**Core**

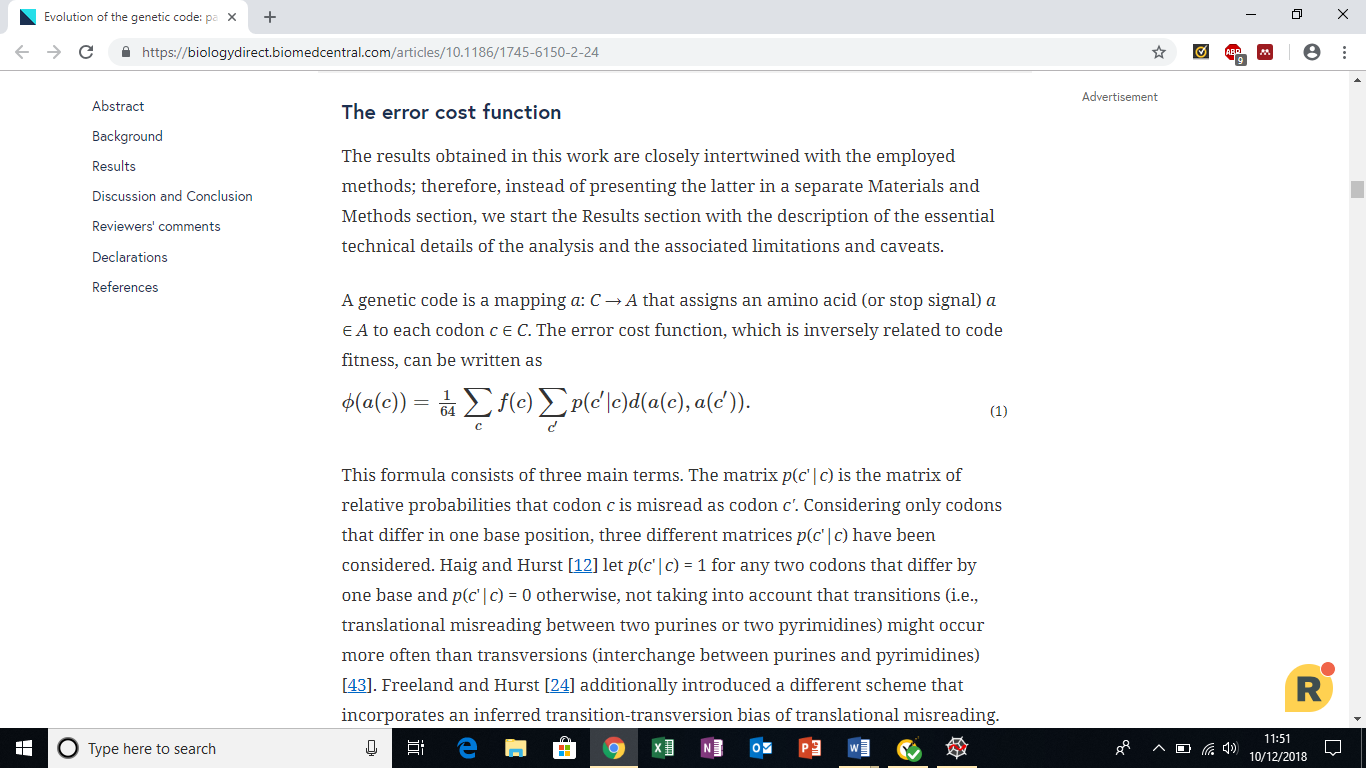
<https://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-2-24>

**Novozhilov, A.S., Wolf, Y.I. and Koonin, E.V., 2007. Evolution of the genetic code: partial optimization of a random code for robustness to translation error in a rugged fitness landscape. Biology Direct, 2(1), p.24.**

The paper compares the “fitness” of the standard genetic code, the one used by nearly all life, to that of optimized codes derived from random codes, these random codes still use the standard codon blocks as part of the random assignment of amino acids, with amino acids assignable only to the block type (2 or 4) that they occupy in the standard code. Trp and the stop codons remained the same as in the standard code.

