To Carry a Longer Genome One Must First Copy More Accurately: the Order Is Fixed

— The error threshold is mutation rate times genome length below one: influenza A stands at **0.312800**, HIV-1 at **0.291000**, SARS-CoV-2 at **0.029900** and herpes, a DNA virus, at **0.001520** — the separator is whether accuracy comes first or length does — [Paper 692]

Yuuki YAMAGISHI (山岸 勇輝)
Yamagishi Complexity Torus Lab, Japan
September 2026

— *From thinking to syncing.* —

Abstract

RNA viruses are said to mutate readily. Combined with the length of their genomes, that rate sits just inside a threshold.That length and accuracy stand in a fixed order is counted here as the room left below the threshold.No new theorem or law is claimed. First, all four lie inside the threshold——the product of mutation rate and genome length is 0.312800 for influenza A, 0.291000 for HIV-1, 0.029900 for SARS-CoV-2 and 0.001520 for herpes, none of them above one (Section 2). **Second, the RNA side lies just inside it——the error-free fraction for influenza A is 0.731396, so more than one copy in four differs somewhere** (Section 2). **Third, and this is the core. A given accuracy sets a ceiling on length, the reciprocal of the mutation rate——influenza A has only 3.196931 times its present length to spare** (Section 3). **Fourth, the accurate side has room of another order——herpes has 657.894737 times to spare, which is 205.789474 times the room influenza A has** (Section 3). **Fifth, so an order is fixed——improving accuracy tenfold raises the ceiling tenfold, while lengthening first is not available: accuracy must come first** (Section 3). **Sixth, the separator is whether accuracy comes first or length does——the two are not independently choosable; one is the ceiling of the other** (Section 4). **Keywords**: error threshold, mutation rate, genome length, quasispecies, proofreading, accuracy before length.

> **Scope of this paper (scope note)**: **No new theorem or law is claimed——the error threshold of quasispecies theory, the mutation rates of RNA viruses and coronaviral proofreading are all standard (Eigen 1971, Drake et al. 1998).Evolution is not discussed——rates of adaptation, the acquisition of drug resistance and vaccine design are not the subject.Molecular mechanism is not treated——how proofreading enzymes work is not treated; only the numerical fact of higher accuracy is used.The values are not claimed——mutation rates and genome lengths are representative and move with strain and method, and the threshold of one is itself a guide obtained on a flat fitness landscape.No hazard is assessed——the threat posed by any real virus is not treated.Relation to earlier papers**: **Paper 398 counted that the same one-letter error costs 15.75 times more at one position than another, and exactly 0 at one of them——there the cost depended on where the error fell; here each error weighs the same and only the total matters.**What is added is **setting the four products side by side inside the threshold**, **giving the error-free fractions**, **putting the room below the ceiling at 3.196931 and 657.894737**, **putting their ratio at 205.789474**, and **placing the separator on whether accuracy or length comes first**.

1. Notation and the Question

Setting: **at each replication an error enters with probability μ per base. With a genome of length L, the number of errors per copy is μL.**

$$\mu L < 1, \qquad \Pr(\text{error-free copy}) = e^{-\mu L}, \qquad L_{\max} = \frac{1}{\mu}, \qquad \text{room} = \frac{L_{\max}}{L}$$

Beyond $\mu L < 1$ the original sequence cannot be held across generations: this is the error threshold.

Question: **can genome length and copying accuracy be chosen independently?**

2. All Four Lie Inside the Threshold

Computation 2.1: **with representative values, the errors per copy and the error-free fraction were derived.**

virus	mutation rate	length	errors per copy	error-free
influenza A	2.3×10^{-5}	13600	0.312800	0.731396
HIV-1	3.0×10^{-5}	9700	0.291000	0.747516
SARS-CoV-2	1.0×10^{-6}	29900	0.029900	0.970543
herpes (DNA)	1.0×10^{-8}	152000	0.001520	0.998481

Observation 2.2

that all four lie inside is itself the mark of a selection.A combination beyond the threshold cannot hold its sequence and so is not here to be tabulated——what appears in the table is only what did not cross.

