

# Lab 1

Name: \_\_\_\_\_

## USING INTERNET RESOURCES & BIOINFORMATICS FOR BEGINNERS

---

### OBJECTIVES:

- To learn how to find and use reliable scientific and biomedical resources on the web
  - To acquire the literature search skills necessary to write a scientific research paper
  - To learn basic 'bioinformatics'
- 

In this highly technological age, many occupations require specialized computer skills. Biology is no different! With the completion of the human genome project, computing power, information technology, and computer programming have become necessary tools for even the most traditional biologists. Today, we are going to identify and search the different kinds of online news sources for articles relating to hot topics in biology and medicine. Next, we will perform some simple bioinformatics analysis on genes involved in the *cytoskeleton* using internet resources available through the National Center for Biotechnology Information (NCBI). The skills that you will develop in this lab will be useful in finding articles for discussion section, doing literature searches for your research paper, and for exploring your own specific interests in biology or medicine.

### GENERAL

1. Work individually (if there are enough computers) or in pairs.
  - If you do not have a lot of experience with computers, sit next to someone who does.
  - If you are a computer whiz, try to be next to someone who is not.
2. Record notes in your laboratory notebook *as you go*, answering questions posed in this hand-out.
3. Some websites may require you to register as a member. In most cases, membership is free. If a website asks for a paid subscription, please let me know before proceeding.

### USING INTERNET SEARCH ENGINES

1. Go to Google ([www.google.com](http://www.google.com)) and enter the query "extremophile." From the results list, click on the first result (which should be a free encyclopedia). **Organisms from which domains of life can be classified as "extremophiles"? Name and define two different classes of extremophiles.**
2. Now return to the results page and scroll down to an article about Life on Mars. Click the link and scan the article. **According to this article, where are organisms most likely to reside on the Red Planet?**

### USING COMMON INTERNET NEWS SOURCES

3. Go to your favorite online, general news source (mine are: [www.nytimes.com](http://www.nytimes.com), [www.cnn.com](http://www.cnn.com), and [www.seattletimes.com](http://www.seattletimes.com)). Your news source should have a section entitled 'health,' 'medicine,' or ideally 'science.'
4. Go to the Health/Science Section and scan the headlines. **What headlines catch your eye?** Make notes in your lab notebook about what news source you are using, what section you are reading, and the headlines that interest you.

5. Now go to the Search box, enter the query “oil spill,” and wait for the results. Identify recent articles which discuss the environmental impact of the recent oil spill off of Alaska’s Aleutian Islands. If your news source has no related articles, try one of the above suggested sources. Take some time to read your chosen article and record in your notebook the source, article title and date. After reading your article, answer the following questions in your notebook.

- 1) **What organisms were most affected by the recent oil spill in Alaska?**
- 2) **How might this spill impact the local human population?**
- 3) **How might this spill affect the local ecosystem?**

## USING SCIENCE NEWS SOURCES

1. Visit the website of one of the following science news sources:
  - *Scientific American* at [www.sciam.com](http://www.sciam.com)
  - *Discover* magazine at [www.discover.com](http://www.discover.com)
  - *National Geographic* at [www.nationalgeographic.com](http://www.nationalgeographic.com)
  - *New Scientist* at [www.newscientist.com](http://www.newscientist.com)
2. Scan the headlines on the home page. **What headlines catch your eye?** Record the source and date for these headlines in your notebook.
3. Go to the search box and enter the query “mad cow”. Identify articles that discuss mad cow disease and other related diseases (called ‘spongiform encephalopathies’). Read one or more articles and answer the questions below in your notebook.

- 1) **What is the name of the infectious particle that causes mad cow disease?**
- 2) **Can these types of infectious diseases be treated with antibiotics?**
- 3) **How do these diseases spread?**

## USING THE SCIENTIFIC LITERATURE

1. Visit the website of one of the following top science/medical journals:
  - *Science* at [www.sciencemag.org](http://www.sciencemag.org)
  - *Nature* at [www.nature.com](http://www.nature.com)
  - *Cell* at [www.cell.com](http://www.cell.com)
  - *Public Library of Science* at [www.plos.org](http://www.plos.org) (free access!)
  - *New England Journal of Medicine* at [content.nejm.org](http://content.nejm.org)
2. Go to the current issue and scan the table of contents. **What are the names of the sections found in this journal?**
3. Return to the journal home page. Find the search box and enter the query “tuberculosis.” **How many hits came back from this search?**
4. Now go to the search box again and enter the query “stem cells.” **How many hits came back from this search?** Record these results in your notebook.
5. **Which of these topics do you think receives more federal funding, tuberculosis research or stem cell research?**

## INDEPENDENT WEB EXPLORATION TIME

Now that you are aware of the resources that are available to you on the internet, take some time to peruse these websites to find articles related to your discussion topic and make note of them. You may wish to find your discussion team partners and discuss your ideas together. As you are doing these searches, pay attention to the headlines and look for topics that interest you. You can use this opportunity to identify a subject for your research paper (topics are due one week from today!).

