**1. Is your sequence a protein sequence or nucleotide sequence?**

My protein is a nucleotide sequence (mRNA of the Prlr gene).  
 **2. What species or organism does the sequence come from?**

Mus musculus (or house mouse)

**3. What are the associated references, either journal articles or submission references?**

**Direct Submission**

Moore, R.C., Submitted (30-MAR-1993) R.C. Moore, Laboratory of Molecular and Cellular Biology, National Institute of Diabetes, and Digestive and Kidney Diseases, NIH, Bethesda, MD 20892, USA

**Journal Articles**  
Moore, R. C., & Takami, O. (1993). Cloning and sequencing of the cDNA encoding the murine mammary gland long-form prolactin receptor. *Gene*, *134*(2), 263–265. doi:10.1016/0378-1119(93)90104-B

**4. When was the sequence submitted? Has the sequence been revised? Has the flatfile been revised?**

It was submitted on March 30, 1993. The sequence has not been revised, it is still in version 1 - L13593.1. The flatfile appears to have been revised, because the last modified date is February 9, 1994, almost a year after the submission date.

**5. How long is the sequence?**

1992 bp  
 **6. Can you find all the related sequences? For example, if you have a protein sequence can you find the mRNA or if you have a genomic sequence can you find the mRNA sequence. Of course bacterial and viral sequences will be a bit different, i.e. no mRNA.**

# The protein product listed in GeneBank is AAC37641.1 (prolactin receptor [mus musculus]).

# Searching in BLAST, I found the complete genomic sequence “Mus musculus chromosome 15, clone RP24-326H11, complete sequence”. In BLAST, I also found several transcript variants of the mRNA, clone cDNA, mutant alleles, as well as orthologues of prolactin receptors across several species.

# RefSeq revealed 8 alternative splicing sequences for the Prlr gene. 7. Are there any unique features associated with your sequence?

The sequence has two unique features noted in GeneBank, a signal peptide coding sequence at 59..115 bp, and a mature peptide coding sequence at 116..1882 bp whose produce is a prolactin receptor.

References:

1. <https://www.ncbi.nlm.nih.gov/nuccore/347398>
2. <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
3. <https://www.ncbi.nlm.nih.gov/nuccore/1039748485,1039748484,568990391,568990387,568990385,359385719,359339022,254675323>

**Using Entrez gene to search for more information about your sequence from unit 1, what additional information can you find? For example, are there SNPs, ESTs, mRNA variants, or orthologs associated with your sequence? Write one paragraph summarizing your findings. Please DO NOT cut and paste the information you have retrieved from the database.**

Prlr (prolactin receptor) in the species, Mus Musculus, is a protein coding gene for the protein prolactin and is located on chromosome 15. It has several homologs, including human, chimpanzee and seven other species. The human Prlr gene has 197 orthologs. It is involved in 17 biological pathways, such as adipogenesis, studied in obesity and diabetes, and pathways involving ovarian infertility. The Prlr gene has 16 variants in mice. Studies involving the Prlr gene include obesity due to leptin resistance, cardiac dysfunction, ovarian and neural studies to name a few.

<http://www.informatics.jax.org/marker/MGI:97763>

**Subject: Ensembl and your ACC sequence (required)**  
**Using Ensembl gene to search for more information about your sequence from unit 1, what additional information can you find? For example, are there SNPs, ESTs, mRNA variants,or orthologs associated with your sequence? Write one paragraph summarizing your findings. Please DO NOT cut and paste the information you have retrieved from the database.**

Using Ensembl and searching for the Prlr gene in mus musculus, I found that the gene has 12 splice variants, 8 of them protein coding, which matches up with what I found in RefSeq. In Ensembl, I found 74 orthologs, the most closely related species being Rat (*Rattus norvegicus*). It also has 3 paralogs. The Prlr gene is a member of one Ensembl protein family, that also includes the gene Gm21973, which is a predicted gene. The Prlr gene has 46 phenotypes. There are numerous SNPs, the variant table show 4380 of them.

Choosing the first transcript in the list with 608 amino acids, I found 10 exons, 8 of them coding exons.

FEATURES CAN BE FOUND BY SEARCH ENSEMBL FINDING THE TRANSCRIPT AND GOING TO UNIPROT

**Subject:****UCSC and your ACC sequence (required exercise)**

**Using UCSC to search for more information about your sequence from unit 1, what additional information can you find? For example, are there SNPs, ESTs, mRNA variants,or orthologs associated with your sequence? What contigs are associated with your gene?**

Using the UCSC Genome browser, it is showing 142 SNPs for my mRNA. This is conflicting with what I found in previous weeks, so I must investigate that more. There is one contig for the sequence shown in Genome browser that covers the entire sequence. The gene has 38 spliced ESTs. It appears that the gene has no CpG Islands, but if I zoom out, I find them on surrounding genes. I found five simple repeats.

I used the Comparative Genome section to visual conservation across species, which I found particularly interesting. I added a screenshot of the image below. As you can see, this gene is highly conserved in the Rat species, as I found last week in Ensembl. It also is strongly conserved across human, rabbit, shrew, dog and elephant species.