The paper concludes that the standard code is a partly optimised random code (based against resistance to error). It also entails that the reason the standard code didn’t optimise further is in part due to the severity of effect of changing codon assignment in a growing system and the trade off against the benefits. It concludes that code adaptation is partly due to adaptation and partly to frozen accident.





<https://www.sciencedirect.com/science/article/pii/S0960982216309174>

Keeling, P.J., 2016. Genomics: evolution of the genetic code. Current Biology, 26(18), pp.R851-R853.

This is a review article, and while it isn’t indepth, it does make mention of the frozen accident model, and also mentions a paper:

K. Záhonová, A.Y. Kostygov, T. Ševčíková, V. Yurchenko, M. Eliáš: An unprecedented non-canonical nuclear genetic code with all three termination codons reassigned as sense codons:Curr. Biol., 26 (2016), pp. 2364-2369

The review explains that the main findings of zahonova is that a protist symbiont in insects all three stop codons encode amino acids, with UAA and UAG encoding glutamate as well as stop signals. With the method by how the distinction is made not being known. The symbiont is a trypanosome, which according to the article commonly alter the standard genomics practices.

<https://www.sciencedirect.com/science/article/pii/S0960982216307059>

**Záhonová, K., Kostygov, A.Y., Ševčíková, T., Yurchenko, V. and Eliáš, M., 2016. An unprecedented non-canonical nuclear genetic code with all three termination codons reassigned as sense codons. Current Biology, 26(17), pp.2364-2369.**

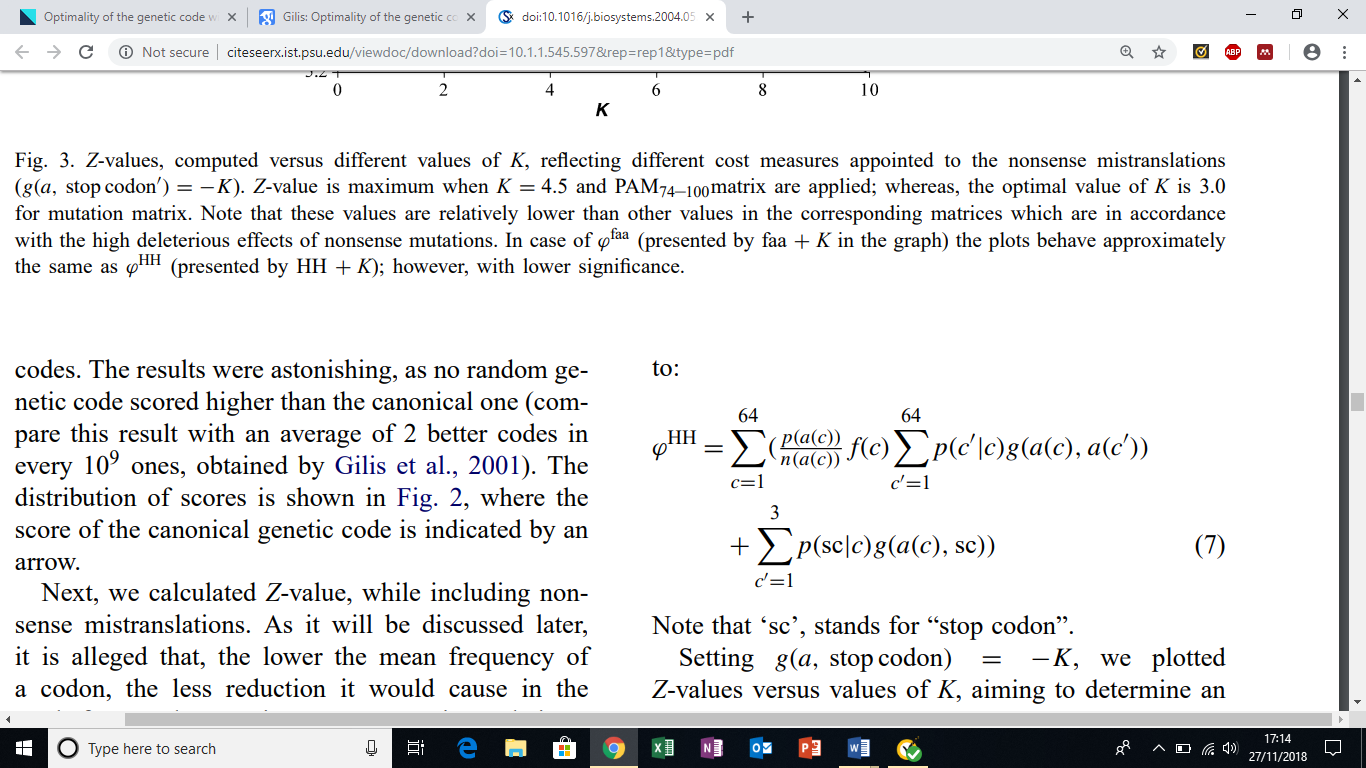
While studying the transcriptome sequence assembly of L. hesperus the auhors found a contaminant of a type of trypanosomatid organism, a protist. The trypanosomatid was assigned the genus Blastocrithidia. In this organism UAA and UAG encoded glutamate, and UGA encoded tryptophan.

