THE CONTINUITY PRINCIPLE AND THE EVOLUTION OF REPLICATION FIDELITY

Seymour Garte

Department of Pharmacology and Toxicology, Rutgers University, Piscataway, NJ USA.

Author information:

Seymour Garte

Department of Pharmacology and Toxicology, Ernest Mario School of Pharmacy, Rutgers University, 160 Frelinghuysen Road, Piscataway, NJ 08854-8020, USA.

Corresponding author: phone – 917 526 3980, sy.garte@rutgers.edu

Abstract

Evolution in modern life requires high replication fidelity to allow for natural selection. A simulation model utilizing simulated phenotype data on cellular probability of survival was developed to determine how self-replication fidelity could evolve in early life. The results indicate that initial survivability and replication fidelity both contribute to overall fitness as measured by growth rates of the cell population. Survival probability was the more dominant feature, and evolution was possible even with zero replication fidelity. A derived formula for the relationship of survival probability and replication fidelity with growth rate was consistent with the simulated empirical data. Quantitative assessment of continuity and other evidence was obtained for a saltation (non-continuous) evolutionary process starting from low to moderate levels of survival probability and self-replication fidelity to reach the high levels seen in modern life forms.

1. INTRODUCTION

The Continuity Principle (Wolpert 1994) in evolution has been described as the "…general Darwinian principle… [that] evolution must proceed via consecutive, manageable steps, each one associated with a demonstrable increase in fitness" (Wolf and Koonin 2007). While Darwin envisaged a strict requirement for very small steps, we now know that there are many exceptions to this, wherein it is possible for a major evolutionary change to occur that leads to a relatively sudden and dramatic increase (or jump - "saltation") in fitness (Eldredge et al. 1972, Koonin 2007, Fontana and Schuster 1998, Raggi et al. 2016). Examples of saltation in evolution include the endosymbiosis event leading to the origin of eukaryotes (Margulis 1996, Martin et al. 2015), the whole genome duplications at the origin of the vertebrates (Dehal and Boore 2005, Coate and Doyle 2011), and the major radiation of the Cambrian explosion (Morris 1989, Lee et al. 2013). Mechanisms for these rare violations of the continuity principle are now part of standard evolutionary biological theory (Minelli et al. 2009, Theissen 2009, Laland et al. 2015).

Continuity is a major issue in the origin of life (Gabora 2006). Starting with chemical evolution, there are many steps in the process to reaching a living cell that have yet to be elucidated, with unknown mechanisms and unknown degrees of continuity. One of the most important of these to biologists is the origin of evolution, which allows for the transition from chemical to biological complexity (Raggi et al. 2016).

Evolution is not synonymous with life in its simplest form, although all known living forms contain the molecular apparatus and mechanisms that make evolution possible. One of the key features of proto-cells required for biological evolution is self-replication with some degree of accuracy.

The understanding that replication (meaning self-replication) is of fundamental importance in the origin of life underpins the general concept of “replication first” in origin of life theories (Szathmáry and Smith 1997, Levin and West 2017). Since the cell phenotype is the target of natural selection, some mechanism to replicate the phenotype is critical to allow for evolution by natural selection; therefore replicators (such as RNA or DNA) must not only be able to replicate themselves, but also have the capacity to

allow for directing phenotype replication. This is achieved by RNA ribozymal activity in the hypothetical RNA world (Cech 2012, Robertson and Joyce 2012), and by DNA-directed translation in modern life forms. If phenotypic replication does not occur, then any beneficial changes giving a selective advantage to daughter cells will be lost.

The issue of self-replication fidelity has been addressed in several biological contexts, such as ”error catastrophes” in viruses and cells (Eigen 2002, Summers and Litwin 2005), in evolution (Goel and Yčas 1975, Simons 2002), and the relationship of mutation rate to evolution (Edelmann and Gallant 1977). The requirement for minimal replication fidelity to allow for survival and evolution in early cells in any hypothesized RNA world has been estimated at approximately 0.98 (Szostak 2012), but scenarios for the evolution of replication fidelity in informational molecules are unknown.

Some have questioned the reality of RNA world (Preiner et al. 2020, Carter 2016), and other theories of origin of life that do not rely on informational molecule self-replication have been postulated (Xavier et al. 2020).

The molecular aspect of the evolution of self-replication (or the evolution of evolution) is not the focus of this paper. The question addressed here is

how replication fidelity evolved; and, specifically, whether the evolution of high replication fidelity could follow the continuity principle. A theoretical and statistical approach using simulations of cell division and population growth was used to answer some basic questions on the role of replication fidelity in the evolution of the earliest living cells.

