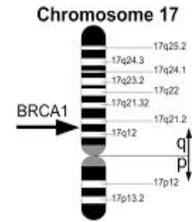


## Hands-On Ten The BRCA1 Gene and Protein



**Objective:** To review transcription, translation, reading frames, mutations, and reading files from GenBank, and to review some of the bioinformatics tools, such as BLAST for pairwise alignment and CLUSTAL OMEGA for multiple sequence alignment.

**Assignment:** The BRCA1 Gene

**A)** We would like to retrieve the record with accession number U14680 from NCBI.

- Go to the NCBI web site: [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov).
- Select “Nucleotide” from the “All Databases” drop-down menu and type “U14680” in the search window.
- Click on the blue “Search”

1. Determine the organism and the name of the gene from which the sequence comes. What is its locus?

Organism: \_\_\_\_\_ Gene: \_\_\_\_\_

Locus: \_\_\_\_\_.

2. What type of sequence (DNA, RNA, protein) is described in this NCBI record?  
\_\_\_\_\_.

3. How many exons does the gene have? \_\_\_\_\_. Are they all coding exons? \_\_\_\_\_.

Explain: \_\_\_\_\_.

4. Why does the sequence under /translation start with following specific amino acids: MDLS?  
\_\_\_\_\_.

5. Scroll up to the top of the page and click on “FASTA” under “GenBank: U14680.1”.

- Copy and save the sequence in FASTA format in a file.

6. Change the display back to the default, GenBank format by clicking on “GenBank” at the top of the page.

Based on the annotation, what is the accession number (protein\_id) of the resulting **protein sequence**? \_\_\_\_\_.

7. Click on the accession number (protein\_id) of the protein. How many amino acid residues are there in the protein? \_\_\_\_\_.

8. Scroll up to the top of the page and click on “FASTA” under “GenBank: AAA73985.1”.

- Copy and save the sequence in FASTA format in the same file you created in number 5.

9. We will perform a pairwise alignment between translations of the mRNA sequence (from part 5) and the protein sequence (from part 8).

- Open a second browser window, go to the NCBI home page, and select the BLAST tool by clicking on “BLAST” under “Popular Resources” at the right-hand side of the page.
- From the main page of BLAST, choose “Align two (or more) sequences using BLAST (bl2seq)” under “Specialized BLAST”.
- From the new page, select the translated BLAST (blastx) from the top of the page
- Scroll down towards the bottom of the page and click on “Algorithm parameters”
- Go to “Filter” and deselect “Low complexity regions”
- Paste the sequence from part 5) in the first window and the amino acid sequence from part 8) in the second window. Alternatively, we can type “U14680” and “AAA73985.1” in the first and second windows, respectively. Note that the order of entering the sequences is important.
- Choose “Show results in a new window” next to “BLAST”.
- Click on the blue “BLAST”.

The following three questions are about the very first alignment (and most significant one) displayed under “Alignments”.

- a) How similar are the two sequences? \_\_\_\_\_.
- b) Did you expect the degree of similarity you found? \_\_\_\_\_.
- Why? \_\_\_\_\_.

10. Pairwise alignment methods can also be used to detect mutations. We are going to align the wild-type BRCA1 mRNA sequence (with accession number U14680) with a mutant version of the mRNA (with accession number U64805).

- Go back the BLAST web page
- Choose “Align two (or more) sequences using BLAST (bl2seq)” under “Specialized BLAST”.
- Make sure you have “blastn” at the top of the page.
- To make the alignment, simply enter the accession numbers: U14680 in the first window and U64805 in the second window. You do not have to retrieve and paste the two sequences in the windows.

- Choose “Somewhat similar sequences (blastn)” under “Program Selection”
- Choose “Show results in a new window” next to “BLAST”.
- Click on “BLAST”.

Examine the top two alignments in the result web page, and answer the following question.

Given the results of your alignment, what type of mutation do you think has occurred to produce the mutant mRNA? (i.e. insertion, deletion, point mutation, etc). Be specific. Retrieve the mutant version of the BRCA1 mRNA and check your answer.

Hint: Click on “Dot Matrix View” and carefully examining the graph produced by BLAST.

**B)** We are going to use the multiple sequence alignment tool CLUSTAL OMEGA available at EBI to find regions of similarity between the human BRCA1 wild-type and seven other BRCA1 proteins from other organisms.

- Retrieve the seven sequences: “sequences\_BRCA1.txt”.  
Note that the first sequence is the human BRCA1 wild type.
- Go to <http://www.ebi.ac.uk/Tools/msa/clustalo>
- Paste all seven sequences in the CLUSTAL OMEGA data window.
- Under STEP 2, click on “More options ...” and choose “input” under “ORDER”.
- Click on “Submit” to align the seven sequences and wait for the results.
- Scroll down and click on “Show Colors”.
- Carefully examine the alignment. Note that:
  - “\*” at the bottom of a column means that we have an alignment where all the amino acids are identical.
  - “:” at the bottom of a column represents strongly conserved alignment.
  - “.” at the bottom of a column represents a weakly conserved alignment.
- Note that by clicking on “Page Summary” at the top of the page, and by choosing the JAVA alignment viewer and editor “JalView”, we get to see a different way of showing the alignment of the eight sequences.