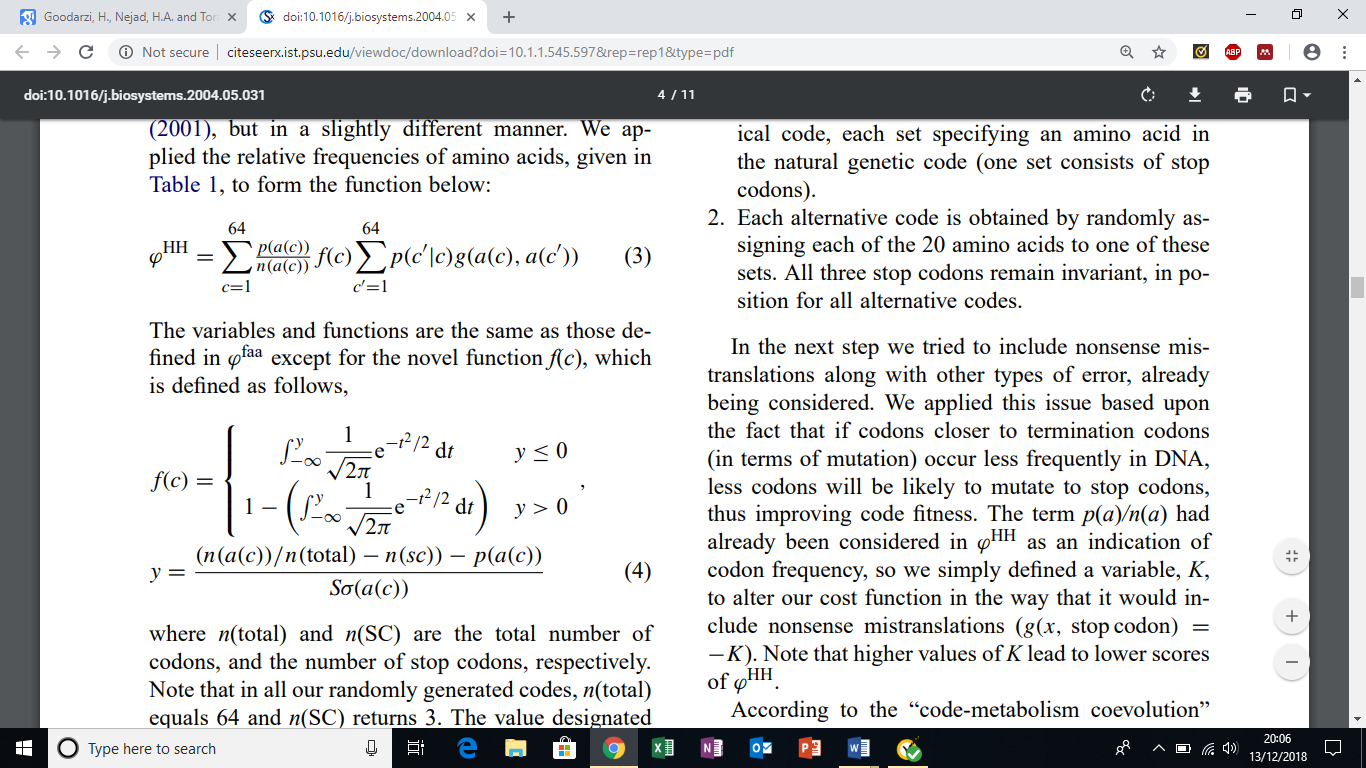
As for also serving as termination codons, no UGA terminations were inferred, while of the 115 transcripts of which as termination codon could be confirmed, 109 were UAA and 6 UAG i.e. UAA was clearly favoured as a termination.

