**Matthew Morassutti**

**BIOL550 SPR’24 Week 6 Project Proposal**

**Draft Title:**

Resolving the Ranges of Two Species of Arborimus

**Question and Relevance:**

Red tree voles (*Arborimus longicaudus*) and Sonoma tree voles (*A. pomo*) are arboreal rodents endemic to the Pacific Northwest and Northern California. They are unique to other wildlife in this region, as they spend almost all of their life in the canopy of coniferous trees, feeding exclusively on their leaves (Taylor 1915).This project will clarify the ranges of two species of tree voles, red tree voles and Sonoma tree voles, that have previously been inferred though not directly assessed. This research aims to directly assess the northern range limit of the Sonoma tree vole and the southern range limit of the red tree vole by conducting phylogenomic analysis of ultraconserved elements (UCEs) of the nuclear genome and the entire mitochondrial DNA (mtDNA) genome. The primary source of molecular samples will be existing tissue collections and museum study skins, with some potential augmentation with samples extracted from feces collected from tree vole nests in trees north and south of the Klamath River. With these data I will resolve previous inferences of the range delimitation of both species most recently explored by Blois and Arbogast (2006). In this phylogenomic analysis, the flanking regions of UCEs and mtDNA will be examined to determine if these two species form reciprocally monophyletic clades, and if not, if this is the result of incomplete lineage sorting and/or recent introgression, and lastly to determine the range limits of these two species (McCormack et al. 2012, McLean et al. 2019, Zhang et al. 2019).Because *A. pomo* is a Species of Special concern in California (CNDDB 2023), this study is important to understand the taxonomic relationship at the transition zone between *A. pomo* and *A. longicaudus*. If their populations have the potential to be strongly impacted by invasive predators, it will be useful to get a better understanding of gene flow between seemingly isolated species and populations. If these species are able to navigate a purported barrier as wide as the Klamath River, this could be promising to the conservation of these species’ genetic diversity. Furthermore, by elucidating the complexity of the taxonomy of these species, resource managers may be able to advocate for their conservation with greater salience.

**Focal Organism Introduction:**

*A. longicaudus* and *A. pomo* display a nearly completely arboreal life history (Howell 1926). Terrestrial activity appears only to occur when adjacent trees are not connected by branch pathways (Swingle 2009). Both species build nests in the canopies of trees, between two and 65 meters in elevation above the forest floor (Biswell et al. 2000). Nests are generally spherical with a flat roof, contain one or more inner chambers, and display multiple entrance and escape tunnels (Howell 1926). The nests of males and subadult females are usually smaller than that of adult females, which have larger nests to include space for their offspring (Howell 1926, Taylor 1915). Nests may be as large as one meter in diameter and the result of several generations of tree vole activity (Adam and Hayes 1998). Nests are built against the boles of trees, in hollows or cavities, on forked trunks or broken treetops, or on closely spaced limb whorls of the forest canopy. Tree voles have also been documented using the dense branching structures of dwarf mistletoe (*Arceuthobium sp.*) brooms as nesting substrate. In older trees with larger branches, tree voles may construct nests away from the bole of the tree, and nests have been observed in the outer crown of trees (Thompson and Diller 2002, Nicholas Kerhoulas, pers. comm.). Tree voles typically nest in Douglas-fir trees but may also be found in grand fir (*Abies grandis*), Sitka spruce (*Picea sitchensis*), Western hemlock (*Tsuga heterophylla*), coast redwood (*Sequoia sempervirens*), and Pacific madrone (*Arbutus menziesii*), among others (Zentner 1977, Vrieze 1980, Meiselman and Doyle 1996, Nicholas Kerhoulas pers. comm.). Because many of the aforementioned canopy structures used by tree voles are associated with mature trees, tree voles appear to favor stands of trees older than 80 years because of the more suitable nesting structures supported by a complex and mature canopy, though they may also be found in young stands (<80 years) (Huff et al. 1992, Meiselman and Doyle 1996, Lesmeister and Swingle 2017).

**Methods Overview**

Genomic data will be extracted from dried museum study skins (n = 8) using QIAGEN DNeasy Blood and Tissue (QIAGEN, Valencia CA) DNA extraction kit following the manufacturer’s protocol and the DNA extraction protocol outlined in Kerhoulas et al. 2010. Genomic data will also be sourced from preserved tissue samples and/or existing extractions (n = 24, Murray 1995, Blois and Arbogast 2006).

