The Mouse Genome Informatics online resource:

Open a browser and go to www.informatics.jax.org to begin.

If necessary for assistance while completing this worksheet, or for future reference, see the "Introduction to mouse genetics" and "How to use MGI" in the Getting Started box below topic specific search tools on the home page.

For assistance with completing Section 6 questions (the Human-Mouse: Disease Connection), or for future reference, see "Take a tour of the Human-Mouse Disease Connection" on the HMDC home page. The HMDC home page can be accessed by clicking the link from the MGI homepage, or bookmarked directly at http://diseasemodel.org.

For assistance with completing Section 7 questions (MouseMine) or for future reference, see the Help link in the top right corner of MouseMine pages. MouseMine can be accessed via Batch Data on the MGI homepage, or bookmarked directly at www.mousemine.org.

Answers to questions are on the final page.

Contact: mgi-help@jax.org
Joanne.berghout@jax.org

Worksheet Outline

Section 1: What does this gene do?
Section 2: Does a (knockout/conditional allele/reporter) mouse exist for this gene? How do I obtain it?
Section 3: Where is this gene expressed?
Section 4: How can I find SNPs between two inbred mouse strains within a gene or region?
Section 5: How can I find a list of alleles annotated to a disease or phenotype?
Section 6: How can I prioritize a gene list using mouse phenotype or disease associations?
Section 7: How can I determine if my gene set is enriched for GO (ontology:function/biological process/cell component) terms? Mammalian phenotype terms? Human disease terms?
Section 8: I read a paper with a mouse described in it, how can I find that mouse in MGI?

Section 1. What does this gene do?

Gene function in model organisms can be inferred from mutant phenotypes, described using the *Molecular Function*, *Cellular Component* and *Biological Process* attributes assigned by the Gene Ontology (GO) consortium, or predicted by sequence homology or structural domains.

MGI allows you to access information on each of these from the Gene Detail page.

1. What is one of the characteristic phenotypes reported for mice with mutations in van gogh like 2 (*Vangl2*) (use the Mutations, phenotypes and alleles row)?

a) acute myelogenous leukemia
b) embryonic lethality and developmental defects
c) liver dysfunction leading to jaundice
d) infertility

2. How many total mutations/alleles have been described for this gene?

a) 3
b) 6
c) 19
d) 22

3. Has a mouse mutant of *Vangl2* been used to model a human disease? If yes, which one?

a) No
b) Yes, Neural Tube Defects
c) Yes, Breast Cancer
d) Yes, Alzheimer Disease
e) Yes, Polydactyly

4. Using the Gene Ontology (GO) classifications row, the product of the *Vangl2* gene has been shown to:
(***hint:*** click the number next to All GO Classifications (#))

a) bind *Celsr1, Dact1, Dvl1, Dvl2, Dvl3, Gnb2l1, Magi3, Scrib* and *Sept2*
b) participate in the *biological process* "Wnt signaling pathway, planar cell polarity pathway"
c) participate in the *biological process* "cochlea morphogenesis"
d) all of the above

Section 2: Does a (knockout/conditional allele/reporter) mouse exist for this gene? How do I obtain it?

MGI indexes information regarding alleles described by the research community worldwide, whether generated (or characterized) by individual academic laboratories, large-scale public projects, commercial organizations or repositories. These may be curated from published literature, submitted directly by researchers or loaded as batch reports.

5. Locate the gene detail page for bone morphogenic protein 4 (*Bmp4*). How many total mutations/alleles have been described for this gene?

a) none
b) 2
c) 10
d) 24

6. Click on the hyperlinked number next to "All mutations alleles". From the resulting table, locate which of the following has the attribute "conditional-ready"

a) *Bmp4*^{tm1Blh}
b) *Bmp4*^{tm1Jfm}
c) *Bmp4*$^{tm2(tetO-Bmp4,lacZ)Jfm}$
d) *Bmp4*$^{tm1e(EUCOMM)Hmgu}$

7. Click on the allele symbol for *Bmp4*^{tm1Jfm}. According to the Mutation details (in the Mutation description row; click to expand), which exon(s) were flanked by LoxP sites in this mutant?

a) the whole gene
b) exon 1 and the transcription start site
c) exon 4
d) exons 7 and 8

8. Locate the "Find Mice (IMSR)" row on the Allele detail page. Is this strain available to purchase through the IMSR?

a) Yes, as ES cells from MMRRC
b) Yes, as live mice on a B6;129S4 background from JAX
c) Yes, as cryopreserved embryos on a B6 background from JAX
c) No, but potentially by contacting Dr. James F Martin directly

Section 3. Where is this gene expressed?

The Gene Expression Database portion of MGI catalogues expression of endogenous mouse genes throughout development using highly confident specific assay results (RT-PCR, immunohistochemistry, reporters, etc) and reported across structured anatomical terms.