A critical biological factor for the success of cells during early abiogenesis can be defined as the summation of all of the ingredients of evolutionary fitness (such as energy capture and usage, metabolic efficacy, homeostasis, etc.) into a single parameter of the probability of survival of each cell in a population, between cell divisions. Overall fitness would then be composed of some combination of this probability with replication fidelity. One goal of this study was to determine if any quantitative relationships between cell survival probability (P_s) and replication fidelity (F_R) can be derived. A second goal was to investigate whether a simulation model for cell division with known starting conditions of P_s and F_R could elucidate the issue of evolutionary continuity in early development of living cells. This study is neutral with respect to any details of early life mechanisms (such as replication or metabolism first), although the results do bear on this issue (see Discussion).

2. METHODS

2.1 Replication Fidelity and Probability of Survival

Replication fidelity (F_R) is defined as the probability that the cellular phenotype is replicated with complete accuracy. This allows for a quantitative measure of replication fidelity from 0 (no fidelity at all) to 1 (perfect fidelity). In modern cells, this value is 0.99999. The mutation rate is $1 – F_R$. The survival probability of cells (P_s) is defined as the probability that a living cell will survive to the point of division. No assumptions about the molecular mechanism of replication or the components or determinants of cell survival probability are made.

2.2 Growth Rates

A simulation program of cell growth based on binary reproduction starting with a single cell with initial input values of P_s and F_R was created to determine the fate of all daughter cells for 6 to 10 generations. Based on the number of surviving cells at each generation, growth rates (k) were calculated using Matlab (Mathworks 2015). Above 10 generations, computer power became an obstacle, although results were checked in a

small number of runs with higher generation numbers (requiring run times of several days). No significant effect of generation number on results was found, at least up to 14 generations.

For each cell division, it was assumed that daughter cells either retained the P_s of the parent cell, or acquired a new, randomly determined value of P_s based on a Monte Carlo subroutine using the initial value of F_R within constraints imposed by other constant pre-set parameters: the magnitude of the mutational effect (the maximal degree of change of the starting P_s value) and the proportion of beneficial to harmful mutations (which determine the direction of the change in P_s).

Living cells were counted at each generation based on a second Monte Carlo simulation routine to convert each cell's survival probability (P_s) into a simulated binary value corresponding to being alive or dead. Surviving cells were used as parent cells for the next generation. Each run consisted of 500 replicates, and average cell counts for every generation were used in further calculations.

Total surviving cell numbers at each generation (C_t) were used to calculate the growth rate constant (k) by

$$C_t = C_{t0}e^{kt} \qquad \text{eq. 1}$$

Where t is measured in generations. For every run, each calculation of k was repeated 500 times to simulate a starting population C_{t0} of 500 cells. If even a single one of the 500 runs produced at least one viable cell after 10 generations, the value of k would be greater than zero.

For bacteria or any single-cell life form that reproduces by fission into two daughter cells, we generally limit the maximal value of growth rates to 2 (population doubling), using

$$C_t = C_0K^t \qquad \text{eq. 2}$$

Where K (sometimes referred to as “r”) varies from 0 to 2. Thus, K can be calculated from cell count data at time = t by

$$K = e^{\ln(C_t)/t} \qquad \text{eq. 3}$$

And the growth rate k = ln(K). For negative growth rates (k < 0, K < 1), the number of generations until extinction or time to extinction (TTE) was calculated by

$$TTE = \frac{\ln\left(\frac{1}{C_{t0}}\right)}{k} \qquad \text{eq. 4}$$

The role of the mutational effect magnitude (from 0.1 to 0.5) and of the proportion of beneficial mutations (from 0.01 to 0.2) was evaluated for all endpoints.

2.3 Quantitative Determination of Continuity

One way to think about continuity is to picture a process that progresses in small steps where each step produces a meaningful difference in an outcome compared to the previous step. If the outcome is a measurable quantity, we can assess meaningful differences by statistical significance. Continuity can then be measured by the smallest steps of some parameter that result in statistically significant differences in outcome. If a large number of steps is required before a significant outcome difference is

observed, then continuity is broken, and the best explanation is saltation. The average minimal distance (D_m) between two parameter values under a fixed set of conditions that give a significant difference in outcome is a quantitative determination of the degree of continuity in the process. In living cells, such discontinuities can arise from multiple mutations, or single mutations acting on regulatory elements or other mechanisms. The D_m statistic, however, is not relevant to any specific mechanism of discontinuity, but only to its quantitation.

