**Amino Acid Tutorial**

When the amino acid numbers are not the same between the horse and human genomes, it is important to determine where the discrepancy in the number of amino acids is coming from. It is also important to look at the gene across multiple species to try and get an understanding of how conserved that region is and if the difference in number could just represent normal species variation.

1. Take a closer look at the transcripts to find where the change in amino acid number is coming from? Is the change in base pairs on just one exon, multiple exons?
2. Before proceeding with multiple species comparison it is important to assess if the other parameters are appropriate. This includes checking to see if the sequence contains both 5’ and 3’ untranslated regions (UTR’s), are start and stop codons present, number and length of exons, and are the splice junctions normal.
3. If all of these other parameters are appropriate then proceed with a multiple species comparison. This is done to look at the gene across many species and helps determine if the changes in AA sequence are normal species variation or if they are significant and could indicate a problem. Variations that are in non-conserved regions of the genome are more likely species variation.
4. If other parameters aren’t appropriate then trouble shoot those issues and proceed with BLAST. Using BLAST allows direct comparison of the human genome to the horse and can provide information to update exon boundaries and improve the gene annotation.

\*If you assess all the above steps and they are appropriate then proceed down the flow chart.\*

**Exon Tutorial**

1. If the number of exons are NOT the same there are many questions we have to answer.

-Is the number different but the amino acid sequence is the same?

-Does the horse possess more/less exons than the human and this is not of significance and just the normal variation?

\*GYS1 (multiple sequence alignment and also transcript table pics)

-Is there a problem with the splice junctions which affected the number of exons?

-Is there a start/stop codon indicating that you most likely have the entire transcript and no exons are missing?

1. Is the number different but the amino acid sequence is the same?

-This may indicate that the horse just has a normal change from the human since the amino acid number is conserved. This can be assessed by comparing the exon lengths and amino acid sequence across species.

\*\*\*\*\*CONSERVATION PROCESS\*\*\*\*\*

1. Is there a problem with the splice junctions which affected the distribution of exons?

-Proceed by looking at all of the splice junctions and ensuring they are correct and start with a “GT” and end with an “AG”.

1. Is there a start/stop codon indicating that you most likely have the entire transcript and no exons are missing?

\*If you assess all the above steps and they are appropriate then proceed down the flow chart.\*

**First Exon Tutorial**

Assessing the first exon includes looking for a proper START codon (ATG) and the presence of a 5’ UTR.

1. Is there a missing START codon?

-Example of VAMP2 gene with pics

-Looking at other transcripts to see if one matches up as that could indicate there is a predicted equine transcript variant that just don’t have data yet

1. Is the 5’ UTR missing?

-If the 5’ region is not present this could indicate that the transcribed sequence does not go far enough upstream and some of the sequence is not transcribed. -To proceed BLAST the human exons to the horse genome to determine the base pair locations of the sequence.

\*If you assess all the above steps and they are appropriate then proceed down the flow chart.\*

**Exon Length Tutorial**

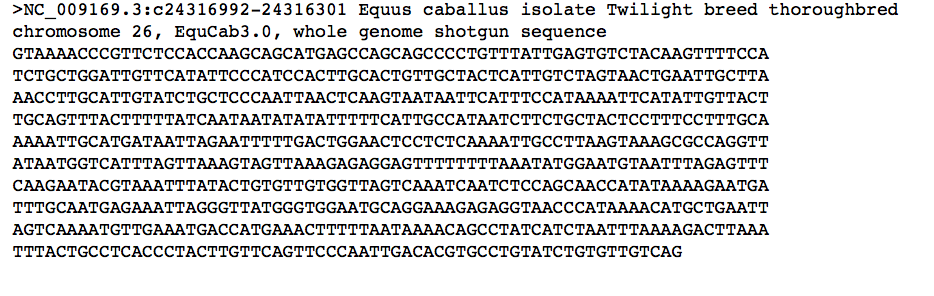
\*\*SAME as amino acid\*\*\*

\*If you assess all the above steps and they are appropriate then proceed down the flow chart.\*

**Splice Junction Tutorial**

1. Is there a problem with the splice junctions which affected the distribution of exons?

-Proceed by looking at all of the splice junctions and ensuring they are correct and start with a “GT” and end with an “AG”.



1. If the splice junctions are not correct🡪 Proceed by BLAST’ing the sequences of the exons to determine where the change is occurring.

\*If you assess all the above steps and they are appropriate then proceed down the flow chart.\*

**Last Exon Tutorial**

Assessing the last exon includes looking for a proper STOP codon (TAA, TAG, TGA) and the presence of a 3’ UTR.

1. Is the 3’ UTR missing?

-Significance if there is a stop codon but no 3’ UTR?

\*\*\*would look to rnaseek but not here

-If there is no stop codon and no 3’UTR proceed with the BLAST (typically more upstream issue)

1. Is there a missing STOP codon?

-If the stop codon is not present this indicates that the current transcript does not go far enough downstream and is missing the end of the sequence and the 3’ UTR.

-To try and locate this missing portion of the sequence proceed with BLAST’ing the sequence of the last few human exons (including the last one containing the stop codon) to the horse genome to locate if/where this sequence is present in the horse.

\*If you assess all the above steps and they are appropriate then proceed down the flow chart.\*

\*\*\*ADD TO NORMAL TUTORIAL

-Add looking at other human transcripts for a match

-Multiple sequence alignment (other programs for options)