Evolutionary Comparison of αlpha Subunits in The CACNB2 Gene of Humans and Rodents

May 26th, 2021
University of Delaware
Gene of Interest Final Project
By: Macie McCarthy

Introduction

It is well known in the world of evolution and genetics that humans share a good portion of their genome with many rodents. Mice, rats, and guinea pigs, are just a few of the rodents that share a lot of the same DNA. In fact, one very prominent thing that humans have in common with rodents are their immune systems. According to Genome.gov, “The rat model is poised to be at the forefront of discovery, providing insight into human health and new treatments for human diseases," said NHLBI Acting Director Barbara Alving, M.D…..The rat sequence draft, which covers more than 90 percent of the genome, represents the third mammalian genome… Almost all human genes associated with diseases have counterparts in the rat genome and appear highly conserved through mammalian evolution.” (1*).

This is where the CACNB2 gene comes into play. The gene of interest, CACNB2, is a voltage-dependent L-type calcium gene. This gene is identified as an antigen target in Lambert-Eaton myasthenic syndrome, an autoimmune disorder. Mutations in this gene are associated with Brugada syndrome, and so, it is Classified as an autoimmune antigen target (4*). CACNB2 is a prominent gene in humans’ autoimmune response and the way the body reacts to disease, infections, and viruses. One of the few subunits of this gene is known to code for an extremely crucial protein called 1T0J which is also known for its functions in the immune system.

The CACNB2 gene can be found on the 20th chromosome in humans as well as the specific location of 10p12.33-p12.31. It has a total of 20 exons and 21 introns as well. Something worth noting is that the protein structure here is highly conserved between orthologs and subtly conserved when it comes to its paralogs. This gene hints to be very crucial within different species and, yet, it does not seem to fight mutations all that well.

Because of all this information and more, the hypothesis of this gene is that this protein’s evolutionary lineage would align most closely with rodents such as mice, rats, and rabbits. The human ortholog of this gene will have a most recent common ancestor with these rodent species and share a high amount of conservation and redundancy via the DNA sequence all the way to its protein structure and function.

Methods

The first step that was carried out in order to find CACNB2’s evolutionary lineage was to pick a subunit that yielded a gene that was relevant to the topic of autoimmune response in humans. The gene was searched in the website for National Center for Biotechnology (NCBI) under the protein database. Once the gene was found under Homo sapiens, the conserved domains tab gave all three of the domains involved: two different beta-2 subunits as well as an alpha-1 subunit. The beta subunits were first analyzed however it never gave any accurate results. All that came in the BLAST hits of NCBI were isoforms which were the same version of the gene with just different splicing versions.

Analyzing the alpha-1 subunit of CACNB2 using BLAST yielded multiple paralogs with little to no isoforms as well as more variance with conservation. The sequences were aligned via the ‘clustalW’ algorithm, yielded a phylogenetic analysis which then created a dendrogram using MegaX. A very similar process was conducted in order to find the orthologs of this gene. The protein, 1T0J, was put through a Blastp search without classifying a specific species or group of species. Conveniently enough, all the hits that followed were mostly vertebrates. The top 12 sequences with the highest hit queries were then isolated to conduct an alignment. The sequences were put through an ‘alignment with clustal W’ algorithm, a phylogenetic analysis, and then generated a dendrogram on MegaX.

In order to uncover the multiple sequence alignments of both the paralogs and the orthologs found above, the hits of interest had to be checked off and aligned together using NCBI. This page then showed the conservation of each sequence with one another along with the graphical overview, descriptions, and the full sequence alignment of each paralog and ortholog. The accession numbers were also labeled next to each sequence in the overview in order to find each proteins' location, exon count, and exon boundaries using the sequence viewer and the various tools available.

Next was achieving a viable structural alignment between humans and the rodent(s) of choice. This was achieved by first searching the CACNB2 gene and 1T0J protein under the 'structure' database in NCBI. The search result mostly yielded structures of humans, rodents, or both together. The structure that included both humans and a rodent species was analyzed using the 'Full 3D viewer tool'. These domains were aligned structure to structure using this viewer tool. The PDB code for Homso sapiens as well as Rattus norvegicus (Norwegian brown rat) were used in the structural alignment. This then showed the complete 3D structure of the ortholog protein subunits interacting as well as the conservation in the model. This process was also completed the exact same except the comparison was between Homo sapiens and Oryctolagus cuniculus (rabbit).

