

# CSC 525 Course Project

## An Annotated Bibliography

Derek Robinson (drobinson@uvic.a)  
University of Victoria

October 19, 2021

This document gives the citation, title, and abstract for references related to the CSC 525 Course Project. It is split into two different sections: core references and possibly related references.

### 1 Core References

**Title:** Long noncoding RNAs and Alzheimers disease

**Bibtex Key:** luo2016long

[1]

Long noncoding RNAs (lncRNAs) are typically defined as transcripts longer than 200 nucleotides. lncRNAs can regulate gene expression at epigenetic, transcriptional, and posttranscriptional levels. Recent studies have shown that lncRNAs are involved in many neurological diseases such as epilepsy, neurodegenerative conditions, and genetic disorders. Alzheimers disease is a neurodegenerative disease, which accounts for .80elderly subjects. In this review, we will highlight recent studies investigating the role of lncRNAs in Alzheimers disease and focus on some specific lncRNAs that may underlie Alzheimers disease pathophysiology and therefore could be potential therapeutic targets.

**Title:** Phylogenetic Analysis of Three Long Non-coding RNA Genes: AK082072, AK043754 and AK082467

**Bibtex Key:** amirmahani2018phylogenetic

[2]

Now, it is clear that protein is just one of the most functional products produced by the eukaryotic genome. Indeed, a major part of the human genome is transcribed to non-coding sequences than to the coding sequence of the protein. In this study, we selected three long non-coding RNAs namely AK082072, AK043754 and AK082467 which show brain expression and local region conservation among vertebrates. Thus, the sequences of these genes are appropriate for phylogenetic analysis. In order to evaluate the evolutionary and molecular trend of lncRNAs in vertebrates, phylogenetic analysis and natural selection process were analyzed during evolution. The nucleotide sequences of selected long non-coding RNAs from different vertebrates were aligned and the phylogenetic trees were constructed using Neighbor Joining method with maximum sequence differences of 0.75. Our analysis of nucleotide sequences to find closely evolved organisms with high similarity by NCBI-BLAST tools and MEGA7 showed that the selected sequence of AK082072 in human and *M. fascicularis* (macaque) were placed into the same cluster and they may originate from a common ancestor. In addition, the human sequence of AK082467 and AK043754 had the closest similarity with cow. Also, bioinformatic analysis showed that the dN/dS ratio is lower than 1 for all three genes which demonstrates purifying selection for the longest predicted ORF of each lncRNA. Together, these results indicate that lncRNAs act as regulatory genes that have important roles in development.

**Title:** MEGA11: Molecular Evolutionary Genetics Analysis Version 11

**Bibtex Key:** tamura2021mega11

[3]

The Molecular Evolutionary Genetics Analysis (MEGA) software has matured to contain a large collection of methods and tools of computational molecular evolution. Here, we describe new additions that make MEGA a more comprehensive tool for building timetrees of species, pathogens, and gene families using rapid relaxed-clock methods. Methods for estimating divergence times and confidence intervals are implemented to use probability densities for calibration constraints for node-dating and sequence sampling dates for tip-dating analyses. They are supported by new options for tagging sequences with spatiotemporal sampling information, an expanded interactive Node Calibrations Editor, and an extended Tree Explorer to display timetrees. Also added is a Bayesian method for estimating neutral evolutionary probabilities of alleles in a species using multispecies sequence alignments and a machine learning method to test for the autocorrelation of evolutionary rates in phylogenies. The computer memory requirements for

the maximum likelihood analysis are reduced significantly through reprogramming, and the graphical user interface has been made more responsive and interactive for very big data sets. These enhancements will improve the user experience, quality of results, and the pace of biological discovery. Natively compiled graphical user interface and command-line versions of MEGA11 are available for Microsoft Windows, Linux, and macOS from [www.megasoftware.net](http://www.megasoftware.net).

**Title:** Building Phylogenetic Trees from Molecular Data with MEGA

**Bibtex Key:** hall2013building

[4]

Phylogenetic analysis is sometimes regarded as being an intimidating, complex process that requires expertise and years of experience. In fact, it is a fairly straightforward process that can be learned quickly and applied effectively. This Protocol describes the several steps required to produce a phylogenetic tree from molecular data for novices. In the example illustrated here, the program MEGA is used to implement all those steps, thereby eliminating the need to learn several programs, and to deal with multiple file formats from one step to another (Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 28:27312739). The first step, identification of a set of homologous sequences and downloading those sequences, is implemented by MEGA's own browser built on top of the Google Chrome toolkit. For the second step, alignment of those sequences, MEGA offers two different algorithms: ClustalW and MUSCLE. For the third step, construction of a phylogenetic tree from the aligned sequences, MEGA offers many different methods. Here we illustrate the maximum likelihood method, beginning with MEGA's Models feature, which permits selecting the most suitable substitution model. Finally, MEGA provides a powerful and flexible interface for the final step, actually drawing the tree for publication. Here a step-by-step protocol is presented in sufficient detail to allow a novice to start with a sequence of interest and to build a publication-quality tree illustrating the evolution of an appropriate set of homologs of that sequence. MEGA is available for use on PCs and Macs from [www.megasoftware.net](http://www.megasoftware.net).

