Michael Boodoo

Genomics Final Project Bibliography

Project Idea: I am interested in performing some sort of statistical analysis on my gene (SCN5A) and it's orthologs to see how it has genetically drifted over time. My current idea is to use the statistical programming language R to perform a multi-dimensional scaling (MDS) of SCN5A with a large number of orthologs and create a 2-dimensional scatterplot; from that data I could then see and write about what inferences come about from it in terms of evolution. My hypothesis is that since sodium voltage gated channels in heart tissue are critical in survival, that the overall genetic drift between orthologs is low. The level of drift could be quantitatively described using tdN/dS ratios amongst other statistical tools.

References

1) Pelé et al.: Bios2mds: an R package for comparing
orthologous protein families by metric multidimensional scaling. BMC
Bioinformatics 2012 13:133.

2) Schliep KP: Phangorn: phylogenetic analysis in R. Bioinformatics 2011,
27:592–593

3) Pelé J, Abdi H, Moreau M, Thybert D, Chabbert M (2011) Multidimensional Scaling Reveals the Main Evolutionary Pathways of Class A G-Protein-Coupled Receptors. PLoS ONE 6(4): e19094. doi: 10.1371/journal.pone.0019094

4) Dayhoff MO, Schwartz RM, Orcutt BC: A model of evolutionary change in proteins. Atlas of Protein Sequence and Structure 1978, 5:345–352

5) Liu, K., Hipkens, S., Yang, T., Abraham, R., Zhang, W., Chopra, N., Knollmann, B., Magnuson, M. A. and Roden, D. M. (2006), Recombinase-mediated cassette exchange to rapidly and efficiently generate mice with human cardiac sodium channels. Genesis, 44: 556–564. doi: 10.1002/dvg.20247

6) Woolley KJ, Athalye M (1986) A use for principal coordinate analysis in the comparison of protein sequences. Biochem Biophys Res Commun 140: 808–813.

7) Higgins DG (1992) Sequence ordinations: a multivariate analysis approach to analysing large sequence data sets. Comput Appl Biosci 8: 15–22.

8) Studer RA, Robinson-Rechavi M (2010) Large-scale analysis of orthologs and

paralogs under covarion-like and constant-but-different models of amino acid evolution. Mol Biol Evol 27: 2618–2627.

9) Gillespie JH. Variability of evolutionary rates of DNA. Genetics. 1986;113:1077–1091.

10) Allman, E.S.; Rhodes, J.A., "The Identifiability of Covarion Models in Phylogenetics," *Computational Biology and Bioinformatics, IEEE/ACM Transactions on* , vol.6, no.1, pp.76,88, Jan.-March 2009
doi: 10.1109/TCBB.2008.52

11) Sayeed MZ, Salam MA, Haque MZ, Islam AK. Brugada syndrome with a novel missense mutation in SCN5A gene: A case report from Bangladesh. Indian Heart J. 2014;66(1):104-7.

12) Glengarry JM, Crawford J, Morrow PL, Stables SR, Love DR, Skinner JR. Long QT molecular autopsy in sudden infant death syndrome. Arch Dis Child. 2014;

13) Juang JM, Lu TP, Lai LC, et al. Utilizing multiple in silico analyses to identify putative causal SCN5A variants in Brugada syndrome. Sci Rep. 2014;4:3850.

14) Schiffman. S. S.. Reynolds, M. L., and Young. F. W. (1981). *Introduction to Multidimensional Scaling.* Academic Press, New York.

15) Penny D, Mccomish BJ, Charleston MA, Hendy MD. Mathematical elegance with biochemical realism: the covarion model of molecular evolution. J Mol Evol. 2001;53(6):711-23.

16) Hout MC, Papesh MH, Goldinger SD. Multidimensional scaling. Wiley Interdiscip Rev Cogn Sci. 2013;4(1):93-103.

17) Tzeng J, Lu HH, Li WH. Multidimensional scaling for large genomic data sets. BMC Bioinformatics. 2008;9:179.

18) Jaworska N, Chupetlovska-Anastasova A. A review of multidimensional scaling (MDS) and its utility in various psychological domains. Tutorials in Quantitative Methods for Psychology. 2009;5:1–10.

19) Wickelmaier, Florian. *An introduction to MDS.* Aalborg Universitetsforlag. 2003. (Aalborg Universitetscenter. Institut for Elektroniske Systemer. Afdeling for Kommunikationsteknologi. Rapport; Journal number R00-6003).

20) Jacoby, William. *Multidimensional Scaling: An Introduction.* 2012. Michigan State University and ICPSR. http://polisci.msu.edu/jacoby/iu/mds2012/