<https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&q=Goodarzi%2C+H.%2C+Nejad%2C+H.A.+and+Torabi%2C+N.%2C+2004.+On+the+optimality+of+the+genetic+code%2C+with+the+consideration+of+termination+codons.+Biosystems%2C+77%281-3%29%2C+pp.163-173.&btnG=>

Goodarzi, H., Nejad, H.A. and Torabi, N., 2004. On the optimality of the genetic code, with the consideration of termination codons. Biosystems, 77(1-3), pp.163-173.

This paper tries to correct previous measurements of code optimality by also assessing the effect of mutation to stop codon. this equation is the new model:





<https://www.webpages.uidaho.edu/~stevel/565/literature/The%20genetic%20code%20is%20one%20in%20a%20million.pdf>

Freeland, S.J. and Hurst, L.D., 1998. The genetic code is one in a million. *Journal of molecular evolution*, *47*(3), pp.238-248.

The paper that came up with the p(c|c’) = (matrix using N). the paper doesn’t actually lay out the matrix in the standard form and gave no indication on how to derive N

**Genetic code evolution**

<https://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-4-16>

Higgs, P.G., 2009. A four-column theory for the origin of the genetic code: tracing the evolutionary pathways that gave rise to an optimized code. Biology direct, 4(1), p.16.

Details evolution of code from four codon blocks coding Gly, Ala, Asp, and Val with other AA’s having codons assigned later forming the current coding block structure. Provides a proposed starting point of the standard codes evolution.

<https://iubmb.onlinelibrary.wiley.com/doi/pdf/10.1002/iub.146>

Koonin, E.V. and Novozhilov, A.S., 2009. Origin and evolution of the genetic code: the universal enigma. IUBMB life, 61(2), pp.99-111.

Details various theories on how standard code evolved, frozen accident, coevolution and error minimilasition.

The papers basis will be useful as part of a literature review and background of the topic.

<https://pdfs.semanticscholar.org/f329/61aa4d5327ae51ad93a2fec174a90d3c6baa.pdf>

Crick, F.H., 1968. The origin of the genetic code. Journal of molecular biology, 38(3), pp.367-379.

Despite the age of the paper it mentioned several points still accepted by or mentioned in more recent papers. For example, the idea that only a few amino acids could have been encoded during the early stages of the codes evolution, the driving concept behind Higgs, P.G., 2009. A four-column theory for the origin of the genetic code: tracing the evolutionary pathways that gave rise to an optimized code. Biology direct, 4(1), p.16.

There also so notable gaps that have been filled in, for example suggesting that tRNA could serve as its own activating enzyme, clearly showing that the paper is before the discovery of ribozymes.

As the paper is old a lot of work has been done on it over the years (over 2000 citations on scholar), but it still remains useful since it covers the basics of the standard code.

<https://www.mdpi.com/2075-1729/7/2/22/htm>

Koonin, E.V., 2017. Frozen accident pushing 50: Stereochemistry, expansion, and chance in the evolution of the genetic code. Life, 7(2), p.22.

Paper is, as the title suggests, a review of frozen accident model of crick. As mentioned in many papers listed in this log, the general idea of the standard code in this model is that it evolved to a certain point and “froze” because any further changes would be deleterious to the point of lethality.