In the present model, the outcome is the value of the growth constant K, as determined by the experimental simulation. The parameter of interest for testing continuity is the replication fidelity, F_R. For a smoothly continuous process, we would expect small differences in F_R to result in statistically significant differences in the value of K. On the other hand, if no difference in K is observed until the distance between any two values of F_R is quite large, then continuity would appear to be broken.

Mean and standard deviations of multiple runs of the simulation algorithm were determined at values of F_R that differed by 0.01 to calculate statistically significant differences in K at various distances between F_R

values (where D is defined as $F_{R,2} - F_{R,1}$) given a constant P_s. The results were tabulated for every distance using multiple values of $F_{R,1}$ and $F_{R,2}$. A quantitative estimate of continuity was given by the average minimum distance (D_m) that produces a p value < 0.05 for differences in F_R values for each constant P_s.

3. RESULTS

The overall fitness of a population of cells can be estimated by its growth rate, given by K. Values of K below 1 lead to eventual extinction, while K > 1 allows for population survival and growth. The maximal value for K for single cells is 2, signifying a doubling of population size per unit of time.

3.1 Survival Probability and Replication Fidelity as Determinants of Growth Rate

Simulations of population growth with different initial values of cell survival probability (P_s) and replication fidelity (F_R) indicate that both parameters contribute, but not equally, to the value of K. The minimum value for P_s as a function of F_R are shown in Figure 1. Values of these parameters below the

line lead to extinction; this curve presents the lower limits of initial survival probability and replication fidelity in the earliest life forms.

One surprising result is that high replication fidelity is not a requirement for long term organismal survival in all circumstances. Contrary to expectation, cell populations with no initial replication fidelity at all can survive and evolve, although at a very low growth rate (K = 1.02 on average). However, the lower the value of F_R, the higher the initial survival probability P_s must be to maintain a positive growth rate. At $F_R = 0$, the minimum value of P_s is 0.95 to allow for population survival and further evolution.

Some representative data points of P_s and F_R are shown in Table 1 along with values for K, and calculated time (in generations) for the population with these initial parameter values to go extinct (TTE).

3.2 Evidence for Discontinuity in the Evolution of Replication Fidelity at Low Survival Probability

Figure 2 shows the influence of P_s on K at different constant values of F_R. It appears from these curves that the survival probability P_s of cells is the primary determinant of K, and of whether a population will survive or die out. While at each fixed F_R there appears to be a linear relationship

between P_s and K at higher values of P_s, this is not seen at lower values, irrespective of the value of F_R.

This is confirmed by the lack of any quantitative relationship between F_R and K at $P_s = 0.1$, as seen in Figure 3. This is extended to each value of P_s from 0.1 to 1.0 in Figure 4, which indicates a linear relationship between fidelity of replication and K at values of $P_s > 0.4$, and the absence of such a relationship below that value.

3.3 The Relationship between Growth Rate, Survival Probability, and Replication Fidelity

The quantitative relation between P_s, F_R, and K was determined to fit the following empirical formula derived from simulation data, for $P_s > 0.3$ and all values of F_R:

$$K = P_S(F_R + 1) - 0.25(1 - P_S) \qquad \text{eq. 5}$$

The dominant role of P_s vs. F_R is derived from the first term in the equation, while the second term is a correction related to the probability of death (1-P_s) of one of the 4 descendants of any pair of cells in the previous generation. (See Methods.)

For values of P_s between 0 and 0.3, no quantitative relationship was found (as illustrated in Figures 3 and 4) between F_R and K.

In Figure 5, it is clear that the data fit this equation quite well at higher values of P_s, less so at $P_s = 0.3$, and not at all below $P_s = 0.3$.

The minimum values of F_R and P_s to allow for K > 1 (or survival of the population) can be calculated from eq. 5:

$$P_S F_R > 1.25(1 - P_S) \qquad \text{eq. 6a}$$

So that

$$F_{R(min)} = \frac{1.25(1-P_S)}{P_S} \qquad \text{eq. 6b}$$

Overlaying the curve using equation 6b on the simulation data shown in Figure 1 gives the result in Figure 6.

Equations 5 and 6A suggest that even with perfect replication fidelity (F_R = 1) values of P_s below 0.625 will give K < 1 and lead to extinction. This result

was confirmed by the experimental simulations, which consistently gave a value of P_s between 0.63 and 0.65 as the minimal value of P_s to allow for population survival and evolution.

