

Cumming School of Medicine



Evolution of variable mutative organisms Summary of Literature

Oliver A.S. Lyon
2021

Supervisor: Dr. Jason de Koning

1 Project Outline

A project proposal may or may not appear here. It depends on how private the material should be.

2 References Discussion

This section is intended for me to write notes connecting the material in the references with observations and comments.

References

- [1] ; on behalf of the ACMG Laboratory Quality Assurance Committee, S. Richards, N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W. W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, and H. L. Rehm, "Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology," *Genetics in Medicine*, vol. 17, no. 5, pp. 405–423, may 2015. [Online]. Available: <http://www.nature.com/articles/gim201530>

This is the guidelines for assessing the pathogenicity of a genetic variant by the ACMG. These guidelines are fairly exustive but are currently a bit out of date in my opinion with regard to in silico methods. The go

through an exhaustive list of impacts. These impacts are ranked as supporting, moderate, strong, and very strong when evaluating pathogeneticity and stand alone, strong, and supporting with regard to benign evidence. To list the impacts and evidence is not useful, but there are some select quotes highlighted in the text about population genetics and the use of in silico solutions.

- [2] S. Balsamo and A. Marin, “Separable solutions for Markov processes in random environments,” *European Journal of Operational Research*, vol. 229, no. 2, pp. 391–403, sep 2013. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0377221713002191>

NOT CLEAR YET.

- [3] A. J. de Koning and B. D. Sanctis, “The rate of molecular evolution when mutation may not be weak,” *bioRxiv*, 2018.

This paper is quite dense. It is well written and covers a large number of topics. These mostly concern the topic of looking at how we define evolution, and how this definition is influenced by the assumptions from neutral theory. This paper does derive a new equations free of assumptions besides the ones used to derive inputs. It then demonstrates the discrepancies between the models.

- [4] M. M. Desai, D. S. Fisher, and A. W. Murray, “The speed of evolution and maintenance of variation in asexual populations,” *Current Biology*, vol. 17, no. 5, pp. 385–394, 2007. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0960982207009840>

This paper reviews CI topics and provides a new model which allows for full CI. Meaning the new clones are allowed to appear in the population and an arbitrary number of them could be competing at a given time. The trade off for this generalizability is the assumptions that all mutants have the same fitness/ selective properties.

- [5] E. Domingo and C. Perales, “Viral quasispecies,” *PLOS Genetics*, vol. 15, no. 10, p. e1008271, oct 2019. [Online]. Available: <https://dx.plos.org/10.1371/journal.pgen.1008271>

A comprehensive review of support for Quasi-Species theory. This paper covers the historical basis and gives a comprehensive coverage of evidence for viral QS. It also discusses the following topics, Error-prone replication and mutant spectra, phenotypic reservoir, adaptive parameters, intra-mutant spectrum quasispecies memory, sequence space, population bottle neck, biological constraints, connections with viral pathogenesis, and QS and long term evolution.

- [6] J. Frazer, P. Notin, M. Dias, A. Gomez, J. K. Min, K. Brock, Y. Gal, and D. S. Marks, “Disease variant prediction with deep generative models of evolutionary

data,” *Nature*, vol. 599, no. 7883, pp. 91–95, nov 2021. [Online]. Available: <https://www.nature.com/articles/s41586-021-04043-8>

This paper covers a neural net work that uses the amino acid sequence of differnt protiens to produce a evolutionary index. The evolutionary index is then used to produce a distribution of of differnt sequences and make a prediction of Benign, unknowen, and Pathogenic: The distribution allows for an estimate of the confidence of the prediction. They used unsupervised learning and were able to achive an avg AUC = 0.92

- [7] P. Gerrish and R. Lenski, “The fate of competing beneficial mutations in an asexual population,” *Genetica*, vol. 102-103, pp. 127–44, 03 1998.

Added notes

- [8] J. Haigh, “The accumulation of deleterious genes in a population muller’s ratchet,” *Theoretical Population Biology*, vol. 14, no. 2, pp. 251–267, 1978. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0040580978900278>

This is a classic paper. Beautiful old math paper with wonky old math notation. Several closed equations are derived to show that positive mutations must appeare in the populations since the pop would go to exinction with out it. The models produced in this paper prodminently show MULLer’s Ratchet in full effect, and produce sumary satistics since computing resorces were scarese at the time.

- [9] P. G. Harrison, “Reversed processes, product forms and a non-product form,” *Linear Algebra and its Applications*, vol. 386, pp. 359–381, jul 2004. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0024379504001089>

NOT CLEAR YET.

- [10] E. C. Holmes and A. Moya, “Is the Quasispecies Concept Relevant to RNA Viruses?” *Journal of Virology*, vol. 76, no. 1, pp. 460–462, jan 2002. [Online]. Available: <https://journals.asm.org/doi/10.1128/JVI.76.1.460-462.2002>

This is a letter to the editor, discussing the issues of applying quasi-species to viral RNA. It starts by critisizing the lack of formal definitions of QS. This letter addresses, critisims (lack of def, lack of drift), and evidence (strong theoretic basis?) it makes arguments but many seem flimsy now.

- [11] M. A. Huynen, P. F. Stadler, and W. Fontana, “Smoothness within ruggedness: the role of neutrality in adaptation,” *Proceedings of the National Academy of Sciences*, vol. 93, no. 1, pp. 397–401, 1996. [Online]. Available: <https://www.pnas.org/content/93/1/397>

This paper is fascianting. It used a simulation to show the importance of nutral mutation on secondary structure evolution. This was achived by simulationg mtation using a diffution model with imperically derived hyper params. This paper then shows that nutral mutation allows for

neutral movement to new wells of fitness, that are otherwise unreachable if selective forces are too strong. Worth Re-reading. :)

- [12] M. Kimura, “Evolutionary Rate at the Molecular Level,” *Nature*, vol. 217, no. 5129, pp. 624–626, feb 1968. [Online]. Available: <https://www.nature.com/articles/217624a0>

A paper on the early scuttle butt of how to compute mutation rates. It draws its own conclusion, which has been since refuted but has realted neutral theory with molecular evolution which is molecular clock theory.