Paper mentions a argument for why there was only one last universal cellular ancestor (LUCA) from a Vetsigian, Woese, and Goldenfeld 2006 paper, that Horizontal gene transfer (HGT) would limit the number of codes that survived, with only one inevitably surviving, because of how inhibitive any change to the code is to the benefit of HGT.

Paper contains a brief outline of stereochemical, co evolution and error minimalization theory. It also mentions that evolution of error minimisation may have evolved neutrally as part of a initially limited code expanding to include more amino acids.

The translation system is universally conserved, but the protein based translation equipment, and tRNA, are paralogues and paralogues families. The paper also details the idea that the translation systems evolution is still largely mystery.

<https://www.jstage.jst.go.jp/article/pjab/84/2/84_2_58/_pdf>

Ohama, T., Inagaki, Y., Bessho, Y. and Osawa, S., 2008. Evolving genetic code. Proceedings of the Japan Academy, Series B, 84(2), pp.58-74.

General review.

<https://is.muni.cz/el/1431/jaro2008/Bi8350/um/5096325/Knight_99_Genetic_code.pdf>

Knight, R.D., Freeland, S.J. and Landweber, L.F., 1999. Selection, history and chemistry: the three faces of the genetic code. Trends in biochemical sciences, 24(6), pp.241-247.

Stereo chemistry stuff

<http://www.webpages.uidaho.edu/~stevel/565/literature/THE%20CASE%20FOR%20AN%20ERROR%20MINIMIZING%20STANDARD%20GENETIC%20CODE.pdf>

Freeland, S.J., Wu, T. and Keulmann, N., 2003. The case for an error minimizing standard genetic code. Origins of Life and Evolution of the Biosphere, 33(4-5), pp.457-477.

Error minimisation.

[https://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-4-44](https://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-4-44" \t "_blank)

Novozhilov, A.S. and Koonin, E.V., 2009. Exceptional error minimization in putative primordial genetic codes. *Biology Direct*, *4*(1), p.44.

Papr goes over evolution of primitive genetic code, including a two base codon system and order amino acids were acquired in, compared to Higgs 2009 has some different ideas on how code evolved.

Koonin, E.V. and Novozhilov, A.S., 2017. Origin and evolution of the universal genetic code. *Annual review of genetics*, *51*, pp.45-62.

Another paper reviewing the evolution of the genetic code, is recent and still pretty useful.

**(pdf in genetic code evolution papers email)**

<https://pure.mpg.de/rest/items/item_2539179/component/file_2548048/content>

Kuruoglu, E.E. and Arndt, P.F., 2017. The information capacity of the genetic code: Is the natural code optimal?. *Journal of theoretical biology*, *419*, pp.227-237

Paper on code optimality didn’t get through full paper but seemed useful and is more recent.

<http://www.webpages.uidaho.edu/~stevel/565/literature/THE%20CASE%20FOR%20AN%20ERROR%20MINIMIZING%20STANDARD%20GENETIC%20CODE.pdf>

Freeland, S.J., Wu, T. and Keulmann, N., 2003. The case for an error minimizing standard genetic code. Origins of Life and Evolution of the Biosphere, 33(4-5), pp.457-477.

Covers error minimisation theory.

<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-56>

Santos, J. and Monteagudo, Á., 2011. Simulated evolution applied to study the genetic code optimality using a model of codon reassignments. BMC bioinformatics, 12(1), p.56.

The paper focuses on the evolution of the code, so will be useful for information on the evolutionary theories on the code.

<http://hamedahmadi.com/publications/Optimality_Genetic_Code_Coevolution.pdf>

Goodarzi, H., Najafabadi, H.S., Hassani, K., Nejad, H.A. and Torabi, N., 2005. On the optimality of the genetic code, with the consideration of coevolution theory by comparison of prominent cost measure matrices. Journal of theoretical biology, 235(3), pp.318-325.

Another paper by goodarzi, focusing on the coevolution theory it also voices that the PAM74-100 mutation matrix isn’t a good choice when assessing the evolution of the genetic code

<https://arxiv.org/ftp/arxiv/papers/1210/1210.4322.pdf>

Chechetkin, V.R. and Lobzin, V.V., 2011. Stability of the genetic code and optimal parameters of amino acids. Journal of theoretical biology, 269(1), pp.57-63

This paper is related to the stereo-chemical theory on how the code evolved.