Results

- Reference Figures -

Evolutionary history of autoimmune response of a retroviral infection from 'Frontiers In Microbiology'

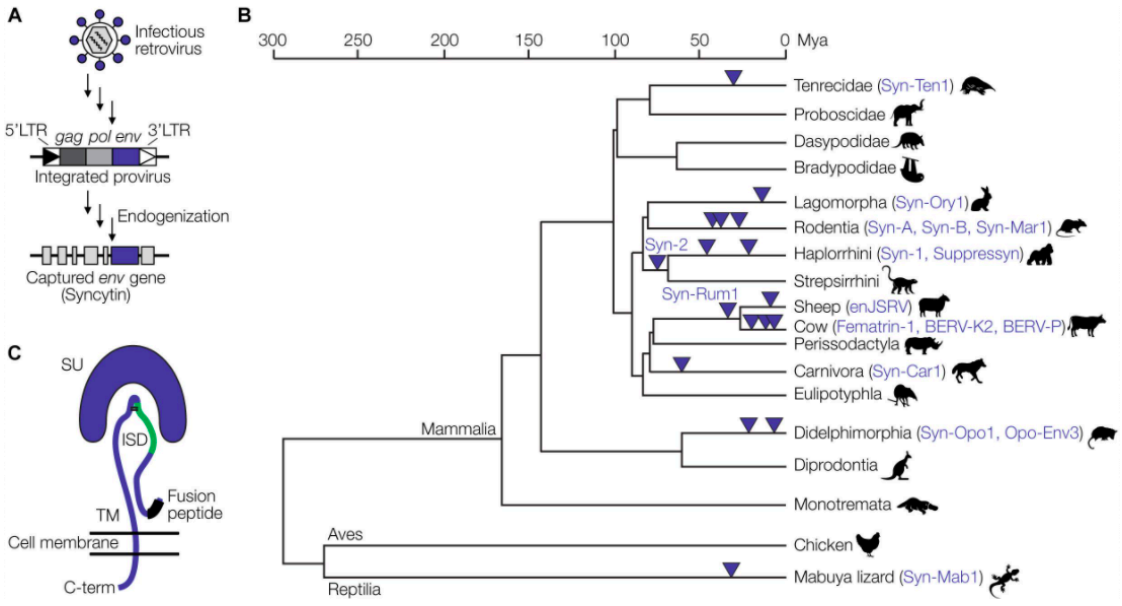

Figure 1. In the image above found from 'Frontiers in Microbiology', it indicates the evolutionary tracing of the retroviral envelope gene that fights off the infection used in this annotation. The evolutionary history above is shown via a dendrogram of multiple mammalia. About 90 million years ago, there was a clade that began to speciate into bunnies, rats, and gorillas (primates). This clade has a high density of the presence of the retroviral envelope gene to fight off the infection. This indicates a shared, most recent common ancestor between rodents and primates in the bodily response to a viral infection (5*).

