Navigating the NCBI Instructions CLASS COPY

**Aim:** To become familiar with the resources available at the National Center for Bioinformatics (NCBI) and its internal search engine, *Entrez*.

**Instructions:** Write the answers to your questions on the Student Worksheet.

1. Go to the NCBI homepage <https://www.ncbi.nlm.nih.gov/>

2. Take a few minutes to look around the site. The goal is to familiarize yourself with a few key components of the NCBI.

a. What is the name of one interesting resource or database shown in the blue menu box on the left? What do you think is its function or purpose?

b. What is one interesting resource listed in the Popular Resources menu on the right? What do you think is its function or purpose?

3. Find the search box at the top of the webpage. This search box uses the NCBI search engine *Entrez* to look for your search term (or “**query**”) across ***all*** of the databases at the NCBI.

4. Type “BRCA1” into the **Search** box. Make sure there is *no* space between BRCA and 1. Click **Search**.

*BRCA1* is a tumor suppressor gene that normally prevents cancer. Mutations in this gene are associated with increased risk of hereditary breast cancer and ovarian cancer when normal function is lost.

The number to the right of each database contains the number of “hits” returned from that database (see screen shot, above). This is like searching in iTunes® without specifying categories like ringtones, podcasts, movies, TV, or songs.

a. Why are we searching for *BRCA1*?

b. The **Nucleotide** database (under Genomes) has DNA sequences that have been loaded onto the NCBI database. How many times is ‘BRCA1’ cited in the Nucleotide database?

c. The **PubMed** database (under Literature) has the articles that have been published about a specific gene or disease. How many times is ‘BRCA1’ cited in the PubMed database?

d. Compare the numbers you got for Questions **a** and **c**. Do these relative numbers surprise you? What does this tell you about the *BRCA1* gene? Explain.

5. Go back to the NCBI homepage by clicking the **NIH** logo in the upper left corner of the screen.

This search shows that there is a lot of information at the NCBI! It can be challenging to try to make sense of it all. Let’s start with something more familiar

6. Click the “**All Resources**” link from the list of resources on the left side of the screen.

7. Find “**Genome Data Viewer (GDV)**.” By clicking on the “Tools” tab and scroll through the alphabetical list. Then Click on the “**GDV**” link.

The resulting page is called **GDV** and it allows us to search the genomes of many different organisms, including humans.

8. Select ***Homo sapiens (Humans)*** from “Select Organism” menu (it should be the default).

To the left is a green labeled box containing access to information to the Human genome. Click on “Browse Genome”

9. Now we can see the ***Homo sapiens* (human) genome view. A genome** is all of the genetic information in an organism. Each figure you see in the “genome view” represents a pair of chromosomes. Most of the chromosomes are numbered, but a few are not. The abbreviations “X” and “Y” refer to the human sex chromosomes.

a. How many different types of chromosomes do you see?

b. “MT” represents the Mitochondrial genome. How big is it? (click on it to find out)

c. With the exception of MT, the chromosomes of the human genome are in pairs. X and Y are a pair. Using this information and the information from your answer to Question 9A, how many **pairs** of chromosomes are in the human genome?

10. The Breast Cancer Susceptibility gene *BRCA1* is on chromosome 17 in humans. [Click on the link below chromosome 17.] Explore some of the links and views.

What do you see when you click on chromosome 17? Explore some of the links on the picture, and write down two things you found interesting, such as the description of other genes that are also found on chromosome 17.

11. To find the location of the *BRCA1* gene, type “BRCA1” in the “**Search**” box at the left of the screen.

Draw a picture of chromosome 17 and show the *approximate* location of *BRCA1* on this chromosome.

12. Go back to the NCBI Homepage and again enter **BRCA1** into the search box. Click on the “Gene” database. Notice how many different organisms have exact copies of this gene. Then click on the top BRCA1 gene for humans.

List three of the functions that the BRCA1 protein performs.

13. To learn about all of the **phenotypes** associated with mutations of *BRCA1*, return to the top of the web page and from the “**Table of Contents**” on the right, select “**Phenotypes**.” This will bring you to the portion of the web page that contains the phenotype information for *BRCA1*.

a. Based on what you’ve learned in class, what is a **phenotype**?

b. What\_ **phenotypes** are associated with mutations in the *BRCA1* gene? (read the “Associated conditions” section)

14. Return to the **Table of Contents** at the top of the page and click “**NCBI** **Reference Sequences**.” This will take you to the portion of the webpage that contains the actual genetic sequence of the *BRCA1* gene.

**15 Reference sequences** are DNA or protein sequences that scientists, doctors and genetic counselors use to study genes like *BRCA1*. You can download these sequences in different formats. For this exercise click “**FASTA**” in the “Range Download” box

16. This link takes you to the FASTA sequence for *BRCA1*. Scroll through the web page. This gene is *very* large!

a. What four letters make up this long sequence?

b. Based on what you’ve learned in class, what do these letters represent?

17. Return to the NCBI homepage by clicking on the NCBI icon on the top left of the web page.

18 Type **BRCA1** in the **Search** box and select “**Nucleotide**” from the “Genomes”box, to limit your search to the database containing all of the DNA and RNA (Nucleotide) sequences.

19. *What other organisms have* BRCA1 *genes?* You can scroll through the list of organisms, but note that these are listed by the **scientific name** of the organism (Genus and species), and the common name. For example, *Homo sapiens* is the scientific name for humans. Also, the **Top Organisms** (or the organisms with the most “hits”) are listed from top to bottom.

List three organisms other than humans that have *BRCA1* genes. Include both the scientific and common names.

20. Look back at your list of functions for the *BRCA1* gene (question #12).

Does it surprise you that so many organisms share the *BRCA1* gene? Explain.

21. What kind of information can you find at the National Center for Biotechnology Information?

Summarize what you have learned today by listing three types of information found at the NCBI