9. Find the **Gene Expression Data Query** form on the Expression homepage, or by using the Search drop down menu. Enter *Nrg2* (for neuregulin 2) into the Genes search box. Use the Assay Results tab of the results table to determine which of the following structure & Theiler developmental stage combinations has *Nrg2* expression *detected* (***hint:*** pay attention to "detected" column).

a) TS12 (E8.0) rhombomere
b) TS15 (E9.5) endocardial lining
c) TS28 (P adult) uterus
d) TS28 (P adult) pancreas

10. Which assay type was used to make the annotation for Nrg2 expression in TS15 endocardial lining?

a) RNA in situ
b) Reporter knock in
c) RT-PCR
d) Immunohistochemistry

11. Which of the following structure & developmental stage combinations was specifically examined for *Nrg2* expression, but expression was not detected?

a) TS12 (E8.0) rhombomere
b) TS15 (E9.5) endocardial lining
c) TS28 (P adult) uterus
d) TS28 (P adult) pancreas

12. For additional details on the assay used to examine *Nrg2* expression in the TS28 uterus, which reference should be used?

a) J:42983 Carraway KL, et al., (1997) Nature
b) J:195554 Audouard E, et al., (2012) PLoS One
c) J:135516 Wansbury O, et al., (2008) J Invest Dermatol.
d) J:94401 Brown N, et al., (2004) Biol Reprod

Section 4: How can I find SNPs between two inbred mouse strains within a gene or region?

MGI has a SNP query form available, see homepage for Strains, SNPs & Polymorphisms, or use the Search drop down menu.

However, for this example, please open the Mouse Phenome Database (MPD; http://phenome.jax.org/)

Select the *Genotype* button, then *SNP / variation query* to proceed. In the window that appears, type "Hc" (case sensitive) for the gene hemolytic complement. Then, select "*Variation effect*: will specify next" from the menu on the right and use the radio buttons towards the bottom to choose "Standard with strain filtering".

Proceed to the next page to specify variation as "**Cn**: coding nonsynonymous".

Proceed to the next page to specify strains of interest. Choose A/J and C57BL/6J. De-select "All Sanger 1 strains" or extra data will appear.

13. On the results table, how many polymorphic variants appear between these strains?

a) 3
b) 10
c) 33
d) 408

14. Closely examine the row corresponding to rs27168885. Which of the following pieces of information appear?

a) The genome location of the variant, specifically, Chr2:35028174
b) The impact of the variant on protein sequence, specifically PL:699 (or, P699L)
c) The allele carried by C57BL/6J (G) and the allele carried by A/J (A)
d) All of the above

Section 5: How can I find a list of alleles annotated to a disease or phenotype?

The Phenotypes, Alleles and Disease Models Query provides a flexible search to generate a list of alleles:
(1) corresponding to a given gene
(2) annotated to a given phenotype or described as a model for a human disease
(3) located within a defined genomic region
(4) created using a generation method
(5) possessing specific allele attributes

15. Find the **Phenotypes, Alleles and Disease Models Query** form on the Phenotypes & Mutant Alleles homepage, or by using the Search drop down menu. Enter "prenatal lethality" (use quotes to match the complete phrase) in the Phenotype/disease section to return a list of alleles with at least one genotype exhibiting prenatal lethality, or a sub-term thereof. How many are there?

a) 134
b) 215
c) 4613
d) 11,201

16. Modify or re-run your search, adding the parameters *Chromosome* = [1] and *Allele attributes* = [Null/Knockout]. How many null/knockout alleles on Chr1 have a "prenatal lethality" phenotype?

a) 134
b) 215
c) 4613
d) 11,201

17. The *Acadl*^{tm1Uab} allele should appear on the list, with a Human disease model association to "Acyl-Coa Dehydrogenase, Very Long Chain Deficiency Of (ACADVLD)". Clicking on the disease name hyperlink brings up the Human Disease and Mouse Model Detail page. Which of the following statements is supported by this page:

a) *ACADL* gene mutations have been reported in human clinical cases of ACADVLD
b) *ACADVL* mutations have been reported in human clinical cases of ACADVLD; and *Acadvl* gene mutations have been used to model ACADVLD in mice.
c) mutating both *Acadvl* and *Acadl* are required to cause ACADVLD
d) all of the above

Section 6: How can I prioritize a gene list using mouse phenotype or disease associations?

This can be done similarly to the previous section, however for this example, please go to the Human-Mouse: Disease Connection (HMDC), accessible from MGI's home page, or direct at diseasemodel.org

Background: You have an exome sequencing result come back for a patient who has a familial susceptibility to **glomerulonephritis** (*renal/urinary system phenotype*), **vasculitis** (*cardiovascular system phenotype*) and **leukemia** (*tumorigenesis phenotype*), which appears to be inherited as a monogenic trait.