‘Cold Spring Harbor Laboratory’ experiment of a System-Humanized rodent model

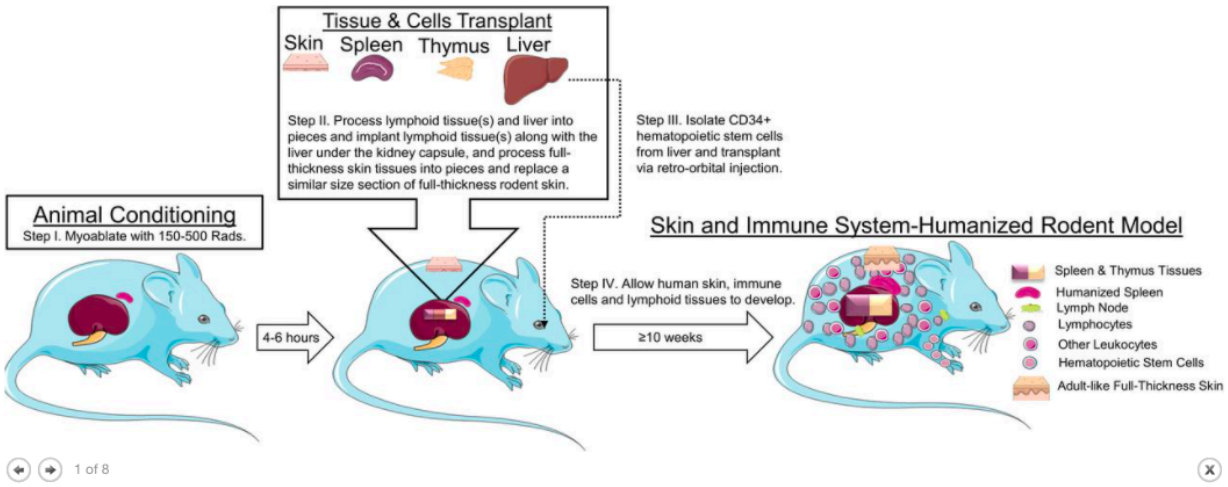

Figure 2. The display above shows a visual of an experiment done on rats using human tissue conducted by ‘Cold Spring Harbor Laboratory’. Severely immunodeficient rats and mice went through a series of procedures and treatments with human lymphoid, liver, and fetal skin tissue. The rodents were then administered antibiotics and further tests in order to measure the stem cell count after the tissues were inserted. After roughly 10 weeks, the results accumulated were an immune system-humanized rodent model of both rats and mice (3*)

- Paralogs -

Dendrogram of paralogous Homo sapien proteins

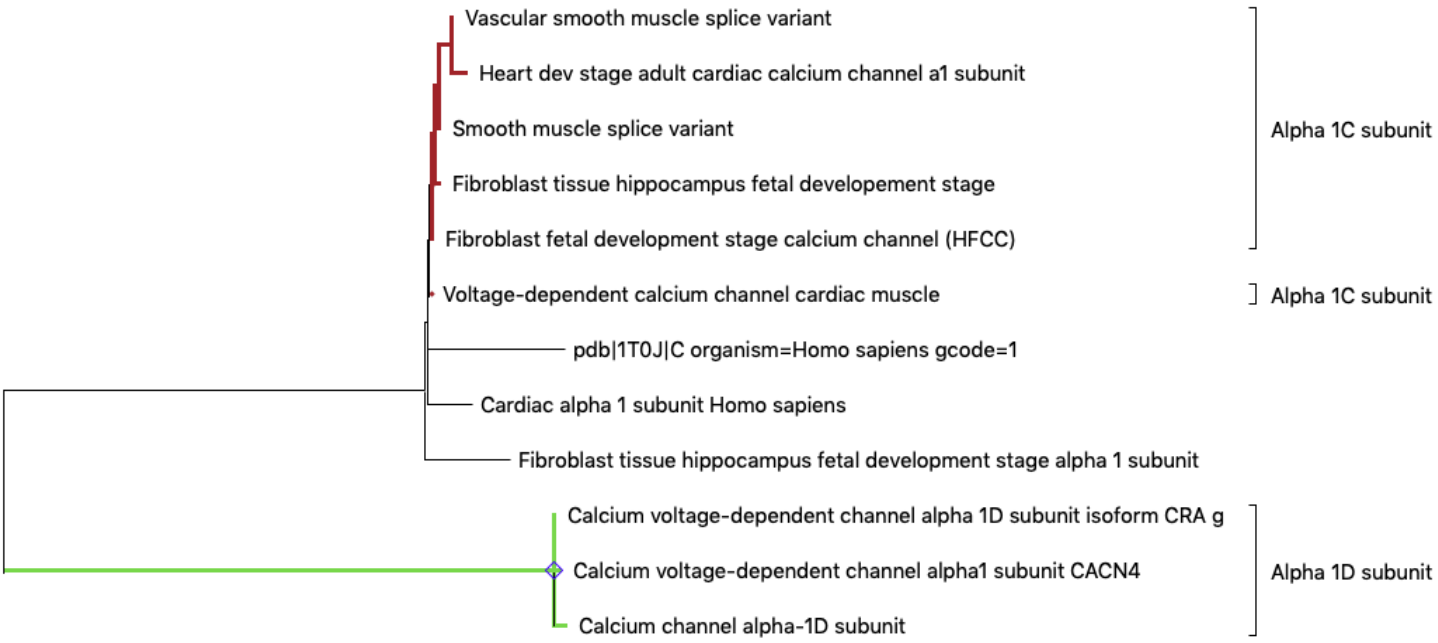

Figure 3. The dendrogram above depicts the evolution of the CACNB2 gene in humans over time rooted using the protein 1T0J. The green branches represent the Alpha 1D protein subunits and red branches

represent the Alpha 1C subunit. The proteins that lie in the middle of the two color-coded clades are proteins of the Alpha 1 subunit and do not specify either of the C or D subgroups. These proteins do not show much connection to a clade like the other subunits do so in the dendrogram. The fact that the alpha 1C and 1D subunits are separated via clade, indicates an evolutionary trend going on in the different protein subunits of CACNB2.