3.4 Effects of Other Initial Conditions.

Two other initial condition parameters were assessed for their effect on growth rate and fitness. These were the effect magnitude of mutations, and the proportion of beneficial to deleterious mutations. All data shown here were calculated assuming a mutation effect magnitude of 10%, and beneficial to deleterious mutation ratio of 0.05 (1 beneficial mutation for every 20 deleterious ones). Raising the mutation effect magnitude to 30% had a negligible effect on the results. Similarly, changing the proportion of beneficial mutations to deleterious mutations in the range from 0.01 to 0.10 (Keightley and Lynch 2003) had almost no influence on K. Effects of this ratio on K were seen only above the unrealistic value of 0.2. There was a small effect of the number of generations on the minimum P_s to allow for population survival. For generation times > 10, there was an asymptotic trend toward the theoretical limit as given by equation 1. No examples of rare mutations with very high P_s values were seen as outliers in the standard run of 500 replicates, nor when the number of runs were

increased up to 2500, suggesting that a mutation allowing for a large jump in survival probability is extremely rare.

3.5 Quantitative Assessment of Continuity

While the results presented so far are consistent with saltation rather than continuity in the development of replication fidelity in cells with low initial survival probability, it was possible to use the statistical properties of the simulation-determined values of K to directly quantify the extent of continuity at various levels of survivability.

Table 2 shows two examples of the quantitative approach using an initial P_s value of 0.5, and two examples of results for K at various ranges of initial F_R. The table shows means and standard deviations (SD) for K at every initial F_R, with progressive steps of 0.01. On the right side of the table are shown the results of statistical tests for significance using Student's t test comparing an initial F_R ($F_{R,1}$) to every possible increased value ($F_{R,2}$).

In the examples in Table 2, the minimum distance between $F_{R,1}$ and $F_{R,2}$ that gave a significantly different result in K (D_m) were between 0.03 and 0.04. This procedure was repeated for all 100 starting values of F_R (from 0.01 to

1.0), and values of D_m for each of several selected P_s values were averaged over the complete range. No effect of the initial magnitude of $F_{R,1}$ was seen for D_m, as shown by the relatively low standard deviations in Table 2.

As expected, for $P_s = 1$, D_m was close to the minimal possible value, or the size of the step (C = 0.01), indicating maximal continuity, whereas for low values of P_s the break in continuity was shown by higher values of D_m. Surprisingly, the curve of D_m vs. P_s (Figure 7) showed a close fit to the inverse square law.

$$D_m = \frac{C}{P_S^2} \qquad \text{eq. 7}$$

Where C is a constant equal to the size of the step.

4. DISCUSSION

The origin of the universal, critical biological phenomenon of self - replication has been studied using statistical mathematical approaches (Szathmáry 2006, England 2013). In all modern life, there is a threshold of replication fidelity below which cells cannot survive. At “high” mutation rates (which can be as low as 0.1 or $F_R = 0.9$), modern organisms undergo an

error catastrophe (see above) from which they cannot recover. The existence of this threshold for mutation rate has been used as an antiviral strategy (Eigen 2002), and it also probably puts a hard limit on the rate of evolution. It is not at all clear that primitive, less complex cells at the origin of life also had such mutational limits, and for the purposes of being conservative in assumptions, the phenomenon of mutational thresholds was not considered in the model.

In complex modern cells, survival probability and replication fidelity are linked, so that deficits in replication fidelity (at either the DNA replication, transcription, or translation phases) result in decreased P_s in future generations. In this report, it is assumed that in early primitive living cells with much less interactive chemical complexity, no such linkage existed. While this assumption cannot be supported by any evidence, the fact remains that even if some degree of error catastrophe did play a role in cell survival, that would result in an even stronger degree of discontinuity in the evolution of both P_s and F_R. The assumption of no linkage between the two parameters (and no threshold for error catastrophe in early cells) is therefore conservative in relation to the hypothesis of saltation.

During the origin of life, it's likely that both P_s and F_R were not near the very high values they hold today for almost all species under normal circumstances but were closer to 0 than 1. The simulation data show that at higher levels of P_s (>0.6), the evolution of efficient and accurate replication can proceed with a high degree of continuity through the entire range of initial replication fidelity. Since every improvement in F_R will lead to a significant increase in overall fitness (as measured by the value of K), this suggests a smooth evolutionary process leading to the very high value of F_R seen in all modern cells since the Last Universal Common Ancestor (LUCA).

However, the converse does not hold true. For low initial P_s values (<0.4), the evolution of increased replication fidelity requires large jumps in F_R that suggest a saltational evolutionary path. Furthermore, there is no continuity from low P_s to higher values, depending on the simultaneous value of F_R. Even at perfect replication fidelity ($F_R = 1$), the minimal level of P_s must be at least 0.6 in order to allow for any possibility of growth, survival and evolution.