Answer the following questions based on the results of the alignment:

11. Which two sequences are the “least” similar?

12. Which sequence is the most similar to the human BRCA1?

C) We are going to look for more information about the human breast cancer susceptibility gene, BRCA1.

- Go to the main page of NCBI: <http://www.ncbi.nlm.nih.gov> and click on “Genomes & Maps” at the left-hand side of the page.
- From the new page, click on “Map Viewer” under “Quick Links” at the right-hand side.
- From the “Map Viewer Home” page, choose the “*Homo sapiens*” row and click on “Annotation Release 107”, under “Build”, to get to the “*Homo sapiens (human) genome view*” page. The Map Viewer provides a wide variety of genome mapping and sequencing data.
- On this page, type in “BRCA1” in the “Search for” window.
- Click on “Find” and carefully examine the new page and answer the following questions.

13. a) How many hits are reported by “Search results for query BRCA1”? \_\_\_\_ .  
 b) On which chromosome do we find the majority of the hits? \_\_\_\_ .  
 c) What chromosome contains the actual BRCA1 locus? \_\_\_\_ .

14. To see the names of some of the genes whose encoded proteins interact with the BRCA1 protein:

- Click on “Advanced Search” at the top of the page on the right hand side. This will take us to the page “*Homo sapiens (human) advanced search*”.
- In “Type of mapped object:” deselect all options by clicking on “clear” and then select “Gene”. Click on “Find” at the top of the page.

How many hits are reported by “Search results for query "BRCA1 AND (Gene[obj\_type])"? \_\_\_\_ .

Note that that we have information wrt “reference” “CHM1\_1.1” and “HuRef” under “Assembly”. Consider the information reported by “reference” only.

15. Fill in the following table for the 6 genes that produce proteins that interact with BRCA1. The first row in the table is already entered.

Chromosome	Match	Map element	Type
2	OTTHUMP00000163973	BARD1	GENE

Consider the information under “17 reference”.

- Go to the row that starts with “BRCA1: protein phosphatase 1, regulatory subunit 53”.
- Click on “Genes seq” under “Maps”.

Examine the map view. It illustrates a portion of the chromosome with some of its genes. Study the information at the top and bottom of the page and answer the following questions:

16. What is the total number of genes on this chromosome? \_\_\_\_\_ .

17. What is the locus (chromosome location) of the BRCA1 gene? (e.g., 18p14.25) \_\_\_\_\_ .

18. What region of the chromosome is illustrated (from nucleotide to nucleotide)? \_\_\_\_\_ .

19. Write down the names of the other genes that are identified in the illustrated region of the chromosome.

20. Are these genes encoded on the same or on different strands of the chromosome? \_\_\_\_\_. On what strand is BRCA1? \_\_\_\_\_.

Hint: See the orientation column (with header "O").

21. What is the total length in bp of the entire chromosome? \_\_\_\_\_

Hint: You can zoom out completely (left side of the page) or enter "1" in the upper box of "Region Shown" and clear the lower box and leave it empty and then hit "Go".

Go to the line that has: "BRCA1  OMIM HGNC sv pr dl ev mm hm sts SNP" and click on "BRCA1". You will now find yourself in "Entrez Gene" which provides a variety of information on BRCA1. Check out the "Updated Date" at the top left-hand corner. What is it? \_\_\_\_\_ .

By carefully reading "Summary", we learn that BRCA1 encodes a nuclear phosphoprotein that plays a role in maintaining genomic stability. The following is a list of the normal functional roles of BRCA1:

- It is a tumor suppressor in human breast cancer cells
- BRCA1 may function as a transcriptional regulator
- It plays a role in DNA repair of double-stranded breaks
- It also plays some role in recombination.

22. What alterations may lead to breast cancer?

In higher eukaryotes there are often variations in gene transcripts due to alternative splicing of mRNAs. BRCA1 has a number of such splice variants. Scroll down to: "NCBI Reference Sequences (RefSeq)". This section describes the gene-specific NCBI reference sequences (RefSeqs) that have been established for this gene. In addition to enumerating the accessions and providing links to the appropriate Entrez sequence database, this section

includes descriptions of each transcript variant, accessions of the public sequences used to support any transcript, and a listing of computed domains in an encoded protein.

23. What is the phenotypic effect on the BRCA1 protein of the following transcript variants? In other words, what is the resulting BRCA1 protein obtained from the following transcript variants?

- i) mRNA Sequence NM\_007294.3
- ii) mRNA Sequence NM\_007297.3
- iii) mRNA Sequence NM\_007298.3

- Scroll up of the “Entrez Gene” page and stop at “Genomic regions, transcripts, and products”.