The genes that come back from your experiment with predicted pathogenic or damaging variants are: ACOX1, CHAT, SH2D3C, TUSC2 and ZYX.

18. Use the HMDC to search by gene. Based on mouse phenotypic annotations, which of these genes is the most likely candidate for your observations?

a) Acox1
b) Chat
c) Sh2d3c
d) Tusc2
e) Zyx

19. Which of these genes would be expected to be associated with **infertility** (a *reproductive system* phenotype)? ***Hint:*** click on the shaded cells at row-column intersections to see more specific phenotypes.

a) Acox1
b) Chat
c) Sh2d3c
d) Tusc2
e) Zyx

20. Which of these genes has variants associated with **Myasthenic Syndrome, Congenital, Associated with Episodic Apnea** in human patients?

a) ACOX1
b) CHAT
c) SH2D3C
d) TUSC2
e) ZYX

Section 7: How can I determine if my gene set is enriched for a variety of terms?

For enrichement analyses, as well as many other kinds of customizable queries, use MouseMine - MGI's batch access tool - found on the Batch Data & Analysis Tools link from MGI's homepage, or direct at: http://www.mousemine.org/

For this example, find a list of 321 genes that pre-loaded into MouseMine. Go to the "Lists" tab, "View" and click on "Mouse DNA repair genes 2013-01-06".

21. At the top of the page in list analysis tools, use Manage Columns (may compress to Columns in smaller windows) to ad a column for "NCBI Gene Number" to the list of 321 DNA repair genes. What is the NCBI Gene Number for *Hmga2*?

a) 15364
b) 22210
c) 13248
d) 35510

22. Scroll down to find the **Mammalian Phenotype Ontology Enrichment** widget. Using the default of Holm-Bonferroni test correction, what is the associated p-value for the MP term: "abnormal induced cell death"? This is calculated using a hypergeometric distribution. Click the "*i*" adjacent to "p-value" for details.

a) p-value: 1.329207e-67 with 51 genes matched
b) p-value: 1.091947e-62 with 63 genes matched
c) p-value: 4.872743e-27 with 23 genes matched
d) 0.05

23. Scroll further down the page to Template Results. Locate and click the header for the table "*Function: Mouse Features -> Functions (GO terms)*".
According to the Column Summary (bar graph icon) in column: "Code", how many functional annotations have the evidence code IMP which corresponds to "Inferred from Mutant Phenotype"?

a) 13
b) 9062
c) 1291
d) 3304

Note: Template queries do not need to be run from lists, explore the "Templates" tab, or see Popular Queries listed on the landing page.

Section 8: I read a paper with a mouse described in it, how can I find that mouse in MGI?

The **Reference Query** form can be found from the Search drop down menu (follow Search > References > Reference query)

24. Using the Reference query form, mice carrying which phenotypic allele were described in the publication: Chisaka O and Capecchi MR., "Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox-1.5" *Nature* 1991 350(6318):473-9

a) Hoxa3^{tm1Mrc}
b) Hoxa3^{tm1Kth}
c) Apoe^{tm1Unc}
d) Msx1
e) Hoxa

Answers

Section 1. What does this gene do?

1. b) embryonic lethality and developmental defects
2. d) 22
3. b) Yes, Neural Tube Defects
4. d) all of the above

Section 2: Does a (knockout/conditional allele/reporter) mouse exist for this gene? How do I obtain it?

5. d) 24
6. b) *Bmp4*^{tm1Jfm}
7. c) exon 4
8. b) Yes, as live mice on a B6;129S4 background from JAX

Section 3. Where is this gene expressed?

9. b) TS15 (E9.5) endocardial lining
10. a) RNA in situ
11. c) TS28 (P adult) uterus
12. d) J:94401 Brown N, et al., (2004) Biol Reprod

Section 4: How can I find SNPs between two inbred mouse strains within a gene or region?

13. a) 3
14. d) All of the above

Section 5: How can I find a list of alleles annotated to a disease or phenotype?

15. c) 4613
16. a) 134
17. b) *ACADVL* mutations have been reported in human clinical cases of ACADVLD; and *Acadvl* gene mutations have been used to model ACADVLD in mice.

Section 6: How can I prioritize a gene list using mouse phenotype or disease associations?

18. d) Tusc2
19. a) Acox1
20. b) CHAT

Section 7: How can I determine if my gene set is enriched for a variety of terms?

21. a) 15364
22. b) p-value: 1.091947e-62 with 63 genes matched
23. c) 1291

Section 8: I read a paper with a mouse described in it, how can I find that mouse in MGI?

24. a) Hoxa3^{tm1Mrc}