Regardless of the degree of replication accuracy (and mutation rate), cell populations whose initial value of P_s is less than approximately 65% cannot survive. Table 1 shows the number of generations until extinction for various combinations of P_s and F_R.

This suggests the possibility of a “frozen accident” at the origin of viable cells, analogous to one theory of the origin of the genetic code (Crick 1968, Tamura 2016). While millions of cells could have formed with low to moderate survival probability, which would give rise to a few generations of viable offspring before the population went extinct, the accidental formation of a cell with high survival probability (in other words with efficient metabolic activity and homeostasis) could have led to further evolution, and a thriving population, even with poor replication fidelity.

The results on the relative importance of survival probability and replication fidelity have implications for the debate on whether replication or metabolism developed first during the origin of life. If we assume that cell survival depends on many metabolic and non-genetic functions such as

energy usage, homeostasis, biosynthesis, cellular cohesion, etc., then the requirement for a minimum level of cell survival greater than 50% probability, regardless of replication fidelity, strongly indicates a metabolism-first scenario. It is also clear that at low levels of cell survival probability (due, for example, to incomplete or inefficient metabolic power), improvements in replication fidelity have little or no bearing on the overall fitness of the cell as represented by the population growth rate, K.

On the other hand, the data indicate that increased replication fidelity is a strong advantage for cells with moderate to high P_s, and F_R could be expected to evolve to higher values within populations that have reached positive growth dynamics.

Given these conclusions, it could be profitable for origin of life researchers to focus on highly stable metabolic systems, regardless of their initial simplicity, rather than on the evolution of replication systems within early cells with questionable degrees of survival probability.

The data shown in Figure 7 most directly indicate the discontinuity (saltation) in the evolution of increasing K as a function of P_s. The reason

that this relationship follows the simple inverse square law (eq. 7) is not clear. Derivation of this relationship from eq. 5 is not possible, and it appears to be an intrinsic property of the model system using survival probability to determine overall fitness of early life forms. In classical physics, the inverse square law applies to forces as a function of distance, whereas in this case, “distance” (D_m) is the difference between two independent variables, $F_{R,1}$ and $F_{R,2}$. If anything, the result is closer to the central limit theorem, where a probability is a function of an inverse square root. Further clarification of the significance of the empirically determined function given in eq. 7 is beyond the scope of the present investigation.

Specific biochemical mechanisms are not addressed in this report, which deals only with the statistical features of binary reproduction, such as the impact of the probability that a given cell will survive in its environment to the time or reproduction, and the probability that such reproduction will involve high fidelity of replication of the original cellular phenotype (which includes its survival probability). The results therefore are not applicable to the specific pre-biotic and early biotic processes that had to develop in order to produce a living cell, other than the survival and replication fidelity characteristics of such processes.

Several possible scenarios involving metabolic cycles in pre-biotic chemistry that may have given rise to living cells have been proposed (Peretó 2012, Muchowska et al. 2019). While much emphasis has been devoted to the initial development and evolution of replication mechanisms involving autocatalytic sets (Hordijk 2016, Vasas et al. 2012), the present results can be applied to any replicative system, including an RNA-world paradigm of a collection of individual self-replicating, catalyzing ribozymes; the central information storage and translation system of DNA and protein synthesis; or any other, more primitive system for replication of cellular components and characteristics.

The model system used to generate the data presented here is based on a deliberately simplified model of self-replication that ignores a great deal of current knowledge concerning putative cell growth in early life. It is also based on a number of assumptions that might not be borne out once more information on the transition between prebiotic chemistry and the origin of living cells is further elucidated. This includes the assumption that the earliest living cells had a lower probability of survival and a lower fidelity of replication than that seen in modern cells. Other assumptions include

binary reproduction, and even that cell reproduction was truly a feature of early life forms; that survival probability is largely a function of cell phenotype; and that mutations can directly affect survivability.

However, it should be noted that no assumptions were made on the relationship between cellular complexity and survivability. While the results suggest that low survival probability precludes evolution in a continuous stepwise manner, it says nothing about evolution from simple to more complex early life forms, assuming that biochemical complexity is not tightly linked to survival probability. There is, in fact, no *a priori* reason to believe that simple life forms could not have high innate survivability. In that case, further evolution of simple forms would be possible even if initial replication fidelity is low.

5. REFERENCES

Carter CW, Jr. 2016. An Alternative to the RNA World. *Natural history*, *125*(1), 28–33.