**(see pdf in code evolution papers email)**

Sengupta, S. and Higgs, P.G., 2015. Pathways of genetic code evolution in ancient and modern organisms. Journal of molecular evolution, 80(5-6), pp.229-243.

This paper focus is on how the code could have evolved in early life such as the idea that the first codons related to the most naturally abundant amino acids

**Variant codes (except Blastocrithidia)**

<https://onlinelibrary.wiley.com/doi/pdf/10.1002/bies.201600221>

Kollmar, M. and Mühlhausen, S., 2017. Nuclear codon reassignments in the genomics era and mechanisms behind their evolution. BioEssays, 39(5), p.1600221.

talks about many different species with codon reassignments, makes reference to zahonova.

**(use pdf from alt codons email)**

<https://onlinelibrary.wiley.com/doi/pdf/10.1002/bies.201600213>

Alkalaeva, E. and Mikhailova, T., 2017. Reassigning stop codons via translation termination: How a few eukaryotes broke the dogma. BioEssays, 39(3), p.1600213.

more of a summary of different codon reassignments, and also makes direct reference to zahonova.

**(use pdf from alt codon email)**

<https://www.sciencedirect.com/science/article/pii/S0960982201000288>

Lozupone, C.A., Knight, R.D. and Landweber, L.F., 2001. The molecular basis of nuclear genetic code change in ciliates. Current Biology, 11(2), pp.65-74.

Referenced the changes in ciliates ( changes to UAR codons) potentially can **test code robustness** test if there time.

**LUCA**

<http://www.docentes.unal.edu.co/vcorredore/docs/nrmicro-127-03_Koonin.pdf>

Koonin, E.V., 2003. Comparative genomics, minimal gene-sets and the last universal common ancestor. Nature Reviews Microbiology, 1(2), p.127.

<https://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-3-29>

Glansdorff, N., Xu, Y. and Labedan, B., 2008. The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. Biology direct, 3(1), p.29.

**Error cost matrixes**

<https://genomebiology.biomedcentral.com/articles/10.1186/gb-2001-2-11-research0049>

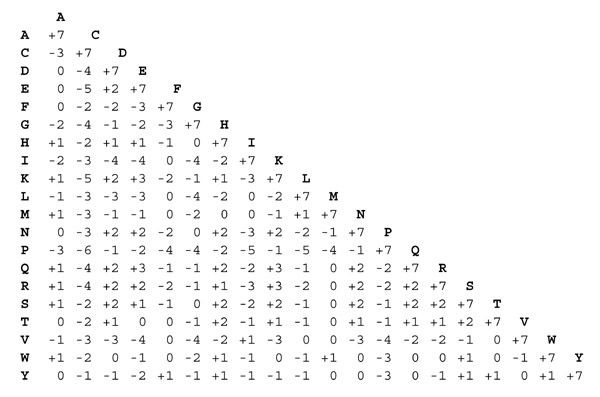
Gilis, D., Massar, S., Cerf, N.J. and Rooman, M., 2001. Optimality of the genetic code with respect to protein stability and amino-acid frequencies. Genome biology, 2(11), pp.research0049-1.

Used as an error cost equation in the Novozhilov code robustness test

Matrix defined as g code (a,a') = 3 - Δ(a,a') where Δ(a,a') is 0 if the same AA is encoded and otherwise is equal to the minimum number of base changes needed to change the encoded residue.

