Using the NCBI Genome Databases to Compare the Genes for Human & Chimpanzee Beta Hemoglobin

Author(s): Susan Offner
Published By: National Association of Biology Teachers
URL: http://www.bioone.org/doi/full/10.1525/abt.2010.72.4.10

BioOne (www.bioone.org) is a nonprofit, online aggregation of core research in the biological, ecological, and environmental sciences. BioOne provides a sustainable online platform for over 170 journals and books published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Web site, and all posted and associated content indicates your acceptance of BioOne’s Terms of Use, available at www.bioone.org/page/terms_of_use.

Usage of BioOne content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.
Using the NCBI Genome Databases to Compare the Genes for Human & Chimpanzee Beta Hemoglobin

ABSTRACT

The beta hemoglobin protein is identical in humans and chimpanzees. In this tutorial, students see that even though the proteins are identical, the genes that code for them are not. There are many more differences in the introns than in the exons, which indicates that coding regions of DNA are more highly conserved than non-coding regions.

Key Words: Evolution; beta hemoglobin genes; beta hemoglobin protein; humans; chimpanzees

This paper presents a tutorial in which students compare the gene for human beta hemoglobin with the gene for beta hemoglobin in the chimpanzee. This comparison is interesting because the two proteins are identical in amino acid sequence but, as students will find out, the genes that code for these proteins are similar but not identical. This tutorial introduces students to the NCBI (National Center for Biotechnology Information) genome databases and gives them some practice in using the powerful tools associated with them. This is where DNA sequences from labs all over the world are deposited. This enormous database is searchable and available at no cost to anybody with Internet access.

After students complete the tutorial, they will find that, although the beta hemoglobin proteins of humans and chimpanzees are identical, the genes that code for these proteins differ by one base pair in the exons (which code for proteins) and by 15 base pairs in the introns (which do not code for proteins). They can then understand that because the genetic code is degenerate (more than one codon can code for the same amino acid), two different genes can code for the same protein. They can also appreciate that introns evolve more rapidly than exons. This is because, although it is assumed that mutations occur in all DNA at the same rate, a mutation in an exon that changes the order of amino acids in the protein will probably be selected against, whereas most mutations in introns will likely have no selection pressure against them and so are more likely to remain in the gene pool. Thus, even though there are fewer than twice as many base pairs of introns (1068) than of exons (538), there are 15 times more mutations in the introns than in the exons.

It is suggested that teachers complete this activity themselves before giving it to students. Finally, there are other Web sites, including the Bioserver tool at the Dolan DNA Learning Center (dnalc.org), that have useful tools for genome analysis.

○ BLAST Tutorial: Beta Hemoglobin

DNA sequencing has become very easy and rapid since the 1990s. By now, over 100 billion base pairs of DNA have been sequenced, and more sequence is being added at an astonishing rate. All this DNA sequence is deposited in an enormous database maintained by NCBI. This database is free and available to the public. It can be easily searched using a free program called BLAST (Basic Local Alignment Search Tool; Altschul et al., 1990). In this tutorial, you will learn how to use these databases and BLAST to compare human and chimpanzee hemoglobin.

Hemoglobin is a protein found in red blood cells that carries oxygen. One hemoglobin molecule consists of four amino acid chains: two alpha hemoglobin chains and two beta hemoglobin chains. The alpha and beta chains are coded for by separate genes; the alpha hemoglobin gene is found on chromosome 16 and the beta hemoglobin gene is found on chromosome 11. Each of these amino acid chains contains a heme group that can carry one oxygen molecule. This means that one molecule of hemoglobin can carry four molecules of oxygen. One beta hemoglobin chain contains 147 amino acids.

Humans and chimpanzees last had a common ancestor between 5 and 8 million years ago. Their beta hemoglobin proteins are identical in amino acid sequence and, of course, in secondary, tertiary, and quaternary structures. In this activity, you will explore whether the gene that codes for the beta hemoglobin protein is identical in humans and chimpanzees.