Cech TR. 2012. The RNA worlds in context. *Cold Spring Harbor Perspectives in Biology, 4* (7): a006742. doi:10.1101/cshperspect.a006742

Coate JE, Doyle JJ. 2011. Divergent evolutionary fates of major photosynthetic gene networks following gene and whole genome duplications. *Plant Signal Behav*., 6:594-597.

Crick FHC. 1968. The origin of the genetic code. *Journal of Molecular Biology,* 38:367-379.

Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol*., 3:E314.

Edelmann P, Gallant J. 1977. On the translational error theory of aging. *Proc Natl Acad Sci, 74* (8) 3396-3398. https://doi.org/10.1073/pnas.74.8.3396

Eigen M. 2002. Error catastrophe and antiviral strategy. *Proc Natl Acad Sci, 99*(21): 13374–13376. doi: 10.1073/pnas.212514799

Eldredge N, Gould SJ. 1972. "Punctuated equilibria: an alternative to phyletic gradualism" in Schopf TJM (ed) *Models in Paleobiology.* Freeman Cooper and Co. San Francisco. pp 82 -115.

Fontana W, Schuster P. 1998. Continuity in Evolution: On the Nature of Transitions. *Science*, 280:1451-1455. doi: 10.1126/science.280.5368.1451

Gabora L. 2006. Self-other organization: Why early life did not evolve through natural selection. *Journal of Theoretical Biology,* 241:443-450. https://doi.org/10.1016/j.jtbi.2005.12.007

Goel NS, Yčas M. 1975. The error catastrophe hypothesis with reference to aging and the evolution of the protein synthesizing machinery. *Journal of*

Theoretical Biology, 55:245-282. https://doi.org/10.1016/S0022-5193(75)80118-4

Hordijk W. 2016. Evolution of autocatalytic sets in computational models of chemical reaction networks*. Origins of Life and Evolution of Biospheres*, *46*(2):233–245.

Keightley PD, Lynch M. 2003. Toward a Realistic Model of Mutations Affecting Fitness. *Evolution, 57*(3), pp. 683–685.

Koonin EV. 2007. The Biological Big Bang model for the major transitions in evolution. *Biology Direct,* 2:21. https://doi.org/10.1186/1745-6150-2-21

Laland KN, Uller T, Feldman MW, Sterelny K, Müller GB, Moczek A, Jablonka E, Odling-Smee J. 2015. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc. R. Soc. B, 282*: 20151019. http://dx.doi.org/10.1098/rspb.2015.1019

Lee MSY, Soubrier J, Edgecombe GD. 2013. Rates of Phenotypic and Genomic Evolution during the Cambrian Explosion. *Current Biology, 23*:1889-1895.

Levin SR, West SA. 2017. The evolution of cooperation in simple molecular replicators. *Proc. R. Soc. B, 284*: 20171967. http://dx.doi.org/10.1098/rspb.2017.196710

Margulis L.1996. Archaeal-eubacterial mergers in the origin of Eukarya: phylogenetic classification of life. *Proceedings of the National Academy of Sciences of the United States of America, 93*(3): 1071–1076. doi:10.1073/pnas.93.3.1071

Martin WF, Garg S, Zimorski V. 2015. Endosymbiotic theories for eukaryote origin. *Philos Trans R Soc Lond B Biol Sci*., *370*(1678). doi: 10.1098/rstb.2014.0330

Mathworks. 2015. MATLAB Version 8.6.0.267246 (R2015b). Natick, MA.

Minelli A, Chagas A, Edgecomb GD. 2009. Saltational evolution of trunk segment number in centipedes. *Evolution & Development, 11*:3, 318–322. doi: 10.1111/j.1525-142X.2009.00334.x

Morris SC. 1989. Burgess Shale Faunas and the Cambrian Explosion. *Science,* 246:339-346.

Muchowska KB, Varma SJ, Moran J. 2019. Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature,* 569:104-107.

Peretó J. 2012. Out of fuzzy chemistry: from prebiotic chemistry to metabolic networks. *Chem. Soc. Rev*., *41*, 5394-5403.

Preiner M, Asche S, Becker S, Betts HC, Boniface A, Camprubi E, Chandru K, Erastova V, Garg SG, Khawaja N, Kostyrka G, Machné R, Moggioli G, Muchowska KB, Neukirchen S, Peter B, Pichlhöfer E, Radványi Á, Rossetto D, Salditt A, Schmelling NM, Sousa FL, Tria FDK, Vörös D, Xavier JC. 2020. The Future of Origin of Life Research: Bridging Decades-Old Divisions. *Life*,*10*, 20.