when a or a' corresponds to a stop codon, we set g(a,a') = 0

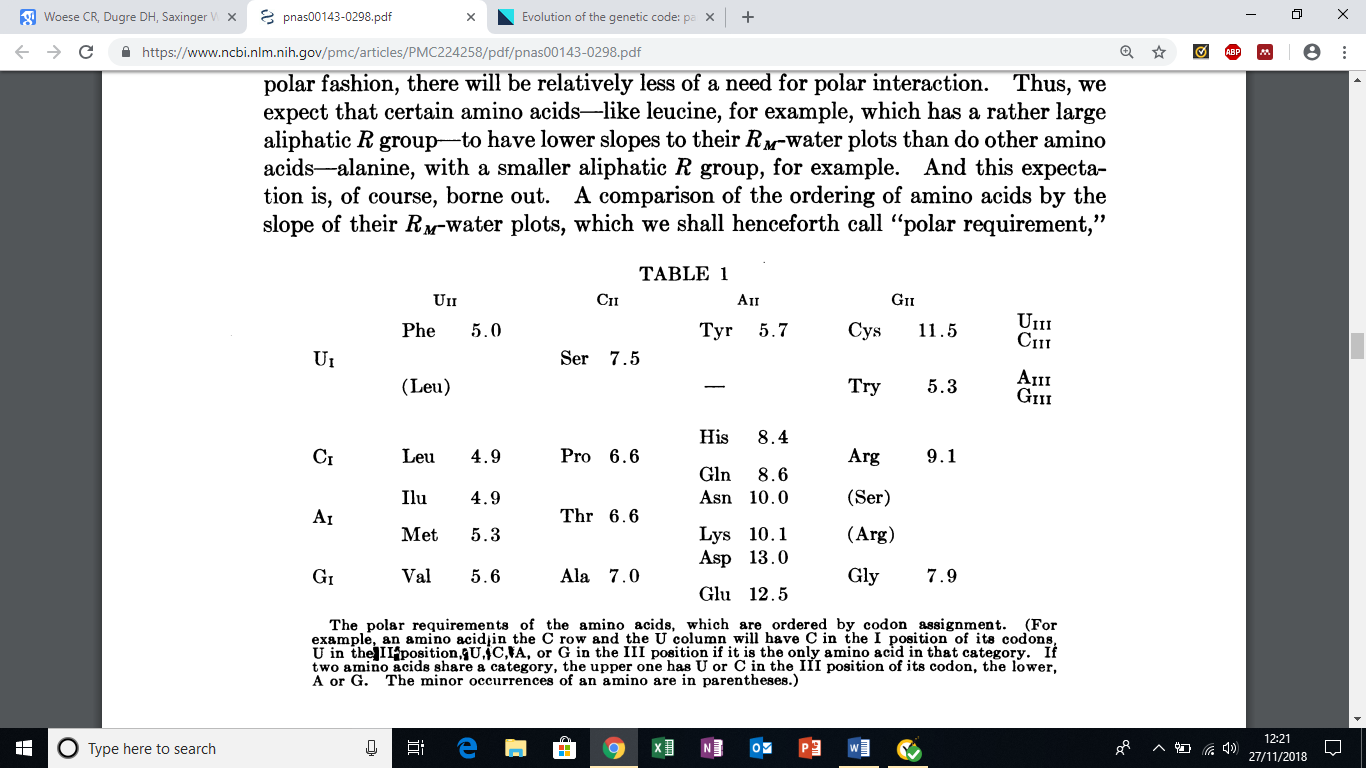
Having gone back over Gilis et al 2001, the equation g mutate (a, a') = M(a, a') is the equation that novozhilov et al 2007 most likely used to define d(a(c),a’(c’)) have used, M(a,a’) is a matrix of free energy changes when the amino acids are substituted.

Below is the matrix for M(a,a’)

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC224258/pdf/pnas00143-0298.pdf>

Woese, C.R., Dugre, D.H., Saxinger, W.C. and Dugre, S.A., 1966. The molecular basis for the genetic code. Proceedings of the National Academy of Sciences of the United States of America, 55(4), p.966.

This paper gives the Polar requirement of the amino acids of the standard code which is used as one of the bases for the error cost function in Novozhilov et al 2007. However, it does not give a polar requirement for the stop codon signals.



<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.545.597&rep=rep1&type=pdf>

Goodarzi, H., Nejad, H.A. and Torabi, N., 2004. On the optimality of the genetic code, with the consideration of termination codons. Biosystems, 77(1-3), pp.163-173.

This paper tries to correct previous measurements of code optimality by also assessing the effect of mutation to stop codon.

