

# Name That Gene

**Objective:** In this activity you will use the website for the National Center for Biotechnology Information to identify a sequence of bases from a DNA sample.

Background: The NCBI contains a database of genes sequenced and identified. The work of a number of scientists and a wide variety of areas allows for information to be compiled into this database. The tool used in this activity is BLAST - Basic Logical Alignment Search Tool, which allows a user to submit a sequence of DNA bases (A, T, G, C) into the search engine, BLAST will search the database and find the gene (if any) associated with that sequence. Furthermore, the user can access information about the gene or the disease it causes.

## Instructions

You will be given nucleotide sequences found in real human DNA that is associated with a genetic disease when mutated. Your job is to compare the sequences you are given with the nucleotide sequence of most known genes, using the BLAST tool to search genetic databases. Keep this window open while you perform your tests, this will make it easy to copy and paste gene sequences.

1. Go to the homepage for NCBI ( [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov))
2. Click on the word "BLAST" located under "popular resources"
3. Click on "Nucleotide BLAST"
4. Copy and paste one of the nucleotide sequences below into the top box. Basically, the program will run through all the known sequences and find a match. Depending on how busy the server is, this may take a while.
5. The closest match will be shown with a short description, for more information on your sequence click the link for GENE. This will give you the description of the gene that your sequence matched.

**Analysis:** On a separate page (or typed and printed) identify each of the genes associated with the below sequences. Write a brief summary that describes the gene and its importance to human and biological studies.

### Gene Sequence 1

ATG GCG ACC CTG GAA AAA GCT GAT GAA GGC CTT CGA GTC CCT CAA GTC CTT  
CCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GCA GC

---

### Gene Sequence 2

ATG GCG GGT CTG ACG GCG GCG GCC CCG CGG CCC GGA GTC CTC CTG CTC CTG  
CTG TCC ATC CTC CAC CCC TCT CGG CCT GGA GGG GTC CCT GGG GCC ATT CCT  
GGT GGA GTT CCT GGA GGA GTC TT

---

### Gene Sequence 3

ATG CTC ACA TTC ATG GCC TCT GAC AGC GAG GAA GAA GTG TGT GAT GAG CGG  
ACG TCC CTA ATG TCG GCC GAG AGC CCC AGC CCG CGC TCC TGC CAG GAG GGC  
AGG CAG GGC CCA GAG GAT GGA G

---

### Gene Sequence 4

ATG TTT TAT ACA GGT GTA GCC TGT AAG AGA TGA AGC CTG GTA TTT ATA GAA  
ATT GAC TTA TTT TAT TCT CAT ATT TAC ATG TGC ATA ATT TTC CAT ATG CCA  
GAA AAG TTG AAT AGT ATC AGA TTC CAA ATC T

---

### Gene Sequence 5

ATG CGT CGA GGG CGT CTG CTG GAG ATC GCC CTG GGA TTT ACC GTG CTT TTA  
GCG TCC TAC ACG AGC CAT GGG GCG GAC GCC AAT TTG GAG GCT GGG AAC GTG  
AAG GAA ACC AGA GCC AGT CGG GCC

---

### Gene Sequence 6

ATG CCG CCC AAA ACC CCC CGA AAA ACG GCC GCC ACC GCC GCC GCT GCC GCC  
GCG GAA CCC GGC ACC GCC GCC GCC GCC CCC TCC TGA GGG ACC CAG AGC AGG  
ACA GCG GCC CGG AGG AC

---

### Gene Sequence 7

ATG TTG TGC AAT ATC CAT CTA CTG TAG TTA AGA TAT TCA GTA GTT TGT TTT  
TCA TAA GCA TGT AAT TGA TCA TAT TTC TGC CAA GGA TGT GCC TTC AAC TTT  
ATA ATT ATA GTG TTG TAA AAT ATT TTT GTC TG

---

### Gene Sequence 8

ATG CCA TCT TCC TTG ATG TTG GAG GTA CCT GCT CTG GCA GAT TTC AAC CGG  
GCT TGG ACA GAA CTT ACC GAC TGG CTT TCT CTG CTT GAT CAA GTT ATA AAA  
TCA CAG AGG GTG ATG GTG GGT GAC CTT

---