To determine if there is evidence of introgression among the genomes of A. longicaudus and A. pomo, ultraconserved elements (UCEs) will be used to resolve the phylogenetic history of the genus Arborimus (Faircloth et al. 2012, Meiklejohn et al. 2016). These are sections of the genome that remain constant across many species, in this case all tetrapods (Bejerano et al 2004, Siepel et al. 2005, Stephen et al 2008, Janes et al. 2011). Because of this, these sections will be targeted for sequencing, and assess the variation and similarities among the flanking regions of UCEs across the genome to give insight to phylogenetic patterns across the ranges of A. longicaudus and A. pomo (Mills et al. 2023). This analysis also recovers the entire mtDNA genome of each individual. Currently there is no full nuclear genome for either species available in GenBank, so this research will contribute directly to the knowledge of the genetic diversity within each species.

UCE sequencing will be outsourced to Arbor Biosciences (Ann Arbor, Michigan) and their MyReads NGS services, including library preparation, target enrichment with myBaits UCE Tetrapods 5kv1 probe set (http://ultraconserved.org), pooling, and sequencing will be used. All indexed reads will be demultiplexed with the Arbor Biosciences Standard DNA (8-plex capture) and Degraded DNA (4-plex capture) packages.

Phylogenetic analysis will follow methods outlined in Mills et al. 2023. In phyluce software, the v.1.6.2 bioinformatics pipeline will be used. Raw UCE reads will be trimmed using Illumiprocessor (Faircloth 2013), a wrapper for Trimmomatic v.036 (Bolger et al. 2014) included in the phyluce pipeline (Faircloth 2015). Loci will be assembled in Trinity v1 (Grabherr et al. 2011) using the phyluce script assemblo\_trinity. Assemblies will be matched to the UCE probes, assemblies will be extracted, and individual UCE loci will be aligned in MAFFT (Katoh and Standley 2013), and edges will be trimmed with GBlocks (Castresama 2000). UCE flanks will be conservatively trimmed to remove poorly represented nucleotide positions within core UCEs, and assembly errors will be removed (McLean et al. 2019).

To assess for presence of genetic introgression, the loci of the UCEs from each genome in the dataset will be partitioned using the entropy-based method SWSC-en to isolate the core and both flanks of each UCE (Tagliacollo and Lanfear 2018). Further consolidation of each partition will be run through PartitionFinder2, which also will identify the best substitution model for each partition based on AICc score (Lanfear et al. 2017). Maximum-likelihood phylogenies will be estimated using these consolidated partitions using IQ\_TREE (Nguyen et al. 2015), the ultrafast bootstrap method (UFboot; Hoang et. al. 2018), and the approximate likelihood ratio test (SH-aLRT; Anisimova et al. 2011) with 1,000 replicates.

Phylogenies will be estimated using a coalescent-based method, as many UCEs contain a weak phylogenetic signal. UCE loci will be sorted using the phyluce\_align\_get\_informative\_sites script and the dataset will be reduced to the top quartile of loci. A maximum-likelihood phylogeny will be constructed using IQ-TREE with branch supports obtained from UFboot and SH-aLRT and all branches with less than 10% bootstrap support will be collapsed (Mills et al. 2023, Zhang et al. 2019). Phylogenies will be used to estimate species using ASTRAL-III (Rabiee et al. 2019). If case species do not form monophyletic groups, individuals will be left as operational taxonomic units (OTUs).

**Works Cited:**

Adam, M. D., Hayes, J. P. 1998. *Arborimus pomo*, *Mammalian Species*, 593, 1–5.

Anisimova M., Gil M., Dufayard J. F., Dessimoz C., Gascuel O. 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Systematic Biology*, 60, 685–699.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W. J., Mattick, J. S., Haussler, D. 2004. Ultraconserved elements in the human genome. *Science*, 304, 5675, 1321–1325.

Biswell, B., Blow, M., Finley, L., Madsen, S., Schmidt, K. 2000. Survey protocol for the red tree vole, A*rborimus longicaudus* (=*Phenacomys longicaudus* in the Record of Decision of the Northwest Forest Plan). Version 2.0. Portland, OR: U.S. Department of Agriculture, Forest Service and U.S. Department of the Interior, Bureau of Land Management. 32.

Blois, J. L., Arbogast, B. S. 2006. Conservation Genetics of the Sonoma Tree Vole (*Arborimus pomo*) Based on Mitochondrial and Amplified Fragment Length Polymorphism Markers. *Journal of Mammalogy*, 87, 950–960.

California Natural Diversity Database (CNDDB). April 2023. Special Animals List. California Department of Fish and Wildlife. Sacramento, CA.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17, 540–552.

Faircloth, B. C. 2013. Illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. <http://dx.doi.org/10.6079/J9ILL>.

Faircloth, B. C. 2015. PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, 32, 786–788.

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., Glenn, T. C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, 61, 717–726.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29, 644-652.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., Vinh, L. S., 2017. UFBoot2: Improving the ultrafast bootstrap approximation. *BioRxiv*, 35, 518–522.