## 2 Possibly Related

**Title:** Asynchronous Evolutionary Origins of A and BACE1

**Bibtex Key:** moore2013asynchronous

[5]

Neurodegenerative plaques characteristic of Alzheimers disease (AD) are composed of amyloid beta (A) peptide, which is proteolyzed from amyloid precursor protein (APP) by  $\gamma$ -secretase (beta-site APP cleaving enzyme [BACE1]) and  $\alpha$ -secretase. Although  $\gamma$ -secretase has essential functions across metazoans, no essential roles have been identified for BACE1 or A. Because their only known function results in a disease phenotype, we sought to understand these components from an evolutionary perspective. We show that APP-like proteins are found throughout most animal taxa, but sequences homologous to A are not found outside gnathostomes and the cut site is only conserved within sarcopterygians. BACE1 enzymes, however, extend through basal chordates and as far as cnidaria. We then sought to determine whether BACE1 from a species that never evolved A could proteolyze APP substrates that include A. We demonstrate that BACE1 from a basal chordate is a functional ortholog that can liberate A from full-length human APP, indicating BACE1 activity evolved at least 360 My before A.

**Title:** RNA Dynamics in Alzheimers Disease

**Bibtex Key:** rybak2021rna

[6]

Alzheimers disease (AD) is the most common age-related neurodegenerative disorder that heavily burdens healthcare systems worldwide. There is a significant requirement to understand the still unknown molecular mechanisms underlying AD. Current evidence shows that two of the major features of AD are transcriptome dysregulation and altered function of RNA binding proteins (RBPs), both of which lead to changes in the expression of different RNA species, including microRNAs (miRNAs), circular RNAs (circRNAs), long non-coding RNAs (lncRNAs), and messenger RNAs (mRNAs). In this review, we will conduct a comprehensive overview of how RNA dynamics are altered in AD and how this leads to the differential expression of both short and long RNA species. We will describe how RBP expression and function are altered in AD and how this impacts the expression of different RNA species. Furthermore, we will also show how changes in the abundance of specific RNA species are linked to the pathology of AD.

**Title:** Identification of the biological affection of long noncoding RNA BC200 in Alzheimers disease

**Bibtex Key:** li2018identification

[7]

BC200 is a long noncoding RNA expressed at high levels in the Alzheimers disease (AD), and blocking of BC200 by siRNA is assumed to be an effective method for various disease therapy. We have established an AD cell model overexpressing amyloid -peptide (A)1-42 to observe the effects of BC200 on the cell viability and apoptosis, and to investigate the associated underlying mechanisms. Efficient knockdown and overexpression of BC200 were established using BC200 siRNA and BC200 mimics, respectively. Cell viability following BC200 knockdown and overexpression was assessed by 3-(4, 5-dimethyl-2-thiazolyl)-2, 5-diphenyltetrazolium bromide assay, and cell apoptosis was monitored by flow cytometry. We successfully established an AD cell model overexpressing A1-42 gene, and reported the results of change of BC200 on A1-42 levels. Knockdown of BC200 significantly suppressed b-site amyloid precursor protein-cleaving enzyme 1 (BACE1) expression, and overexpression of BC200 increased BACE1 expression. Besides, inhibition of BC200 significantly increased cell viability and reduced cell apoptosis in the AD model via directly targeting BACE1, which can be increased by overexpression of BC200. BC200 regulated AD cell viability and apoptosis via targeting BACE1, and it may be one of the putative target in AD development and provides potential new insights into genetic therapy against AD.

## References

- [1] Q. Luo and Y. Chen, “Long noncoding rnas and alzheimers disease,” *Clinical interventions in aging*, vol. 11, p. 867, 2016.
- [2] F. Amirmahani and K. Jamshidi Goharrizi, “Phylogenetic analysis of three long non-coding rna genes: Ak082072, ak043754 and ak082467,” *Journal of Genetic Resources*, vol. 4, no. 1, pp. 56–64, 2018.
- [3] K. Tamura, G. Stecher, and S. Kumar, “Mega11: molecular evolutionary genetics analysis version 11,” *Molecular Biology and Evolution*, vol. 38, no. 7, pp. 3022–3027, 2021.

- [4] B. G. Hall, “Building phylogenetic trees from molecular data with mega,” *Molecular biology and evolution*, vol. 30, no. 5, pp. 1229–1235, 2013.
- [5] D. B. Moore, M. A. Gillentine, N. M. Botezatu, K. A. Wilson, A. E. Benson, and J. A. Langeland, “Asynchronous evolutionary origins of  $a\beta$  and bace1,” *Molecular biology and evolution*, vol. 31, no. 3, pp. 696–702, 2013.
- [6] A. Rybak-Wolf and M. Plass, “Rna dynamics in alzheimers disease,” *Molecules*, vol. 26, no. 17, p. 5113, 2021.
- [7] H. Li, L. Zheng, A. Jiang, Y. Mo, and Q. Gong, “Identification of the biological affection of long noncoding rna bc200 in alzheimers disease,” *Neuroreport*, vol. 29, no. 13, pp. 1061–1067, 2018.