Paralogous Protein Information

Protein Name	**Location**	**Exon Count**	**Subunit (color coded)**
Vascular Alpha 1C splice variant	12p13.33	56	Alpha 1C
Cardiac a1 subunit	12p13.33	56	Alpha 1
Smooth alpha 1C splice variant	12p13.33	56	Alpha 1C
Fibroblast alpha 1C	12p13.33	56	Alpha 1C
Fibroblast alpha 1C (HFCC)	12p13.33	56	Alpha 1C
Cardiac alpha 1C	12p13.33	56	Alpha 1C
Cardiac alpha 1	12p13.33	56	Alpha 1
Fibroblast alpha 1	12p13.33	56	Alpha 1
Calcium alpha 1D Isoform CRA	3p21.1	53	Alpha 1D
CACN4	3p21.1	53	Alpha 1D
Calcium alpha-1D	3p21.1	53	Alpha 1D

Table 1. The set up above goes into further depth on the similarities between the paralogous proteins displayed on the dendrogram in **Figure 3**. Each protein name is in its own row with the genome location, the exon count, and the subunit group it is a part of. All of the Alpha 1 and Alpha 1C subunits are, not only on the same chromosome, but also have the same location in the genome (12p13.33). They also have the same exon count of 56. Meanwhile, the Alpha 1D subunits all have the same chromosomal location (3p21.1) as well as the same exon count of 53.