24. What nucleotide sequence region of the chromosome is shown (from which nucleotide to which nucleotide)? \_\_\_\_\_.

25. What does the orientation of the numbers tell us about the strand of DNA the mRNA is derived from? \_\_\_\_\_.

**D)** We are going to explore the Cancer Genome Anatomy Project (CGAP), which is maintained by the National Cancer Institute.

- Go to the CGAP web site: <http://cgap.nci.nih.gov>

CGAP pulls together a large amount of public data on genes altered in cancer cells. This ranges from gene sequence, mutational analysis, cytogenetic localization, chromosomal aberration, biochemistry, and clinical manifestation for cancers. By making all of this information easily accessible, CGAP helps the researcher analyze new information and relate it to what is already known. Since cancer ultimately is a problem of expression of a mutated gene product(s) and consequent inappropriate expression of other genes to give the cancer phenotype, analysis of gene expression has been one of the major efforts of CGAP.

- Click on the “Cancer Genome Anatomy Project” link in blue, to go to CGAP’s homepage (<http://cgap.nci.nih.gov/cgap.html>).

As we unravel the biochemical function of various genes and gene products, we can place them in functional pathways using all of the information available to us. As mentioned in the CGAP homepage, under “Pathways”, the latter are diagrams of biological pathways and proteins, with links to genetic resources to known proteins.

- Click on “Pathways” to obtain a new page.

- Click on the “BioCarta Pathways on CGAP” to obtain a new page.

BioCarta has an extensive list of biochemical pathway charts.

- Under A, click on “ATM Signaling Pathway” to go to a new page.

Notice the very first line at the top of the map.

28. What are the phases G1, S, G2, and M part of?  
\_\_\_\_\_ .

The ATM pathway is important for delaying cell cycle progression to allow time for repair if there has been damage to DNA. This helps to prevent permanent damage or chromosome breakage and the consequent genomic instability. Maps of this type help the researcher place a particular gene product into a broader context. The BRCA1 gene product is shown just to the left of the center. Its precise role, since mutations in BRCA1 lead to a high risk of breast cancer, is the focus of much research around the world. Clicking on any of the icons in the map takes us to the Gene Info page for that gene.

29. The diagram contains many products that interact with one another. Do you recognize some of the proteins that interact with BRCA1? \_\_\_\_\_ .  
If yes, which ones? \_\_\_\_\_ .

Notice that the gene products have different shapes.

- Click on “legend” right above the display panel to get an explanation of the various shapes.

30. What is BRCA1 according to the “legend”? \_\_\_\_\_ .

- Go back to the ATM Signaling Pathway.

31. Explain how one can go to the “ATM Signaling Pathway” of the mouse.

32. Go to the “ATM Signaling Pathway” of the mouse. Is it very different than the human “ATM Signaling Pathway”? \_\_\_\_\_ .

**The Information provided in this hands-on exercise is for instructional purposes only.**

The Information provided in this hands-on exercise is for instructional purposes only.

### More about the BRCA1 Gene

#### A) Some basic information about the BRCA1 gene

- 1) What is the official name of the BRCA1 gene?
- 2) What is the normal function of the BRCA1 gene?
- 3) What conditions are related to the BRCA1 gene?
- 4) Where is the BRCA1 gene located?
- 5) Where can I find additional information about BRCA1?
- 6) What other names do people use for the BRCA1 gene or gene products?
- 7) What glossary definitions help with understanding BRCA1?

Answers to the above questions can be found at the Genetics Home Reference at: <http://ghr.nlm.nih.gov/gene=brca1>

#### B) What is the role of heredity in susceptibility to cancer?

- 1) Given that individuals normally inherit a predisposition to cancer, not the cancer itself, how is the predisposition transmitted?
- 2) If women with BRCA1 and BRCA2 mutants have an increased risk of breast cancer, is the risk associated with a specific protein product of these genes or the lack of a functional product?
- 3) Are women with BRCA1 and BRCA2 mutations at higher-than-average risk for cancers other than breast cancer?

The article found at <http://www.breastdiseases.com/genebr3.htm> on the characteristics of inherited forms of breast cancer will help in answering the above questions.

#### C) What is the role of heredity in susceptibility to cancer?

A wealth of information can be found at:

<http://www.cancerindex.org/geneweb/BRCA1.htm>

- 1) Gene Database Entries for BRCA1
- 2) Other BRCA1 Related Resources
- 3) Hereditary Breast and Ovarian Cancer Syndromes
  - i) BRCA1 Screening
  - ii) BRCA1 and Genetic Counseling
- 4) BRCA1 Functions and Carcinogenesis
- 5) BRCA1 and Breast Cancer
  - Prophylactic Treatments for Women with BRCA1 mutations
- 6) BRCA1 and Ovarian Cancer
  - TP53 Mutations in BRCA1/BRCA2 Ovarian Cancers
- 7) BRCA1 - Somatic mutations
- 8) BRCA1 Mutations
  - i) 185delAG Mutations
  - ii) 5382InsC Mutations
- 9) BRCA1 and Colon Cancer Risk
- 10) BRCA1 and Fallopian Tube