Beginning with Biology Workbench

Sign into Biology Workbench or register for an account if you do not have one already.

Click on the link to run Biology Workbench.

Start a new session by highlighting the command and clicking run. You can name your session anything you want. I chose "Cystic Fibrosis Tutorial."

Our First Search – Protein Tools

Once you have begun a new session, highlight Ndjinn Multiple Database Search. Run.

Choose the SDSCNR database for this search. It is a larger database, but we will be using this specific one because it is non-redundant in its sequences and includes many species, including humans. It will give us the sequence we want.

Search "NM_000492 AND Homo sapiens." We are searching this because NM_00492 is the name of the locus where the cystic fibrosis gene is located. We are searching Homo sapiens because we will be looking at the genetics of the disease in humans before looking at other species.

This database, since it is large, will take a few minutes to load the matches. Once the search is complete, you will see that there is only one sequence matching our criteria. Check the box to the left of it and import the sequence.

Editing the Protein Sequence for Exon 10

We will be looking at exon 10 for cystic fibrosis mutations. Exon 10 is where the most common mutations occur, and is simple to find.

Check our previously imported protein sequence and highlight Edit Protein Sequence. Run.

Change the format of the sequence from Fasta to PIR/CODATA. This allows you to see numbers that help you to find the position of a specific amino acid used in the sequence.

Let's copy exon 10 from the sequence. It runs from amino acids 465-528.

Abort the protein sequence. This is the same as exiting a file without saving. We cannot save the sequence in this file because Biology Workbench looks at it as corrupted and will not save it.

Once at the protein sequence screen, select Add New Sequence and paste Exon 10 into the box. Name the file Exon 10 and Save.

Thr Ser Leu Leu Met Val Ile Met Gly Glu Leu Glu Pro Ser Glu Lys Ile Lys His Ser Gly Arg Ile Ser Phe Cys Ser Gln Phe Ser Trp Ile Met Pro Gly Thr Ile Lys Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr Arg Ser Val Ile Lys Ala Cys Gln Leu Glu

Let's Mutate a Protein Sequence – Deletion of Phe at AA 508

We will be using the mutations found at http://www.genet.sickkids.on.ca/cftr/PicturePage.html?domain_id=10 for this tutorial.

Check the box next to your Exon 10 sequence and select Copy Protein Sequence. Edit the sequence just as you have done before. Change the format to 3-letter Protein Codes.

You can use the find command on your computer (ctrl + f) to create a mutation. We will be doing the deletion of Phe between nucleotide 1653-1655. The amino acid is number 508.

In order to delete the correct Phe, sue the find command on your computer and search Phe Gly. This will ensure that we find the right one. Once you find it, delete it.

Thr Ser Leu Leu Met Val Ile Met Gly Glu Leu Glu Pro Ser Glu Lys Ile Lys His Ser Gly Arg Ile Ser Phe Cys Ser Gln Phe Ser Trp Ile Met Pro Gly Thr Ile Lys Glu Asn Ile Ile Gly Val Ser Tyr Asp Glu Tyr Arg Tyr Arg Ser Val Ile Lys Ala Cys Gln Leu Glu

Next, we will align our normal Exon 10 with our mutated version and see what happens to the sequence.

Check both versions of Exon 10 and select CLUSTLW - Multiple Sequence Alignment and Run. Choose Submit on the next page and look at our alignment.

The correct alignment should have a – for our deletion with an F directly under it, which corresponds to Phe in the normal sequence.

Our Second Search – Nucleic Tools

This will be just like our search for a protein sequence.

Go to Nucleic Tools and select Ndjinn Multiple Database Search. Run.

Choose the H_sapiens_mRNA database for this search. It only has human sequences in this database, but it is difficult to find the nucleic acid sequence in the larger non-redundant database.

Search "90421312."

Once the search is complete, you will see that there is only one sequence that was found. Check the box to the left of it and import the sequence.

Editing the Nucleic Acid Sequence for Exon 10

Like editing a protein sequence, this will be just as simple. Open the sequence for editing in Nucleic Tools and change the format to PIR/CODATA from Fasta.

According to our webpage of exon 10, the portion we are looking for is from base pairs 1525-1716. Copy this sequence and choose Abort.

We will now add our own sequence and name it exon 10. Look back at how to edit the protein sequence for exon 10 for help.

ACTTCACTTCTAATGGTGATTATGGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTCATTCTGTTCTCAGTTTTCCTGGATTATGCCTGGCACCATTAAAGAAAATATCATCTTTGGTGTTTCCTATGATGAATATAGATACAGAAGCGTCATCAAAGCATGCCAACTAGAAGAG

Let's Mutate Some Nucleic Acid Sequences

These mutations may or may not cause Cystic Fibrosis, but they provide great examples of different mutations that occur and can cause disease.

Frameshift – Deletion of 10 Base Pairs at BP 1540

This mutation would have been called a deletion of amino acids, but because it is not a multiple of three and it does not start at the beginning of an amino acid, it shifts the types of amino acids all the way down the sequence.

Go to Nucleic tools and check your exon 10 sequence and edit it. While in edit mode, search for the base pair sequence from 1541-1550. The sequence is GATTATGGGA. Delete it.

Then, use CLUSTLW – Multiple Sequence Alignment to see the difference between a normal exon 10 and the mutated exon 10. Be sure to check both boxes and select CLUSTLW. Run.

Now, align this mutated sequence with the normal exon 10 sequence. You should notice that there are 10 dashes were the base pairs once were on the one we frame shifted.

ACTTCACTTCTAATGGAGAACTGGAGCCTTCAGAGGGTAAAATTAAGCACAGTGGAAGAATTTCATTCTGTTCTCAGTTTTCCTGGATTATGCCTGGCACCATTAAAGAAAATATCATCTTTGGTGTTTCCTATGATGAATATAGATACAGAAGCGTCATCAAAGCATGCCAACTAGAAGAG

Translating the Nucleic Acid Sequence to Protein Sequence

Using biology workbench you can change your nucleic acid sequence into a protein sequence to see what the change is in the amino acids. There could be a stop codon that halts the rest of the sequence. There could also be a minimal change where the protein still functions properly because the size of the amino acid or it's properties are the same, or a bigger change where the protein is made, but does not function properly.

In order to change your sequence into an amino acid sequence, check the box of the sequence you'd like to see as amino acids and select SIXFRAME – Generate & Import 6 Frame Translations of NS. Run and choose Submit.

The next page will show many different options for the amino acid sequence. Scroll to the bottom and choose the one labeled as having the longest ORF. Change the name and import the sequence and there you go. Your nucleic sequence is now in amino acid form, too.

From the Protein Tools page you can use CLUSTLW – Multiple Sequence Alignment to see the changes made in the sequence of amino acids from the normal exon 10. This shows you how much of an impact your mutation actually made considering the amino acids.

Met Glu Asn Trp Ser Leu Gln Arg Val Lys Leu Ser Thr Val Glu Glu Phe His Ser Val Leu Ser Phe Pro Gly Leu Cys Leu Ala Pro Leu Lys Lys Ile Ser Ser Leu Val Phe Pro Met Met Asn Ile Asp Thr Glu Ala Ser Ser Lys His Ala Asn

Sequence Variation (similar) – G to T at BP 1570

This mutation produces a similar amino acid in place of the original. The mutation is G to T at base pair 1570.

What you want to do is edit the exon 10 sequence in Fasta format and search for the portion GTA. When you find it, replace the G with T. There is your mutated sequence.

Save the sequence as something to remind you of the mutation. I chose "Sequence Variation (similar) – G to T."

Look at the alignment using CLUSTLW and use SIXFRAME – Generate and Import 6 Frame Translations of NS to see your base pair alignment and translate the sequence into amino acids.

Once again, you can use CLUSTLW to align the normal exon 10 and the mutated version to see the difference the mutation creates.

Sequence Variation (different) – A to G at BP 1670

Go back to the Nucleic Tools tab and edit exon 10 and search for the portion of the sequence ATGAA. Replace the A with G.

For this mutation, follow the same procedure as the previous variation. I named my mutation edit "Sequence Variation (different) – A to G."

Exon 11 Missense Mutation that is a Common Cause of Cystic Fibrosis – C to T

This mutation on exon 11 creates a stop codon in the middle of the amino acid sequence, therefore terminating the rest of the sequence following it. The change of an Argenine to a Stop codon is dramatic and a common cause of Cystic Fibrosis.