- Orthologs -

Dendrogram of various orthologous 1T0J proteins of the CACNB2 gene

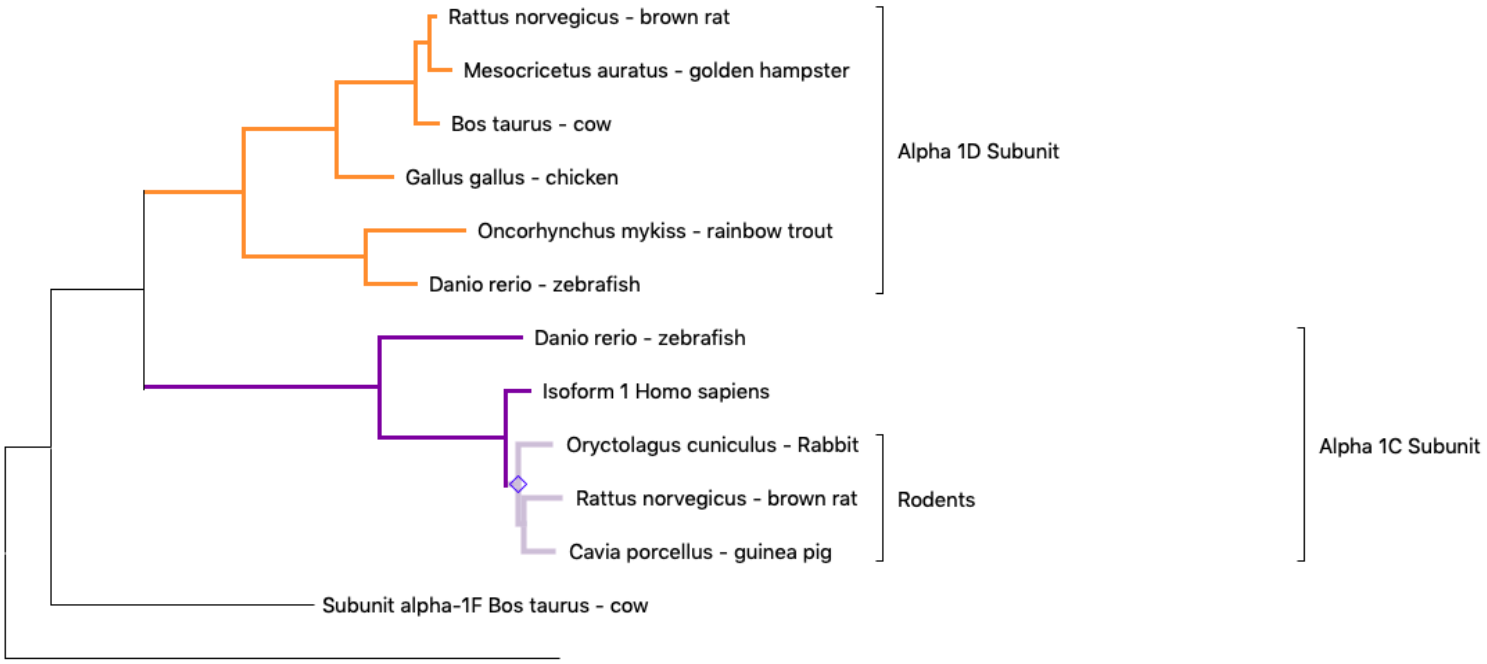

Figure 4. The dendrogram above depicts the Alpha 1D and the Alpha 1C subunit clades between various species of orthologs. The Alpha 1D subunit clade is shown in the orange color. The only commonality in this clade is the subunit involved; there shows no common evolutionary history of homosapiens and rodents, let alone Homo sapiens at all. Looking over at the Alpha 1C subunit shown in the dark purple, there is a recent common ancestor connecting humans and rodents. Better yet, in the light purple color, there is a speciation event directly related to humans that branches off into another clade designated for three different species of rodents: rabbits, norweigan brown rat, and the guinea pig.

Orthologous dendrogram subgroup of interest

Figure 5. Above is another illustration of a section of the dendrogram represented in **Figure 4**. This Alpha 1C subunit clade shown above are the orthologs whose evolutionary history will be analyzed further. This is due to the fact that Homo sapiens share a most recent common ancestor with three different rodents shown in purple: rabbits, guinea pigs, and norwegian brown rats.

- Multiple Sequence Alignment -

Multiple sequence alignment between humans, rodents, and zebrafish

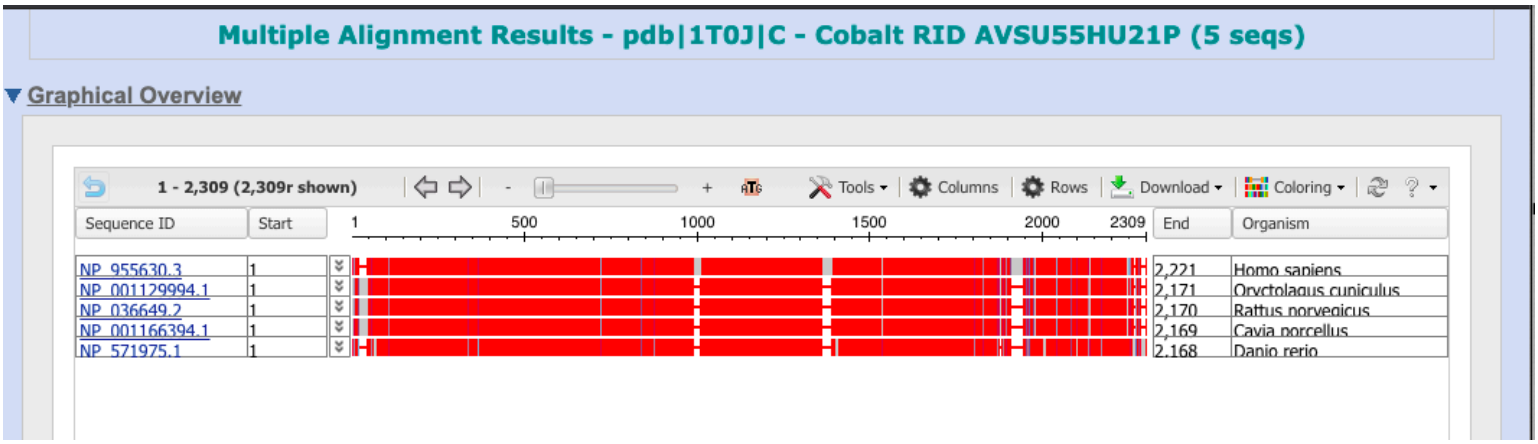

Figure 6. The alignment shown above indicates the graphical overview of the conservation between different orthologous proteins of different species. It aligns the sequence of the proteins of the dendrogram in **Figure 5**: humans, guinea pigs, rabbits, norwegian brown rats, and zebrafish. The image above indicates that these 5 protein sequences are extremely conserved across the different species used here. The conservation is shown via the red shaded parts of the aligned proteins, meanwhile the white and gray gaps show the sections of the proteins that do not align with one another. Thus, these proteins are homologous to one another due to the fact that they share so many similarities in their alignments.

