Katelyn Campbell

Molecular Phylogenetics

Final Project Draft: Phylogenetic Analysis of Myosin Proteins

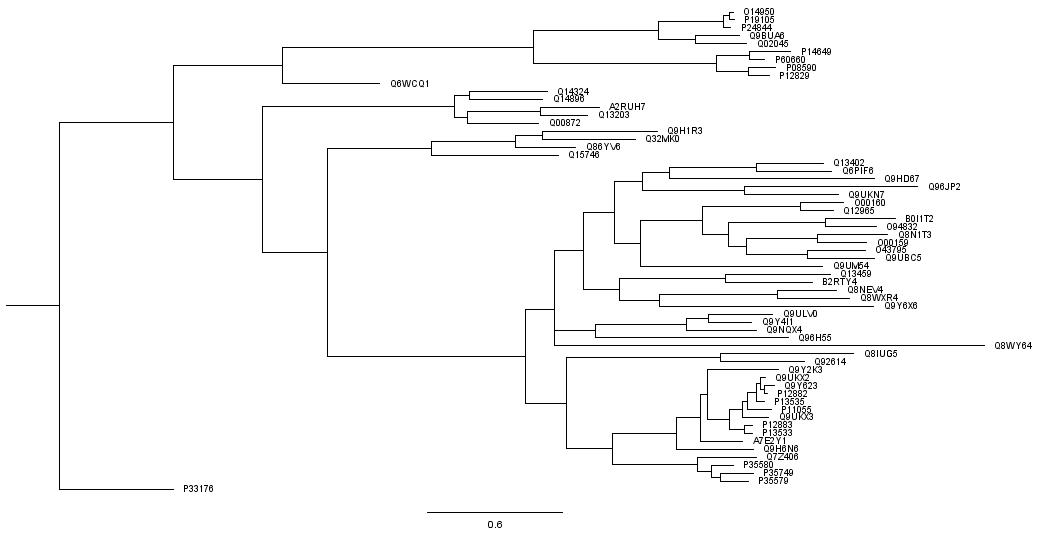
**Introduction**

Myosins, cytoskeletal motor proteins, play a critical role in a wide array of cellular functions within humans. Some pathways that relay on these motor proteins include muscle contraction, cytokinesis, cell signaling, intracellular transport, and organization of actin cytoskeleton1. The general mechanism of myosin entails utilizing energy derived from ATP hydrolysis so that they can “walk” along actin filaments to generate a force1. The myosin protein family consists of 18 classes – 12 in humans – that consist of 38 different myosin-encoding genes2,1. Mutations in myosin proteins are linked to deafness, dilated and hypertrophic cardiomyopathy, Griscelli syndrome, myosin storage myopathy, Usher syndrome, and much more3. The purpose of this project is to look take an evolutionary look into the relation between the different myosin classes by comparing them with Bayesian analysis and seeing if more closely related classes have similar disease phenotype when mutated.

**Methods**

All data used in this analysis was from the UniProtKB – which is a protein knowledgebase. 60 protein sequences for different classes of myosin proteins were used but all isoforms were excluded. The organism, humans, for these protein sequences were held constant and all sequences were reviewed and verified to strengthen the internal validity of the analysis. A human kinesin protein was included to be used as an outgroup to root the consensus tree produced. UniProtKB aligned all the protein sequences and they were exported in a FASTA file. On [online file converter](http://phylogeny.lirmm.fr/phylo_cgi/data_converter.cgi) was used to convert the FASTA file to a NEXUS file. A Bayesian analysis was done via MrBayes and was executed with the Slurm to keep the program running on its own. After MrBayes completed the analysis, a consensus tree was created using the sumt command within the MrBayes program. The consensus tree was then put into the program FigTree to illustrate the tree. Within FigTree, the kinesin outgroup (P33176) was selected and used to root the unrooted tree.

Results

**Figure 1:** Consensus tree of the myosin proteins. This tree displays the most likely tree produced from the Bayesian analysis. The outgroup, P33176, is the kinesin-1 protein which was used to root the tree.

