1844,
Division of Biochemical Sciences,
National Chemical Laboratory,
Pune - 411008.
12 March, 2005

Dear Sir,

I hope this letter finds you in good health. I am well, and am preparing for my interviews for N.C.L. and T.I.F.R.

Sir, I am writing to you with respect to the simulation programmes you had written when I had visited you in January. Unfortunately, I have not been able to pursue that further as I have gotten caught in various other things. Nevertheless, I am still keen on pursuing the line of thought and action that I had discussed with

you fruitfully at C.E.S.

However, herein I have been thinking of putting the simulations to another purpose — that of speculating why the genetic code universal to all organisms is as it is, without any variations except in eucaryotic organellar genomes, which contain their own DNA with slightly altered codes.

It has been speculated previously that the code is accidental — that it just happens to be what it is without having undergone any evolution. Prof. Apoorva Patel speculates that it is optimised to carry one-dimensional information in D.N.A. for transmission

and three-dimensional information in proteins for encoding biological functions. I heard him on December 31, 2004 at a lecture in IUCAA, and he showed by clever quantum computing calculations as to how four bases and twenty amino acids could achieve the information storage, transmission and interpretation objectives.

Nevertheless, I have another idea. Though Dr. Patel suggested that the code is optimal, he did not say why the known four bases and twenty amino acids and not any others are optimal, nor did he say anything about the particular correspondences between the 64 triplet codons and 20 amino acids (plus a sign for stop) that exists.

It is assumed to be optimal since natural selection would have replaced the genetic code with something more optimal had it been there. But this argument may not be entirely true; it may be entirely possible that a sub-optimal system may persist and reach fixation even if a more suitable system might be available.

I can cite this with two examples. The QWERTY keyboard is less optimal than the Dvolak keyboard, yet it persists because a larger number of people over several generations have been trained to use it. Similarly English may not encode human language parameters better than say Latin or Sanskrit,

yet it has emerged as the global language, due to the historical and present control over human resources of the British Empire and USA respectively. Thus both sub-optimal systems have prevailed due to a HEAD START they got.

Similarly a population with a genetic code (that became the universal genetic code of today) might have driven others to extinction due to an accidental (or stochastic) control of a large proportion of general resources (light, metabolites etc), or control of a critical micro-resource (such as enzyme co-factors or minerals). I have calculated that there might be $!21 + 21^{43}$ different codes for

a system with 64 places (codons) and 21 semantic designations (read 20 amino acids + STOP). How only one of these has survived to date is thus intriguing, ignoring for the moment organellar codes. In fact the organellar codes, being as good as the standard code or better, lend to me some credence that it may not have been natural selection at work.

It is here that your simulations might help test the case. I wonder if small fluctuations in the heritability of different codes might lead to a situation wherein one single code might prevail in the end.

I hope this might interest you, and also hope that we may

be able to do something further.

My guide, Dr. Shama Barnabas, evinced interest in your January simulations and said that we could take it forward once she returns from the United States of America where she is visiting her children.

Wishing you all the very best in your endeavours,

Yours truly,

R. Ramesh

P.S. I have attached my calculations on how I arrived at $!21 \times 21^{43}$ possible genetic codes.