Editing Sequence for Exon 11

Exon 11 can be found using the same sequence we used to find exon 10. Exon 11, however, runs from base pair 1717-1811.

Go to the Nucleic Tools tab and edit our original nucleic acid sequence.

Change the format from Fasta to PIR/CODATA and search for base pairs 1717-1811. Copy this sequence and choose to Abort.

Add a new nucleic sequence and paste exon 11 into the blank box. Name the sequence "Exon 11" and save.

GACATCTCCAAGTTTGCAGAGAAAGACAATATAGTTCTTGGAGAAGGTGGAATCACACTGAGTGGAGGTCAACGAGCAAGAATTTCTTTAGCAAG

Let's Create our Exon 11 Mutation – Missense of C to T at BP 1789

Edit our Exon 11 sequence. Use your computers find command to search for the sequence CG. There is only one in the entire exon.

Replace the C that you find with T. This creates our mutation. Save as "Exon 11 – Missense Mutation of C to T" and use CLUSTLW – Multiple Sequence Alignment to see the change we made.

GACATCTCCAAGTTTGCAGAGAAAGACAATATAGTTCTTGGAGAAGGTGGAATCACACTGAGTGGAGGTCAATGAGCAAGAATTTCTTTAGCAAG

Use SIXFRAME to change both Exon 11 and Exon 11 – Missense Mutation of C to T into amino acid sequences, and then use CLUSTLW once again to align the sequences.

Our Third Search - Protein Structure

We will now use Biology Workbench to take a look at the actual protein structure of a normal protein and mutated protein. You will be able to see where the mutation is on the protein and what the effect on its shape will be, and therefore, its function.

Select Ndjnn – Multiple Database Search and select the PDPFINDER database. This will give us a 3 dimensional protein structure for our normal and mutated protein that we just found in Biology Workbench on Exon 11.

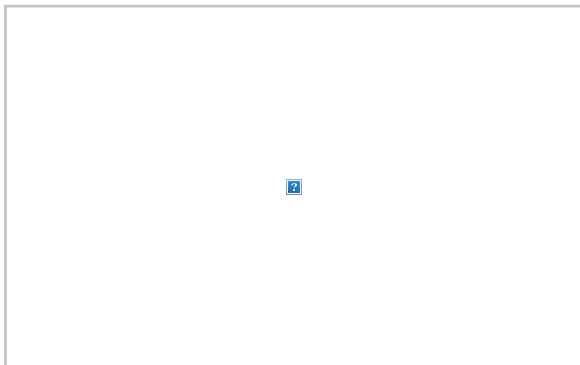Search "2PZE" to find the normal protein. There will only be one structure matching our criteria.

Check the box on structure that we find and choose to Show Record. On the next page, click the link PDB Structure Explorer, which can be found at the top right corner of the page.

Just under the picture of a protein structure, you can find the link called QuickPDB. Click on that and it will take you to our protein structure.

Change the color that you want to highlight the portion of the protein we are looking at by left clicking on the color box and then on a color of your choice. Try to choose a color that you can see against green and black. I chose blue.

In the sequence of letters above the protein and next to the words Chain A, highlight "IIFGV." This will change that section of the protein structure to the color you chose.

Choose float so that the structure is in its own window, and maximize that window. So that you can compare proteins, find the "prt sc" button on your keyboard and press it so that the page you are viewing is copied. Now, open up a word document and paste the photo into the document.

Head back to Biology Workbench.

Select Ndjnn – Multiple Database Search and select the PDPFINDER database.

Search "2PZF" this time to find the mutated protein. There will only be one structure matching this search.

Check the box on structure that we find and choose to Show Record. On the next page, click the link PDB Structure Explorer, which can be found at the top right corner of the page.

Just under the picture of a protein structure, you can find the link called QuickPDB. Click on that and it will take you to our mutated protein structure.

Change the color once again (I chose the same color as before).

In the sequence of letters above the protein and next to the words Chain A, highlight "IIGV." It will be around the same place as the previous highlight but without the F.

Choose float so that the structure is in its own window, and maximize that window. So that you can compare proteins, find the "prt sc" button on your keyboard and press it so that the page you are viewing is copied. Now, paste the photo just under your normal protein.

You may want to type a heading above each picture so that you can recall which one is normal and which one is mutated. Save the file to your computer if you would like to look back at the protein structures, or you can go through the procedure again.