Discussion

References

1. Masters T.A., Kendrick-Jones J., Buss F. (2016) Myosins: Domain Organisation, Motor Properties, Physiological Roles and Cellular Functions. In: Jockusch B. (eds) The Actin Cytoskeleton. Handbook of Experimental Pharmacology, vol 235. Springer, Cham
2. Thompson, Reid F., Langford, George M. (2002) Myosin superfamily evolutionary history. Wiley-Liss, Inc. <https://doi.org/10.1002/ar.10160>
3. Hartman, Amanda M., Spudich, James A. (2012) The myosin superfamily at a glance. Journal of Cell Science. 125(7):1627-1632.

Tony Hodge and M. Jamie T. B. Cope. 1/1/2000. A myosin family tree. Journal of Cell Science. 113(19):3353-3354.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entry** | **Entry name** | **Protein names** | **Gene names** | **Length** |
| P33176 | KINH\_HUMAN | Kinesin-1 heavy chain | KIF5B KNS KNS1 | 963 |
| Q6WCQ1 | MPRIP\_HUMAN | Myosin phosphatase Rho-interacting protein | MPRIP KIAA0864 MRIP RHOIP3 | 1025 |
| Q00872 | MYPC1\_HUMAN | Myosin-binding protein C, slow-type | MYBPC1 MYBPCS | 1141 |
| Q14324 | MYPC2\_HUMAN | Myosin-binding protein C, fast-type | MYBPC2 MYBPCF | 1141 |
| Q14896 | MYPC3\_HUMAN | Myosin-binding protein C, cardiac-type | MYBPC3 | 1274 |
| Q13203 | MYBPH\_HUMAN | Myosin-binding protein H | MYBPH | 477 |
| A2RUH7 | MBPHL\_HUMAN | Myosin-binding protein H-like | MYBPHL | 354 |
| P12882 | MYH1\_HUMAN | Myosin-1 | MYH1 | 1939 |
| P35580 | MYH10\_HUMAN | Myosin-10 | MYH10 | 1976 |
| P35749 | MYH11\_HUMAN | Myosin-11 | MYH11 KIAA0866 | 1972 |
| Q9UKX3 | MYH13\_HUMAN | Myosin-13 | MYH13 | 1938 |
| Q7Z406 | MYH14\_HUMAN | Myosin-14 | MYH14 KIAA2034 FP17425 | 1995 |
| Q9Y2K3 | MYH15\_HUMAN | Myosin-15 | MYH15 KIAA1000 | 1946 |
| Q9H6N6 | MYH16\_HUMAN | Putative uncharacterized protein MYH16 | MYH16 MYH5 | 1097 |
| Q9UKX2 | MYH2\_HUMAN | Myosin-2 | MYH2 MYHSA2 | 1941 |
| P11055 | MYH3\_HUMAN | Myosin-3 | MYH3 | 1940 |
| Q9Y623 | MYH4\_HUMAN | Myosin-4 | MYH4 | 1939 |
| P13533 | MYH6\_HUMAN | Myosin-6 | MYH6 MYHCA | 1939 |
| P12883 | MYH7\_HUMAN | Myosin-7 | MYH7 MYHCB | 1935 |
| A7E2Y1 | MYH7B\_HUMAN | Myosin-7B | MYH7B KIAA1512 | 1983 |
| P13535 | MYH8\_HUMAN | Myosin-8 | MYH8 | 1937 |
| P35579 | MYH9\_HUMAN | Myosin-9 | MYH9 | 1960 |
| Q9BUA6 | MYL10\_HUMAN | Myosin regulatory light chain 10 | MYL10 MYLC2PL PLRLC | 226 |
| P19105 | ML12A\_HUMAN | Myosin regulatory light chain 12A | MYL12A MLCB MRLC3 RLC | 171 |
| O14950 | ML12B\_HUMAN | Myosin regulatory light chain 12B | MYL12B MRLC2 MYLC2B | 172 |
| P08590 | MYL3\_HUMAN | Myosin light chain 3 | MYL3 | 195 |
| P12829 | MYL4\_HUMAN | Myosin light chain 4 | MYL4 MLC1 PRO1957 | 197 |
| Q02045 | MYL5\_HUMAN | Myosin light chain 5 | MYL5 | 173 |
| P60660 | MYL6\_HUMAN | Myosin light polypeptide 6 | MYL6 | 151 |
| P14649 | MYL6B\_HUMAN | Myosin light chain 6B | MYL6B MLC1SA | 208 |
