

## Exercises for GeneSNPs (Version 1, 2007)

**You can choose to do these using the questions as your only guide—or see the following pages for the step-by-step checklist to finding these answers.**

---

**1) Use the GeneSNPs site to identify genes involved in the cell cycle, which have been resequenced for SNP identification by NIEHS and for which a mouse model is available. From the last gene in the list that meets these criteria, identify how many nonsynonymous SNPs have been found.**

*Skills: Accessing gene lists; explore SNPs; filter for SNP type*

**2) Find a gene with a keyword search. Examine a pathway in which this protein participates. Find a downstream gene in this pathway. Examine that protein's structure and think about the implications of the SNPs it has.**

*Skills: keyword search, pathways investigation, PDBViewer*

**3) Find and export a list of tag SNPs for use in a genotyping experiment for DNA repair genes on chromosome 22. Use 'L' algorithm and a 5% cutoff (stringent parameter) to determine tag SNPs in the gene of interest.**



*Skills: search, access tag SNP data, export*

**GeneSNPs Exercises, version 1.0. Correspond to the data available in March 2007.**

**The materials and slides offered are for non-commercial use only. Reproduction, distribution and/or use for commercial purposes are strictly prohibited.  
Copyright 2007, OpenHelix, LLC.**

## Step-by-Step checklist/instructions GeneSNPs introductory exercises

1) Use the GeneSNPs site to identify genes involved in the cell cycle, which have been resequenced for SNP identification by NIEHS and for which a mouse model is available. From the last gene in the list that meets these criteria, identify how many nonsynonymous SNPs have been found.

Step	Action	✓
1	Go to the GeneSNPs homepage: <a href="http://www.genome.utah.edu/genesnps/">www.genome.utah.edu/genesnps/</a>	
2	On the homepage, <b>open the menu called Gene Lists.</b>	
3	From the Gene Lists menu, <b>select Cell Cycle.</b>	
4	On the gene lists results page, examine the table. <b>Identify genes that have been resequenced</b> by NIEHS (they have this star symbol:  ). In addition, <b>find genes for which a mouse model</b> from CMGCC is available (they have this mouse symbol:  ).	
5	From the results list, <b>select the last gene the meets all these criteria:</b> cell cycle, resequenced, mouse model. (At this time that gene is RB1). <b>Click the Gene Models link for RB1</b> ( <a href="http://UCSC:hg16:13">UCSC:hg16:13</a> ) to see the gene structure and SNPs.	
6	Examine the RB1 gene model, and the full list of SNPs in the default state. <b>Scroll down the bottom frame</b> to examine the complete list of SNPs. <i>Note the export options at the bottom of the frame.</i>	
7	Let's focus on a subset of the SNPs: <b>from the All SNPs menu below the diagram, select NONSYNON.</b>	
8	When the table has reloaded with just the nonsynonymous SNPs, find the first one from the EGP_SNPS handle group (currently at global coordinate 46749145). <b>Click the link in the red area that says "nonsynonymous"</b> to bring up a pop-up window with additional detail.	
9	From the SNP detail window, determine if this SNP falls into a protein domain. Yes, it appears to be PF01858. <b>Click the Domain link</b> to determine which domain this represents. A new window should open.	
10	From the new Domain page, <b>determine what PF01858 is.</b> ( <i>Retinoblastoma-associated protein A domain</i> ). <b>Click the PF01858 link</b> to visit Pfam and learn more about this domain. Return to the Domain diagram page. <b>Click the <u>A</u> link</b> to examine the alignment of this peptide to the domain sequence.	
11	<b>Close the domain window and return to the SNPcard for RB1.</b>	

2) Find a gene with a keyword search. Examine a pathway in which this protein participates. Find a downstream gene in this pathway. Examine that protein's structure and think about the implications of the SNPs it has.