Raggi L, Bada JL, Lazcano A. 2016. On the lack of evolutionary continuity between prebiotic peptides and extant enzymes. *Phys. Chem. Chem. Phys*.,*18*, 20028-20032.

Robertson MP, Joyce GF. 2012. The origins of the RNA world. *Cold Spring Harbor Perspectives in Biology*, *4* (5): a003608. doi:10.1101/cshperspect.a003608.

Simons AM. 2002. The continuity of microevolution and macroevolution. *Journal of Evolutionary Biology, 15*: 688-701. doi:10.1046/j.1420-9101.2002.00437.x

Summers J, Litwin S. 2005. Examining the Theory of Error Catastrophe. *Journal of Virology,* 80:20-26; DOI: 10.1128/JVI.80.1.20-26.2006.

Szathmáry E, Smith JM. 1997. From Replicators to Reproducers: The First Major Transitions Leading to Life. *Journal of Theoretical Biology,* 187: 555-571.

Szostak JW. 2012. The eightfold path to non-enzymatic RNA replication. *Journal of Systems Chemistry*, 3:2. http://www.jsystchem.com/content/3/1/2.

Tamura K. 2016. The Genetic Code: Francis Crick's Legacy and Beyond. *Life, 6*(3), 36. https://doi.org/10.3390/life6030036

Theissen G. 2009. Saltational evolution: hopeful monsters are here to stay. *Theory Biosci.,* 128:43–51. doi: 10.1007/s12064-009-0058-z

Vasas V, Fernando C, Santos M, Kauffman S, Szathmáry E. 2012. Evolution before genes. *Biology Direct, 7*(1):1.

Wolf YI, Koonin EV. 2007. On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biology Direct,* 2007, 2:14.

Wolpert L. 1994. The evolutionary origin of development: cycles, patterning, privilege and continuity. *Development,* 1994: 79-84.

Xavier JC, Hordijk W, Kauffman S, Steel M, Martin WF. 2020. Autocatalytic chemical networks at the origin of metabolism *Proc. R. Soc. B.,* 287 20192377. http://dx.doi.org/10.1098/rspb.2019.2377

Figure 1. Minimal values of survival probability (P_s) required for population growth (K >1) as a function of replication fidelity (F_R). Very low variation in simulation results were seen, with average SD equal to about 0.7% of mean values.

Figure 2. The relationship between K and P_s at different constant values of F_R. A) F_R = 0, B) F_R = 0.5, C) F_R = 1.0

Figure 3. The lack of a clear relationship between K and replication fidelity (F_R) at low Ps (= 0.1). This is in contrast to the results at higher P_s (Figures 4 and 5).

Figure 4. K as a function of F_R and P_s. A) from P_s = 0.5 to 1.0, with each line being a different constant value of F_R from 0 (the lowest line) to 1.0 (the uppermost line) with intervals of 0.1. B) from P_s = 0.1 to 0.4 with the same lines for F_R as shown in A).

Figure 5. The relationship between K and F_R at specific values of P_s. A) P_s = 1.0, B) P_s = 0.5, C) P_s = 0.3.

Figure 6. Theoretical and experimental minimal values of survival probability (P_s) required for population growth as a function of replication fidelity (F_R)

Figure 7. Average minimal distance (D_m) between fidelity values for significant differences in growth rate as a function of P_s

TABLE 1

Some simulation determinations of K and Time to Extinction (TTE) for various values of P_s and F_R. Initial cell counts (C_{t0}) = 500

P_s	F_R	K	TTE
0.1	0	0.0032	1.08
0.1	0.5	0.021	1.61
0.1	1	0.016	1.51
0.5	0	0.40	6.72
0.5	0.5	0.57	11.1
0.5	1	0.73	19.4
0.7	0	0.63	13.5
0.7	0.5	0.91	69.4
0.7	0.6	0.97	190
0.7	0.7	1.0	SURVIVAL
1	0	1.1	SURVIVAL
1	0.5	1.6	SURVIVAL
1	1	2.0	SURVIVAL

TABLE 2

Representative calculations of D_m from statistical analysis of K as a function of F_R at a constant $P_S = 0.5$

F_R	Mean K	SD	$F_{R,1}$	$F_{R,2}$	p
0.03	0.4	0.01	0.03	0.04	NS
0.04	0.41	0.01	0.03	0.05	NS
0.05	0.408	0.01	0.03	0.06	0.001*
0.06	0.42	0.01	0.03	0.07	0.001
0.07	0.416	0.007	0.03	0.08	0.01
0.08	0.416	0.01	0.03	0.09	0.001
0.09	0.426	0.007	0.03	0.1	0.001
0.1	0.428	0.011			