- Structural Alignment -

3-D structural alignment between a human and Rattus norvegicus (Brown Rat)

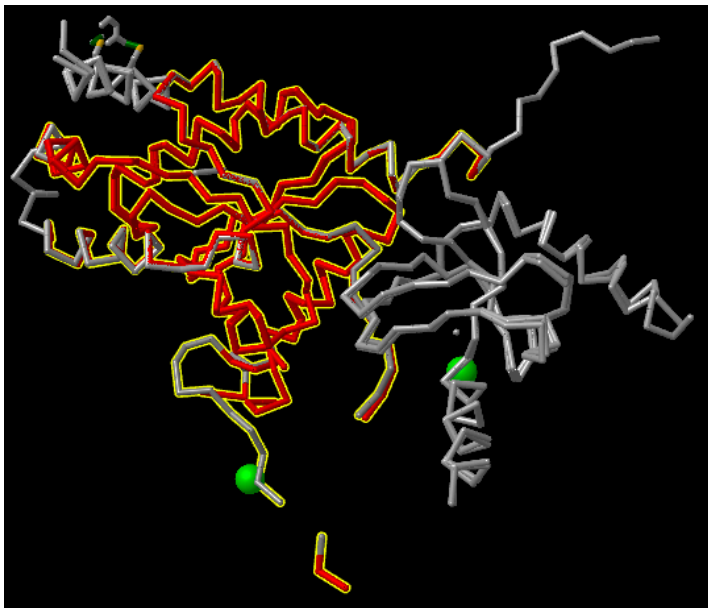

Figure 7. The above image shows the structural alignment of Homo sapiens (PDB: 1T0J) with Rattus norvegicus (PDB: 5V2Q). This was done by using a 3D image of the protein structures and two different domains from the two different species. The red, shaded part of the protein is where the amino acids of the protein here are conserved. Meanwhile, the gray areas are where the two protein domains are not conserved in space. The areas of the protein that are outlined in yellow are the protein domain interactions that are functioning with each other. Take note that the entire conserved region is outlined in yellow as well.

3-D structural alignment between Homo sapiens and Oryctolagus cuniculus (rabbits)

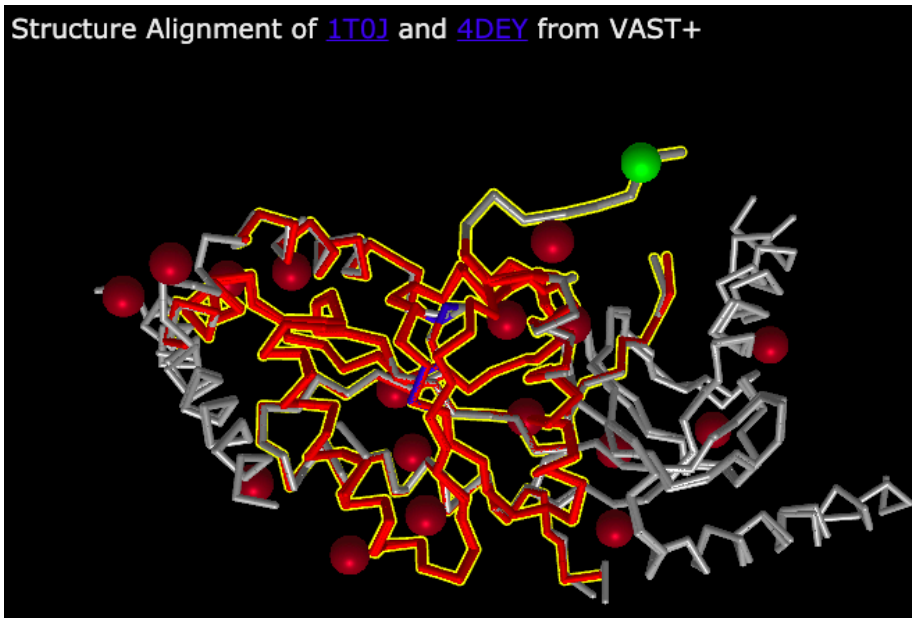

Figure 8. The above image shows the structural alignment of Homo sapiens (PDB: 1T0J) with Oryctolagus cuniculus (PDB: 4DEY). This was done by using a 3D image of the protein structures and two different domains from the two different species. The red, shaded part of the protein is where the amino acids of the protein here are conserved. Meanwhile, the gray areas are where the two protein domains are not conserved in space. The areas of the

