**Specific Aims**

There are many genes in the human body responsible for efficient running economy, which can be quantified by measuring energy utilization while moving at an [aerobic](http://en.wikipedia.org/wiki/Aerobic_exercise) intensity. The gene of interest to this project is COL5A1. COL5A1 codes for the protein pro-alpha-1(V) chain. Three of these proteins come together to form a molecule of type V collagen. Type V collagen helps to strengthen and support the extracellular matrix as well as many tissues such as ligaments, tendons, and muscles. Type V collagen also plays a role in forming these structures with type I collagen. COL5A1 is conserved in most organisms and is necessary for survival during development. DNA sequencing has shown that a single nucleotide polymorphism (SNP) in the 3' untranslated region (UTR) of COL5A1 is linked to enhanced running endurance. This variation has been labeled rs12722. It is unknown how the COL5A1 mutation affects cellular functionality to produce this trait.

The **objective** of this study is to better understand how this variant protein contributes to an improved running economy by looking at its role in collagen fibril formation with type I collagen. There is speculation that rs12722 increases the stability of type V collagen, allowing the protein to efficiently absorb energy produced by impact from running.

**The hypothesis is that type V collagen produced by rs12722 enhances running economy through formation of more durable collagen fibrils and a more stable interaction with type I collagen.** It is known that type V collagen is the precursor to collagen fibril assembly in some muscles. The fibrils made from pro-alpha-1(V) provide the major biomechanical scaffold for cell attachment and allow for the shape and form of tissues to be maintained.

The **long term goal** of this study is to better understand the specific mechanism that causes rs12722 type V collagen to improve running economy. The best way to achieve this is to reach a better understanding of how the rs12722 pro-alpha-1(V) chain affects functionality of ligaments in the body.

**Specific Aims:**

**1. Analyze the human DNA sequence of COL5A1 to see if the SNP is within a DNA motif or regulatory region that affects transcription.** This could determine whether the stability of the protein is due to increased expression from a more stable mRNA transcript, or if the end product protein contains a stronger domain in which to bind type I collagen. This can be tested by using the databank DREAM, which contains known DNA motifs.

**Rationale:** If the SNP is within a motif region of the 3' UTR of the COL5A1 gene, then perhaps that somehow alters function to promote stronger interactions with type I collagen.

**2. Analyze the transcriptome of mice embryos to see if more of rs12722 is being produced in ligaments.** By producing homozygous mutant mice, heterozygous mice, and wild type mice, RNA Sequencing of ligament tissues in mice embryos can be used to see if more copies of the rs12722 pro-alpha-1(V) chain are being produced compared to the normal pro-alpha-1(V) chain.

**Rationale**: Mice are easy to handle and the COL5A1 gene is highly conserved compared to other organisms. By seeing if there is more rs12722 present, an assumption could be made that the quantity of type V collagen in ligament tissues somehow is related to the increased running economy phenotype.

**3. Use a yeast two-hybrid to see what proteins interact with rs12722.** By performing a library yeast two-hybrid approach with both the rs12722 variant and the normal variant of COL5A1, I can see if there are any noticeable or new protein interactions with the rs12722 protein. I can then confirm these interactions are indeed new and stable with a matrix yeast two-hybrid approach.

**Rationale:** There could be a protein that interacts with the rs12722 protein, but not with the normal COL5A1 gene product, that promotes a more stable collagen fibril.