Step	Action	✓
1	From the RB1 SNPcard, <b>click the GeneSNPs logo</b> to return to the homepage and start a new query.	
2	At the top of the homepage, <b>click the button that says Open Query</b> . Make these choices (from left to right): Function: <b>All</b> NIEHS: <b>All</b> Symbol: leave blank Keywords: <b>kinase</b> Chromosome: <b>2</b> <b>Click Search.</b>	
3	From the results list, <b>identify a kinase involved in cell signaling</b> . Use <b>the locus column link</b> to learn the name of this gene ( <i>eukaryotic translation initiation factor 2-alpha kinase 2</i> ). <b>Use the back button to return to the results list.</b>	
4	From the results list, <b>click the link for the gene model for EIF2AK2.</b>	
5	From the SNPcard page, <b>open the Menu near the top. Select Biocarta.</b>	
6	New data will load in the lower frame. This gene participates in several pathways. <b>Select the gsk3Pathway to view.</b>	
7	On the Biocarta pathway page, <b>hover over the EIF2AK2 link</b> on the left to illuminate the gene in the diagram ( <i>The label is a synonym, PKR</i> ). Look downstream of EIF2AK2 (PKR), notice NFkB bound to DNA. Let's go to that gene's SNPcard to learn more: <b>click the NFKB1 link on the left.</b>	
8	On the NFKB1 SNPcard page, <b>use the Menu to select PDBViewer</b> . New data will load in the bottom frame.	
9	Let's examine a structure with a high SDI (substitution/deletion/insertion). <b>Click the 1ikn link</b> in the first column of the PDBViewer table.	
10	<b>A new window will open.</b> A spinning structure molecule should be visible at the top. If not, you may need to load the CHIME software from MDL: <a href="http://mdl.com/products/framework/chime">http://mdl.com/products/framework/chime</a>	
11	<b>Use the menu options on the PDBViewer to change the look of the molecule.</b> The upper left control menus will let you examine a variety of styles and colors.	
12	<b>In the lower table, select different residues in the Chain-Residue menus.</b> This should change the look of the molecule to highlight these locations. Think about these alterations in the context of the structure and the interactions.	
<b><i>The PDBViewer will help you to think about the implications of the SNPs that you may find in genes of interest. Explore!</i></b>		

**3) Find and export a list of tag SNPs for use in a genotyping experiment for DNA repair genes on chromosome 22. Use 'L' algorithm and a 5% cutoff (stringent parameter) to determine tag SNPs in the gene of interest.**

Step	Action	✓
1	Return to the homepage by clicking the GeneSNPs icon, or go to <a href="http://www.genome.utah.edu/genesnps/">http://www.genome.utah.edu/genesnps/</a>	
2	<p>Click the "Open Query" button. In the menus make these choices (from left to right):</p> <ul style="list-style-type: none"> <li>• DNA Repair</li> <li>• NIEHS</li> <li>• leave symbol and keyword textboxes blank</li> <li>• choose 22 for Chrom.</li> </ul> <p>Click search to run the query.</p>	
3	One gene should result from search (DMC1). Click the Gene Model link for this gene (UCSC:hg16:22).	
4	On the SNPcard page, from the NIEHS menu at the top choose TAG SNPS.	
5	In the bottom frame the data will appear. One population is shown: "PDR90." Click the PDR90 link.	
6	You will be presented with several choices in a table. Type of method and algorithm used to select the tag SNPs is shown, and the stringency of the cutoff is offered at several levels. Choose the 5% cutoff for the LD Select method ( <i>this is the recommended method, it is based solely on disequilibrium</i> ).	
7	You will be given a list of 9 SNPs that can be used as tag SNPs in genotyping experiments. Export this list by clicking the Export link.	
8	A dialog box will appear to save the file. Save the file to your desktop. Open the file with a text browser to view the data.	
<p><b><i>Genotyping a few tag SNPs can determine variation in other locations because of linkage disequilibrium between regions.</i></b></p> <p><b><i>Your choice of cutoff and method will depend on needed coverage (number of variable positions determined by experiment), quality of data, and other needs of the experiment.</i></b></p>		