## BASIC BIOINFORMATICS

1. Visit the website of the National Center for Biotechnology Information:  
<http://www.ncbi.nlm.nih.gov/>
2. In the search box, type in the word “actin” and hit enter. Remember that *actin filaments* are one of the main components of the *cytoskeleton*, which is involved in cell structure, shape, and movement.
3. Find “PubMed” (located in the top left of the results page) and note the number of results for “actin” next to it. Click on PubMed. PubMed is a commonly used scientific literature database. Researchers use this database to search for publications related to their work. Scan through the titles of the first twenty articles. **What scientific terms that you recognize from our first class?**
4. Now go back to your “actin” results page and find the term “OMIM” and click on it. OMIM stands for Online Mendelian Inheritance in Man—it is a database of all human genes and the diseases with which they are associated (and more). Click on the top result “Actin, Gamma-1,” read the description, and record it in your notebook. Now scan through the rest of the page and try to find a disease associated with mutations in actin, gamma-1. **What is (or are) the disease(s)? How many mutations are associated with this disease?**
5. Now return to your “actin” results page and find the term “Protein” and click on it. You can clearly see by clicking through these pages how many species have actin proteins and also how many actin related proteins there are.
6. Now go back to your “actin” results page. In the menu bar at the top of the page, click on “BLAST” at the far right. “Blast” allows you to take a given sequence and compare it to other sequences, from whole genomes to a single gene sequence. Under the “Special” heading, go to “Align two sequences.” Beside “Program,” the default option reads “blastn,” which means blast nucleotide (DNA). Click on it and change it to “blastp” (for blast protein). This means we are going to query the PROTEINS, rather than genes. Beside “Sequence 1” enter the accession # NP\_001092 (the entry code for the human actin sequence) and beside “Sequence 2” enter the accession # NP\_031419 (the entry code for the mouse actin sequence). Then scroll to the bottom and click “align”. This blast alignment will show the degree of similarity between the human and mouse actin proteins by a direct comparison. **What is the length of the aligned sequence? How similar are the two sequences?**

7. Now go back to the BLAST home page. Under “Genomes,” click on “Fungi” (listed in the 5<sup>th</sup> bullet). In the search box, enter the accession # for the human actin protein (NP\_001092). Now go to “query” just under the search box and change the default setting, “DNA”, to “Protein.” Scroll down through the list of fungi genomes and click the box next to *Saccharomyces cerevisiae*. This is the scientific name of a fungus familiar to us all, baker’s yeast, used in bread-making and beer-brewing. It is a single-celled eukaryote also used by many biologists to study basic cellular processes. Now scroll back to the top and click the “Blast” button. When the next page appears, click on “Format” and wait for the results to be displayed. By doing this, you are searching the entire yeast genome for a sequence that is similar to that of human actin. **How many significant alignments (matches) are there in the yeast genome to human actin? Now scroll down and look at the first alignment (on yeast chromosome VI)—how similar are the human and yeast actin proteins?** Now think about the description that you read in OMIM for actin and recorded in your notebook. **How important is the role of actin in basic cellular functions based on what you have read and the alignments you have now seen?**
8. Repeat steps 1-4 with the query “lamin.” Lamins are *intermediate filament* proteins (another component of the cytoskeleton) that exist exclusively in the *nucleus*.
9. Now go back to the “lamin” results page and go to “Structure”. There should be four results. Click on the third one down “C-Terminal Globular Domain Of Human Lamin A/C.” From here, you can view the 3-dimensional structure of a domain of this lamin protein (you must first download the Cn3D 4.1 software). Do so and then click on “View 3D Structure.” Once you open the file, you should be able to move the protein around and change the protein’s appearance by going to the “Style” menu and choosing the different options in “Rendering Shortcuts.” **Based on this exercise, what conclusions can you make about the structure of proteins?**