Observation 2.3

the RNA side, though inside, is at the edge.Influenza A leaves 0.731396 of copies error-free, so more than one in four differs somewhere——"mutates readily" names two things at once: a large μ, and a μL near one.

3. The Ceiling Is the Reciprocal of the Accuracy

Computation 3.1: **the longest genome each accuracy allows, and the room left against the present length, were derived.**

virus	ceiling on length	room against present length
influenza A	4.347826×10^{4} nt	3.196931
HIV-1	3.333333×10^{4} nt	3.436426
SARS-CoV-2	1.000000×10^{6} nt	33.444816
herpes (DNA)	1.000000×10^{8} nt	657.894737

Observation 3.2

this is the core: length grows only as far as accuracy permits.Influenza A has 3.196931 times to spare and herpes 657.894737, a ratio of 205.789474——without raising accuracy first, the length cannot be moved.

Observation 3.3

the order does not run the other way.Improving accuracy tenfold raises the ceiling tenfold, but lengthening tenfold first does not improve accuracy; it crosses the threshold and the sequence is lost——the two quantities are not equals: one is the ceiling of the other.

Observation 3.4

the intermediate case supports the order.SARS-CoV-2 is long for an RNA virus at 29900 nt, and its mutation rate is an order lower at 1.0×10^{-6}, leaving room of 33.444816——long RNA genomes are found only on the accurate side.

4. The Separator

Observation 4.1

the separator is whether accuracy comes first or length does.They look like two independent choices, but their product cannot exceed one, so each binds the other——the question settled is not which causes which, but which must come first.

Observation 4.2

beside Paper 398.Paper 398 counted that the same one-letter error costs 15.75 times more at one position than another, and exactly 0 at one of them——there the position of an error fixed its cost; here position does not count and only the total does. One word "error" acts in two ways.

Observation 4.3

what is not said.Evolution and the molecular mechanism of proofreading are not treated, and every value is representative——what is shown is only that length and accuracy stand in a fixed order.

5. Roll Call and Closing

1. **influenza A at 0.312800** (Computation 2.1): **at the edge.**
2. **HIV-1 at 0.291000** (Computation 2.1): **likewise.**
3. **SARS-CoV-2 at 0.029900** (Computation 2.1): **an order inside.**
4. **herpes at 0.001520** (Computation 2.1): **further inside.**
5. **error-free 0.731396** (Computation 2.1): **one in four differs.**

6. **ceiling is $1/\mu$** (Computation 3.1): **the reciprocal of accuracy.**
7. **room 3.196931** (Computation 3.1): **the core.**
8. **herpes has 657.894737** (Computation 3.1): **another order.**
9. **a ratio of 205.789474** (Observation 3.2): **accuracy binds length.**
10. **accuracy first or length first** (Observation 4.1): **the separator; Paper 398.**

Closing: **To carry a longer genome one must first copy more accurately.The product of mutation rate and genome length is 0.312800 for influenza A, 0.291000 for HIV-1, 0.029900 for SARS-CoV-2 and 0.001520 for herpes, all four inside the threshold of one——a combination beyond it cannot hold its sequence, so only those that did not cross are here to be tabulated.The ceiling a given accuracy allows is the reciprocal of the mutation rate, leaving influenza A 3.196931 times its present length and herpes 657.894737, a ratio of 205.789474.Improving accuracy tenfold raises the ceiling tenfold, while lengthening first does not improve accuracy——the two are not equals; one is the ceiling of the other.That SARS-CoV-2, long for an RNA virus, holds room of 33.444816 follows from a mutation rate an order lower.The separator is whether accuracy comes first or length does.Placed among the earlier papers——in Paper 398 the position of an error fixed its cost; here position does not count and only the total does.To be honest——evolution and the mechanism of proofreading are not treated, and every value is representative.**

References

1. Drake, J. W., Charlesworth, B., Charlesworth, D., & Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics*.
2. Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*.
3. Yamagishi, Y. (2026). *The Same One-Letter Typo Costs* **15.75** *Times More at One Position Than Another, and at the Second It Is Exactly* **0** [Paper 398]. Zenodo. 10.5281/zenodo.22699326

— From thinking to syncing. —

Yuuki YAMAGISHI (山岸 勇輝)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)