1. Go to ncbi.nlm.nih.gov. This is the NCBI home page through which you can access both the DNA databases and BLAST.
2. At the top of the home page, search All Databases (default) for HBB. HBB is the abbreviation for beta hemoglobin. Click “Go.”
3. In the middle of the page on the left, click on “Nucleotide: core subset of nucleotide sequence records.”
4. Near the top of the page, click on “HBB (Homo sapiens).” You may have to right-click with your mouse to open this link. A new screen...
will come up that gives you the nucleotide sequence of the gene that codes for the human beta-hemoglobin protein. There is a lot of additional information on this, and you will see links to many other pages. You can explore these after you have finished the basic tutorial.

5. Scroll down a little until you see, in a gray box on the left, “Genomic regions, transcripts and products.” Click on “Reference sequence details.”

6. On the left side of the screen, find “Genome Reference Consortium Human Build” in a gray box. Under this, find “Genomic,” then find “Download.” To the right of “Download,” click on “FASTA.” You will see that the beta hemoglobin gene contains 1606 nucleotides (“letters”), here shown in what we call FASTA format. FASTA is a standard format, used for both genes and proteins, in which the first line begins with a > symbol, followed
Figure 2. The BLAST printout page, showing the comparison between human beta hemoglobin (Query) and chimpanzee beta hemoglobin (Subject). This figure was printed from NCBI’s BLAST site.
Copy the entire DNA sequence. Do not copy the first line.

7. Go back to the NCBI home page (ncbi.nlm.nih.gov) by clicking on the NCBI logo at the top of the page.

8. Under “BLAST Assembled Genomes” near the top of the page, click on “Pan troglodytes.” This will allow you to compare the human beta-hemoglobin gene with the chimpanzee genome, and not with the entire NCBI database. This saves time, because beta hemoglobins have been sequenced for innumerable organisms. Other BLAST searches could allow you to compare your sequence with the entire NCBI database or with different subsets of it.

9. At the top of the page, click on “BLAST.”

10. Paste the human beta-hemoglobin sequence that you have copied into the large box at the top of the page. This is your query. The directions for the box say “enter an accession….”

11. Scroll to the bottom of the page and click on “Begin Search.”

12. In the middle of the page, to the right of “Request ID,” click on “View Report” in a gray box.

13. Scroll down to “Alignments.” The “Query” is human beta-hemoglobin. The “Subject” is chimpanzee beta hemoglobin. Notice that whenever the gene is identical, there is a vertical line between the identical base pairs.

**Figure 2. (Cont’d)**
16. Now go to the top of the page on the left and click on “Download.”
17. Near the top of the page under “Alignment,” click on “text.”
18. Click on “save to disk.”
19. Open the file. This will give you a comparison of the two beta hemoglobins. You can print this out if a printout is not included in the packet your teacher has given you (it is presented here as Figure 2). When this is on your computer screen, you will be able to search for the beginnings and ends of the exons in later steps.
20. Now look at the copy of the human-beta hemoglobin gene included in the packet your teacher has given you (Figure 1). This copy of the gene is from 1980 and uses the three-letter abbreviations for the 20 amino acids. There are three exons and two introns in this gene. You can identify the three exons because the amino acids they code for are given above the base sequences. Look at the sequences at the beginning and end of each exon. Use a highlighter to highlight the nucleotide sequences of each of the three exons in the human beta-hemoglobin gene on this copy of the human beta-hemoglobin gene.
21. Take out your printout of the human–chimpanzee beta-hemoglobin comparison from the BLAST search (Figure 2).
22. Use the find options of the computer (Edit/Find) to identify the beginning and end of each exon of the human beta-hemoglobin gene. Highlight the three exons on your printout of the human–chimpanzee comparison. It is easiest to highlight the exons in the human gene. Hint: search for six bases at the beginning and end of each exon, which you have located in Figure 1. Six bases are enough to identify these positions.

23. You can now answer the following questions (the answers are given in italics):
   a. Are the exons identical? No. Circle or highlight any differences you find. How many differences are there? One.
   b. Are the introns identical? No. Circle or highlight any differences you find. How many differences are there? Fifteen.
   c. Are there more differences in the exons or in the introns? In the introns. Can you think of an explanation for this? There would be no selection against a change in an intron.
   d. How can there be a difference in the exons if the proteins are identical? The genetic code is degenerate, so more than one codon can code for the same amino acid.

Acknowledgment

Thanks to Nadav Kupiek for expert preparation of the artwork.

References


SUSAN OFFNER is a biology teacher at Lexington High School, in Lexington, MA 02421; e-mail: soffner@ix.netcom.com.