| P24844 | MYL9\_HUMAN | Myosin regulatory light polypeptide 9 | MYL9 MLC2 MRLC1 MYRL2 | 172 |
| Q8WY64 | MYLIP\_HUMAN | E3 ubiquitin-protein ligase MYLIP | MYLIP BZF1 IDOL BM-023 PP5242 | 445 |
| Q15746 | MYLK\_HUMAN | Myosin light chain kinase, smooth muscle | MYLK MLCK MLCK1 MYLK1 | 1914 |
| Q9H1R3 | MYLK2\_HUMAN | Myosin light chain kinase 2, skeletal/cardiac muscle | MYLK2 | 596 |
| Q32MK0 | MYLK3\_HUMAN | Myosin light chain kinase 3 | MYLK3 MLCK | 819 |
| Q86YV6 | MYLK4\_HUMAN | Myosin light chain kinase family member 4 | MYLK4 SGK085 | 388 |
| Q9HD67 | MYO10\_HUMAN | Unconventional myosin-X | MYO10 KIAA0799 | 2058 |
| Q9UKN7 | MYO15\_HUMAN | Unconventional myosin-XV | MYO15A MYO15 | 3530 |
| Q96JP2 | MY15B\_HUMAN | Unconventional myosin-XVB | MYO15B KIAA1783 MYO15BP | 1530 |
| Q9Y6X6 | MYO16\_HUMAN | Unconventional myosin-XVI | MYO16 KIAA0865 MYO16B NYAP3 | 1858 |
| Q92614 | MY18A\_HUMAN | Unconventional myosin-XVIIIa | MYO18A CD245 KIAA0216 MYSPDZ | 2054 |
| Q8IUG5 | MY18B\_HUMAN | Unconventional myosin-XVIIIb | MYO18B | 2567 |
| Q96H55 | MYO19\_HUMAN | Unconventional myosin-XIX | MYO19 MYOHD1 | 970 |
| Q9UBC5 | MYO1A\_HUMAN | Unconventional myosin-Ia | MYO1A MYHL | 1043 |
| O43795 | MYO1B\_HUMAN | Unconventional myosin-Ib | MYO1B | 1136 |
| O00159 | MYO1C\_HUMAN | Unconventional myosin-Ic | MYO1C | 1063 |
| O94832 | MYO1D\_HUMAN | Unconventional myosin-Id | MYO1D KIAA0727 | 1006 |
| Q12965 | MYO1E\_HUMAN | Unconventional myosin-Ie | MYO1E MYO1C | 1108 |
| O00160 | MYO1F\_HUMAN | Unconventional myosin-If | MYO1F | 1098 |
| B0I1T2 | MYO1G\_HUMAN | Unconventional myosin-Ig | MYO1G HA2 | 1018 |
| Q8N1T3 | MYO1H\_HUMAN | Unconventional myosin-Ih | MYO1H | 1032 |
| Q8NEV4 | MYO3A\_HUMAN | Myosin-IIIa | MYO3A | 1616 |
| Q8WXR4 | MYO3B\_HUMAN | Myosin-IIIb | MYO3B | 1341 |
| Q9Y4I1 | MYO5A\_HUMAN | Unconventional myosin-Va | MYO5A MYH12 | 1855 |
| Q9ULV0 | MYO5B\_HUMAN | Unconventional myosin-Vb | MYO5B KIAA1119 | 1848 |
| Q9NQX4 | MYO5C\_HUMAN | Unconventional myosin-Vc | MYO5C | 1742 |
| Q9UM54 | MYO6\_HUMAN | Unconventional myosin-VI | MYO6 KIAA0389 | 1294 |
| Q13402 | MYO7A\_HUMAN | Unconventional myosin-VIIa | MYO7A USH1B | 2215 |
| Q6PIF6 | MYO7B\_HUMAN | Unconventional myosin-VIIb | MYO7B | 2116 |
| B2RTY4 | MYO9A\_HUMAN | Unconventional myosin-IXa | MYO9A MYR7 | 2548 |
| Q13459 | MYO9B\_HUMAN | Unconventional myosin-IXb | MYO9B MYR5 | 2157 |

**Table 1:** Table displays the entry number (same identifier used in the analysis), entry name, protein name, gene name, and the length of proteins analyzed. There are 60 myosin proteins, and one kinesin protein. The colors in the table indicate different groups within the proteins. Green = outgroup. Purple = binding proteins. Orange = heavy chain. Blue = light chain. Yellow = mutated proteins. White = proteins that don’t clearly fit into other groups.