protein that are outlined in yellow are the protein domain interactions that are functioning with each other. Just as in **Figure 7**, the red area of conservation between the two species is also where they interact with one another.

Discussion

The hypothesis of this gene evolutionary analysis started off with the fact that the CACNB2 gene was known for producing an antigen target for an autoimmune disorder. Thus, the protein 1T0J functions as an immune response in humans. Since humans and rodents have a lot of similarities between their immune systems, then they must show close evolutionary histories with one another. This connection was made due to the reference **Figure 1** and **Figure 2** along with NCBI. NCBI goes in depth about how mouse and rat models are critical for understanding antibody-antigen interaction and the response to infectious agents (3*). This immune response is just about word-for-word one of the immune functions that CACNB2 codes for. Due to CACNB2 coding as an antigen target, rodents and humans have a recent common ancestry with this gene and protein, thus, their functions, structures, and sequences highly relate to each other.

Looking back at the results of the orthologous dendrogram shown in **Figure 4**, shows the two major clades in purple and orange diverge off of eachother from the same clad drastically diverge from one another thousands of years ago. Usually when there shows divergence branching off into two different clades, it points to genetic drift, usually due to an abiotic factor. Since the focus is on immune response, an example would be a possible autoimmune disorder from consuming a certain diseased prey. This would cause a population to slowly evolve away from each other based on diet, physiology, immune system, or mutations. This would also explain how the two different clades (orange and purple) each separate into different alpha subunits. The alpha 1D subunit in orange and the alpha 1C subunit in purple. The evolutionary history between rodents and humans is prevalent in **Figure 4** because they both diverge together, go through a speciation event, *then* the rodents branch off into their own clade. Not to mention that they all have the alpha 1C subunit. This recent common ancestry supports the immunology evolution trends further.

Figure 6 displays the multiple sequence alignment between humans, rabbits, guinea pigs, and brown rats. As seen in the figure, the sequences show vast coverage of homology between the different species. This is seen by the red shaded portions of the alignments which indicates that these proteins are highly conserved to say the least. Proteins that are more conserved, show a stronger evolutionary history. According to NCBI: “Homology implies that the corresponding residues in homologous proteins are derived from the same ancestral residue and, typically, inherit its function. If the residue in question is the same in a set of homologous sequences, we say that it is (evolutionarily) ***conserved***” (7*). It is known that from high conservation, same subunits present, and the dendrogram that these proteins show a strong evolutionary history between humans and rodents, However, the question is, why are these proteins so similar, so conserved? The main reason as to why a protein may be so conserved is because it serves a crucial role in physiology. According to the ‘Protein Data Bank’ as well as NCBI, the protein 1T0J codes for a subunit that functions as a signaling molecule (8*). This protein’s function surrounds voltage-gated calcium channels. These channels form complexes with signalling molecules and receptors in all kinds of different tissues (9*). A function like this is crucial in any system, especially in the immune-response system in multiple species. A signaling molecule is the beginning of countless different cascade processes in cells to achieve an exponential response in the body. This process is crucial for the immune system and

an organism cannot have an effective immune response without these signalling molecules and the voltage-gated calcium channels present, and functioning to its best ability. A conserved protein cannot handle mutations like a normal protein would. A more conserved protein with a mutation would be ineffective and result in not functioning correctly. If an individual organism had ineffective immune response due to lack of voltage-gated channels and signal transduction, then that organism would not make it to the age of reproduction.

Analyzing the structural alignments of both human vs brown rat domain interactions, and human vs rabbit domain interactions: looking at **Figure 7** and **Figure 8,** at least half of the amino acids between both these comparisons are red which suggests they are structurally very similar, therefore are closely related and stem from a common ancestor. This also indicates that human, rabbit, and brown rats are all structurally similar to one another. This is just yet another level of similarity that these species' have a connection to. Human and rodent CACNB2 proteins are conserved via protein alignment, and structural alignment. This adds to the strong evolutionary relationship that is shared between the two on an immune response level.