- [13] I. Krukov and A. J. de Koning, “Haldane’s probability of mutant survival is not the probability of allele establishment,” *bioRxiv*, 2019. [Online]. Available: <https://www.biorxiv.org/content/early/2019/07/16/704577>

This paper reinterpretes the idea of establishment from the inital notes of Haldane. The idea that establishment is the point at which a population is no longer suseptabl to drift was the orginial and provailing theory for establishment. This paper recontextualizes it as a population frequency where the odds ratio between going to fixation and exinction are K. They derive many nice formaula and show examples.

- [14] S. Matuszewski, L. Ormond, C. Bank, and J. D. Jensen, “Two sides of the same coin: A population genetics perspective on lethal mutagenesis and mutational meltdown ,” *Virus Evolution*, vol. 3, no. 1, 03 2017. [Online]. Available: <https://doi.org/10.1093/ve/vex004>

An exellent review paper. It goes through and cites many of the ideas and evidence for and against many of them. This paper also makes some unfair assumptions in my opinion. Mainly that if mean fitness is less than 1, the population will go extinct, this seems like a gross over assumption. The population will certainly decrease, but depending of the facorts which drive the change in fitness must be considered when making such a statement.

- [15] A. Moya, E. C. Holmes, and G.-C. Fernando, “The population genetics and evolutionary epidemiology of RNA viruses,” *Nature Reviews Microbiology*, vol. 2, no. 4, pp. 279–288, apr 2004. [Online]. Available: <http://www.nature.com/articles/nrmicro863>

This paper reviews viral RNA mechanisms and some quantifications around the variables of mutation. Then it discusses how studies are structured for wet lab research of mutation and evolution studies. Helps to give a large overview of the topic of viral RNA without discussing Quasi-Species.

- [16] on behalf of the ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI), S. V. Tavtigian, M. S. Greenblatt, S. M. Harrison, R. L. Nussbaum, S. A. Prabhu, K. M. Boucher, and L. G. Biesecker, “Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework,” *Genetics in Medicine*, vol. 20, no. 9, pp. 1054–1060, sep 2018. [Online]. Available: <http://www.nature.com/articles/gim2017210>

This paper uses naive bayes method to develop a classifier for pathogeneity of a genetic variant. This method is okay, it is not very strong, but is laying the ground work for future studies. It does a good job of relating the ACMG guidelines to a mathematical framework.

- [17] J. M. Smith, “What determines the rate of evolution?” *The American Naturalist*, vol. 110, no. 973, pp. 331–338, 1976. [Online]. Available: <http://www.jstor.org/stable/2459757>

This paper discusses the pro’s and con’s of using differnt measures for understanding the evolution rates. The main idea is to use the fitness and its related measures as a stand in for the rate of adaptation.

- [18] P. D. Sniegowski and P. J. Gerrish, “Beneficial mutations and the dynamics of adaptation in asexual populations,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1544, pp. 1255–1263, 2010. [Online]. Available: <https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2009.0290>

The paper provides a review of models and developments for clonal interference modles. This sub area of asexual populations has grown significantly since the 1990’s when the first CI model was proposed. This paper breacks the models in the 3 catagories, Strong Seceltion Weak Mutation (SSWM), (SSSM), and (WSSM). The historical analysis and motivation for SSSM was discussed at legnth.

- [19] A. Wagner, “Information theory, evolutionary innovations and evolvability,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 372, no. 1735, p. 20160416, Oct 2017. [Online]. Available: <https://doi.org/10.1098/rstb.2016.0416>

This paper looked at the idea of applying information theory to the mapping between Phenotype space and genotype space. The model selected had significant restrictions for simplicity, but did allow for the establishing of some interesting results, like the fact that “The complexity increase is generally lower if the old site had high information content, which shows that phenotypic information change can depend on ancestral phenotypes and are thus contingent on evolution theory.” These results are based on the understanding of the amount of information stuck in a given moleclar structure.

- [20] C. O. Wilke, “Molecular clock in neutral protein evolution,” *BMC Genetics*, vol. 5, no. 1, p. 25, 2004. [Online]. Available: <http://bmcbgenet.biomedcentral.com/articles/10.1186/1471-2156-5-25>

An interesting paper discussing the impact of population size and mutation rate on the interpitation of molecular clock theory. This was an attempt to show that neutral theory could be used to help awnser the disconnect of emperical results and molecular clock theory, called “overdispersed molecular clock”. They were able to help account for overdispersion when mutation is low and population is resnably low.

- [21] —, “The Speed of Adaptation in Large Asexual Populations,” *Genetics*, vol. 167, no. 4, pp. 2045–2053, 08 2004. [Online]. Available: <https://doi.org/10.1534/genetics.104.027136>

Cool paper, Wilke does derive nice closed form equations, classic Wilke. Shows that as mutation increases rate of evolution decreases, this is because his def of adaptation is related to fixation. Interesting results all the same.

- [22] —, “Quasispecies theory in the context of population genetics,” *BMC Evolutionary Biology*, vol. 5, no. 1, p. 44, dec 2005. [Online]. Available: <https://bmcevolbiol.biomedcentral.com/articles/10.1186/1471-2148-5-44>

This paper is wonderful, it describes and relates the equations from quasispecies and the concepts from evolutionary theory. It shows how the QS equations can be derived from the principals mutation-selection equations.