Drosophila* gene is matched through orthology to *Fkbp5* genes.

3. Click on the *fkbp5* gene symbol for *Danio rerio*.
4. Scroll down to the Function – GO Annotations section and click on the colored cell for protein metabolism.
5. This will show you annotations made using zebrafish.

The Alliance does not yet have phenotype information for zebrafish. For that, go to the mouse gene:

6. Back up to your search results and click on the mouse *Fkbp5* gene symbol to access the mouse gene page.
7. A phenotype summary is provided by MGI in the Description, and MP terms are shown further down the page.

Expression Data

Mutations in the human fibulin 5 gene (FBLN5) cause the disease, Cutis Laxa, which is characterized by pulmonary emphysema. Mice with *Fbln5* mutations also exhibit pulmonary emphysema. *Fbln5* has been assigned the Gene Ontology (GO) Biological Process term, cell adhesion. How can you find cell adhesion genes that are expressed in the lung?

1. Go to Search beneath the MGI logo on the upper left hand side of an MGI web page and select Expression —> Gene Expression Data Query.
2. In the Genes section of the page, use the right hand field and type: *cell adhesion* and select the term from the list. Note that the number of genes annotated to the term is shown.
3. In the Anatomical Structure or Stage section, select “detected in.”
4. In the Anatomical Structures field, type: lung and select the term from the list.
5. To reduce the number of results for this example, we are going to limit the search to Assay type: RNA in situ. An easy way to do this is to click the check mark for “Find expression data in any assay type” to uncheck all assay types and then select RNA in situ.
6. Click the *Search* button.
7. Click the Assays tab and then click on the data link for the *Adamts18* Assay, MGI: 4822986.
8. Click on the Figure link for the Structure, TS23: left lung.
9. Beneath the image, in the Other Database Links section, click on the [Eurexpress](#) link. At the Eurexpress site you can zoom in on the image or view a different serial section.

Batch Queries

How can you find information on several genes at once?

1. On an MGI web page, click on Download in the dark blue banner and select Batch Query.



2. In the text field, enter a column of IDs and/or gene symbols, for example:

P08226
Shank3
App

3. In the Output section of the page, check Nomenclature, Genome Location and Gene Ontology (GO). (GO describes the molecular functions, biological processes and cellular locations of the gene product.)
4. You can export your results in spreadsheet or Tab-delimited formats or forward it to MouseMine (see the next example).

MouseMine

We have seen that MGI makes it easy to find genotypes associated with particular phenotypes but the web forms currently do not allow you to export genotype details in tab-delimited formats. If you used the Mammalian Phenotype (MP) Browser to find genotypes associated with emphysema (MP:0001958) and abnormal lung elastance (MP:0011043), how can you get a text file of those genotypes?

You can do this (and much more) with our MouseMine, www.mousemine.org.

1. Go to Search beneath the MGI logo on the upper left hand side of an MGI web page and select MouseMine.
2. Click on the Templates tab near the top of the page.
3. Click on Mammalian phenotypes (MP terms) --> Mouse genes and models.
4. Enter MP:0001958, MP:0011043 in the LOOKUP field and click the Show Results button.
5. Toggle the 25 results so you can see all the results.

This returns a list of genes and transgenes and genotypes associated with these phenotypes terms and their subterms. Now we can modify the query:

6. Click on the Manage Columns button and then click on Add a Column.
7. Click on MGI Type the first time it is listed (under Subject).
8. Click Add 1 new column.
9. Click Apply Changes.

10. This adds the gene feature types to the table. Now you can remove columns you don't want to see by clicking the X in the column heading and you can filter by clicking the last icon in a column heading, the graph icon.
11. Click the Export button to select your formatting options.

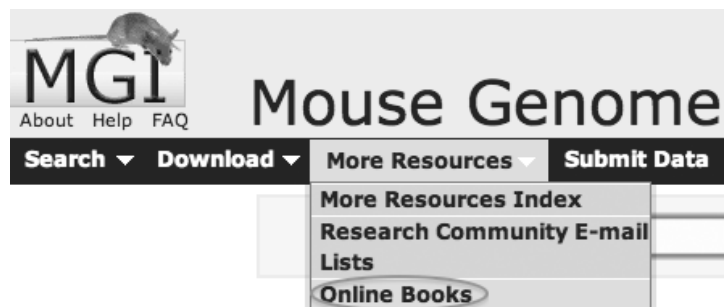
References

An overlooked feature of MGI is the ability to search publications we have curated.

1. Go to Search beneath the MGI logo on the upper left hand side of an MGI web page and select References —> Reference Query.
2. Try an author and/or text search. (The Author and Journal fields feature autofills.)
3. Click on **abstract** ► in the Title column to view an abstract.
4. The first column provides links to PubMed and Journals and the Curated Data column links to data curated from the reference and is also a way to find official nomenclature for genes and alleles researched in a publication.

Mouse Genome Informatics (MGI) has made electronic versions of several great out-of-print books available through our web site, including: *Biology of the Laboratory Mouse* by the Staff of The Jackson Laboratory, edited by Earl L. Green, *The Anatomy of the Laboratory Mouse* by Margaret J. Cook, *Mouse Genetics* by Lee Silver, *The Coat Colors of Mice* by Willys K. Silvers, *Color Atlas of Neoplastic and Non-neoplastic Lesions in Aging Mice* by C. H. Frith and J. M. Ward and *Origins of Inbred Mice*, edited by Herbert C. Morse III.

1. Go to the MGI web site and click on *More Resources* in the dark blue banner and select *Online Books*.
2. Click on *Biology of the Laboratory Mouse*.
3. Click on the cover to open the book.
4. Scroll down to the bottom of the left hand frame of the page and click on *Image Index*.
5. Click on one of the Plates or other Figures.



If you have any questions on the Mouse Genome Informatics databases, contact mgi-help@jax.org, or click the *User Support* link at the very bottom of our web pages.