Looking back at both the paralog and ortholog dendrogram, there shows trends that run through both figures. As indicated in **Figure 3**, the paralogs diverge into two larger clades; one green clade with the alpha 1D subunit, and the other red clade with the alpha 1C subunit. This divergence is significantly similar to the divergence that occurs in the dendrogram of **Figure 4**. The orthologous dendrogram splits itself into two major clades, one with alpha 1D subunits while the other has alpha 1C subunits. Another inference looking at these two phylogenetic trees are the similar exons and locations of each protein. In **Table 1** it shows that all alpha 1D subunits are located at (3p21.1) and have 53 exons; meanwhile all the alpha 1C subunits are located at (12p13.33) with 56 exons. This can be carried over to the proteins in **Figure 4** due to the strong conservation between all the orthologs of humans and rodents. Each species' alpha 1C subunits are most-likely in the same location with the same number of exons and same with their alpha 1D subunits.

Circling back to the hypothesis that humans and rodents have the most closely-related evolutionary history between their orthologs, can be accepted. The phylogenetic tree, the multiple protein sequence alignment, the structural alignment, as well as the research all accept this strong evolutionary lineage between the two. This is due to the fact that the CACNB2 gene is crucial for the voltage-gated calcium channels that it codes for as well as the signalling response it participates in. This gene is necessary for an immune health and immune response due to the extremely highly conserved sequence and structures across different species. One thing that could open more doors the next time around would be to analyze the direct immune response with a common scenario. This would be helpful because it is a good way of application of this gene and the 1T0J protein it codes for. Doing this would help further understanding of the importance of this gene which sheds light on exactly why it is so conserved.

Bibliography

(1*) *Scientists Compare Rat Genome With Human, Mouse*. Genome.gov. (n.d.). https://www.genome.gov/11511308/2004-release-scientists-compare-rat-genome.

(2*) Masopust, D., Sivula, C. P., & Jameson, S. C. (2017, July 15). *Of Mice, Dirty Mice, and Men: Using Mice To Understand Human Immunology*. Journal of immunology (Baltimore, Md. : 1950). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5512602/.

(3*) Agarwal, Y., Beatty, C., Ho, S., Thurlow, L., Das, A., Kelly, S., Castronova, I., Salunke, R., Biradar, S., Yeshi, T., Richardson, A., & Bility, M. (2020, January 1). *Development of humanized mouse and rat models with full-thickness human skin and autologous immune cells*. bioRxiv. https://www.biorxiv.org/content/10.1101/2020.01.14.906255v1.full.

(4*) U.S. National Library of Medicine. (n.d.). *CACNB2 calcium voltage-gated channel auxiliary subunit beta 2 [Homo sapiens (human)] - Gene - NCBI*. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/gene/783.

(5*) Broecker, F., & Moelling, K. (2019, January 14). *Evolution of Immune Systems From Viruses and Transposable Elements*. Frontiers. https://www.frontiersin.org/articles/10.3389/fmicb.2019.00051/full.

(6*) /* custom css */.tdb-post-meta{ margin-bottom: 16px; color: #444; font-family: 'Open Sans', Editors, B. B. D., By: & Editors, B. D. (2017, April 28). *Divergent Evolution - Definition and Examples*. Biology Dictionary. https://biologydictionary.net/divergent-evolution/.

(7*) Koonin, E. V. (1970, January 1). *Evolutionary Concept in Genetics and Genomics*. Sequence - Evolution - Function: Computational Approaches in Comparative Genomics. https://www.ncbi.nlm.nih.gov/books/NBK20255/.

(8*) Bank, R. C. S. B. P. D. (n.d.). *1T0J: Crystal structure of a complex between voltage-gated calcium channel beta2a subunit and a peptide of the alpha1c subunit*. RCSB PDB. https://www.rcsb.org/structure/1T0J.

(9*) Turner, R. W., Anderson, D., & Zamponi, G. W. (2011). *Signaling complexes of voltage-gated calcium channels*. Channels (Austin, Tex.). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3265765/#:~:text=There%20is%20increasing%20evidence%20that,other%20types%20of%20ion%20channels.