Howell, A. B. 1926. Voles of the genus *Phenacomys*, II: Life history of the red tree mouse *Phenacomys*. *North American Fauna*, 48, 39–66.

Huff, M. H., Holthausen, R. S., Aubry, K.B. 1992.Habitat management for red tree voles in Douglas-fir forests. General Technical Report. PNW-GTR-302. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 1–16.

Janes, D. E., Chapus, C., Gondo, Y., Clayton, D. F., Sinha, S., Blatti, C. A., Organ, C. L., Fujita, M. K., Balakrishnan, C. N., Edwards, S. V. 2011. Reptiles and mammals have differentially retained long conserved noncoding sequences from the amniote ancestor. *Genome Biology and Evolution*, 3, 102–113.

Katoh, K., Standley, D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780.

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., Calcott, B. 2017. Partitionfinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, 34, 772–773.

Lesmeister, D. B., Swingle, J. K. 2017. Field Guide to Red Tree Vole Nests. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

McCormack, J. E., Faircloth, B. C., Crawford, N. G., Gowaty, P. A., Brumfield, R. T., Glenn, T. C. 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome Research*, 22, 4, 746–754.

McLean, B. S., Bell, K. C., Allen, J. M., Helgen, K. M., Cook, J. A. 2019. Impacts of Inference Method and Data set Filtering on Phylogenomic Resolution in a Rapid Radiation of Ground Squirrels (*Xerinae*: *Marmotini*). *Systematic Biology*, 68, 298–316.

Meiklejohn, K. A., Faircloth, B. C., Glenn, T. C., Kimball, R. T., Braun, E .L., 2016. Analysis of a rapid evolutionary radiation using ultraconserved elements: Evidence for a bias in some multispecies coalescent methods.*Systematic Biology*, 65, 612–627.

Meiselman, N., Doyle, A. T. 1996. Habitat and microhabitat use by the red tree vole (*Phenacomys longicaudus*). *The American Midland Naturalist*, 135, 33–42.

Mills, K. K., Everson, K. M., Hildebrandt, K. P. B., Brandler, O. V., Steppan, S. J., Olson, L. E. 2023. Ultraconserved elements improve resolution of marmot phylogeny and offer insights into biogeographic history. *Molecular Phylogenetics and Evolution*, 107785.

Murray M. A. 1995. Biochemical systematics of the genus *Arborimus*. M.A. thesis, Humboldt State University. Arcata, California.

Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., Minh, B. Q. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular *Biology and Evolution*, 32, 1, 268–274.

Rabiee, M., Sayyari, E., Mirarab, S. 2019. Multi-allele species reconstruction using ASTRAL. *Molecular Phylogenetics and Evolution*, 130, 286–96.

Siepel, A., Bejerano, G., Pedersen, J. S., Hinrichs, A. S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L. W., Richards, S., Weinstock, G. M., Wilson, R. K., Gibbs, R. A., Kent, W. J., Miller, W., Haussler, D. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Research*, 15, 8, 1034–1050.

Stephen, S., Pheasant, M., Makunin, I. V., Mattick, J. S. 2008. Large-scale appearance of ultraconserved elements in tetrapod genomes and slowdown of the molecular clock. *Molecular Biology and Evolution*, 25, 2, 402–408.

Swingle, J. 2009. Home range areas and activity patterns of red tree voles (*Arborimus longicaudus*) in Western Oregon. *Northwest Science*, 83, 273–286.

Tagliacollo, V. A., Lanfear, R. 2018. Estimating improved partitioning schemes for ultraconserved elements. *Molecular Biology and Evolution*, 35, 1798–1811.

Taylor, W. P. 1915. Description of a new subgenus (*Arborimus*) of *Phenacomys*, with a contribution to the knowledge and distribution of *Phenacomys longicaudus*. *Proceedings of the California Academy of Sciences*, 5, 111–161.

Thompson, J. L., Diller, L. V. 2002. Relative abundance, nest site characteristics, and nest dynamics of sonoma tree voles on managed timberlands in Coastal Northwest California. *Northwestern Naturalist*, 83, 91.

Vrieze, J. M. 1980.Spatial patterning of red tree mouse, *Arborimus longicaudus*, nests. M.A. thesis, Humboldt State University. Arcata, California.

 Zentner, P. L. 1977.The nest of *Phenacomys longicaudus* in northwestern California. M.A. thesis, California State University. Sacramento, California.

Zhang, Y. M., Williams, J. L., Lucky, A. 2019. Understanding UCEs: A comprehensive primer on using ultraconserved elements for arthropod phylogenomics. *Insect Systematics and Diversity*, 3, 1–3.