$D_m = F_{R,2} - F_{R,1} = 0.06-0.03 = .03$

0.94	0.722	0.011	0.94	0.95	NS
0.95	0.717	0.01	0.94	0.96	NS
0.96	0.725	0.015	0.94	0.97	NS
0.97	0.731	0.013	0.94	0.98	0.05*
0.98	0.732	0.009	0.94	0.99	0.01
0.99	0.738	0.012	0.94	1	0.001
1	0.742	0.011			

$D_m = F_{R,2} - F_{R,1} = 0.98-0.94 = .04$

***Minimal distance between values of F_R that show a statistically significant difference in K

Figure 1

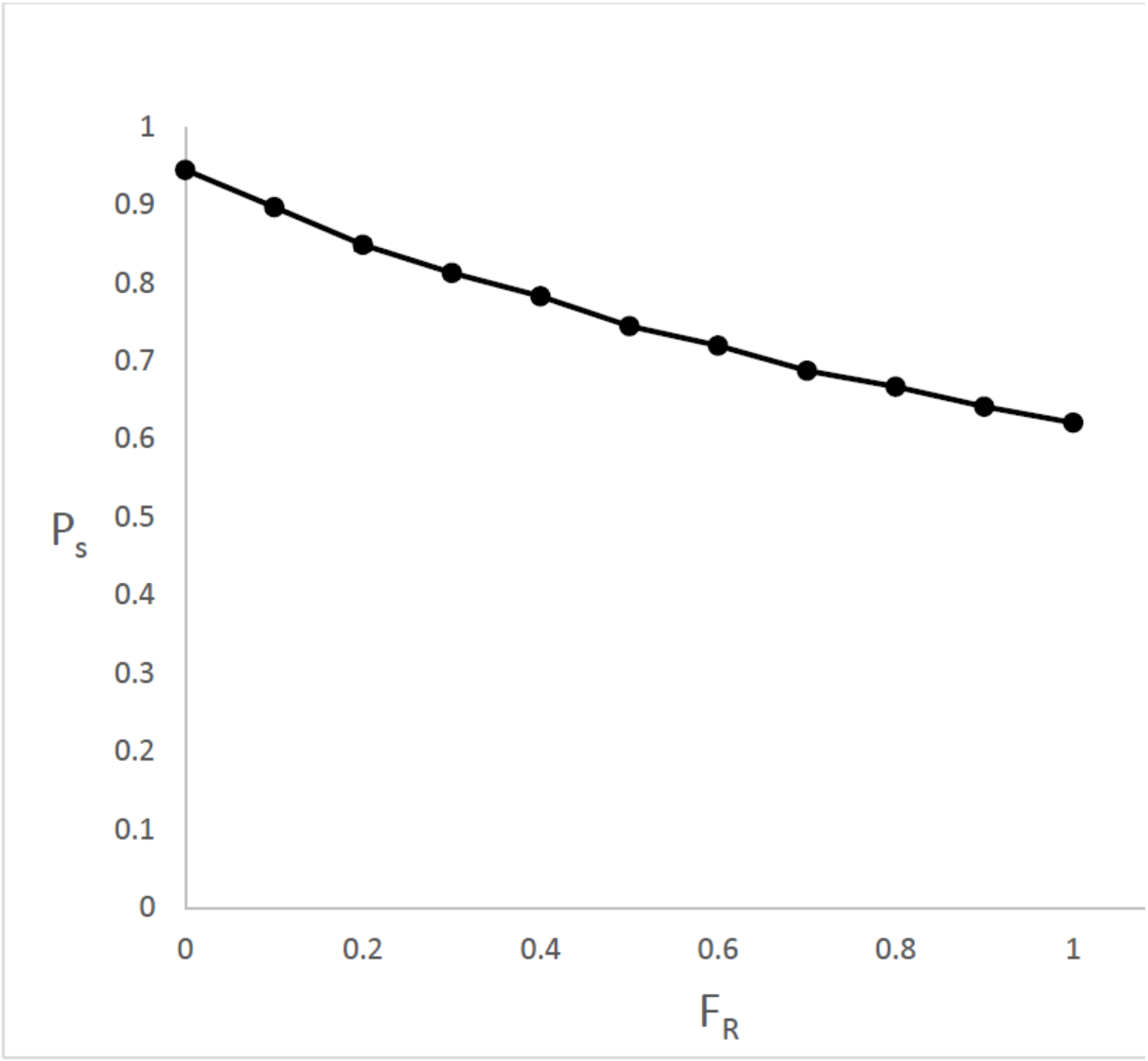

Figure 2

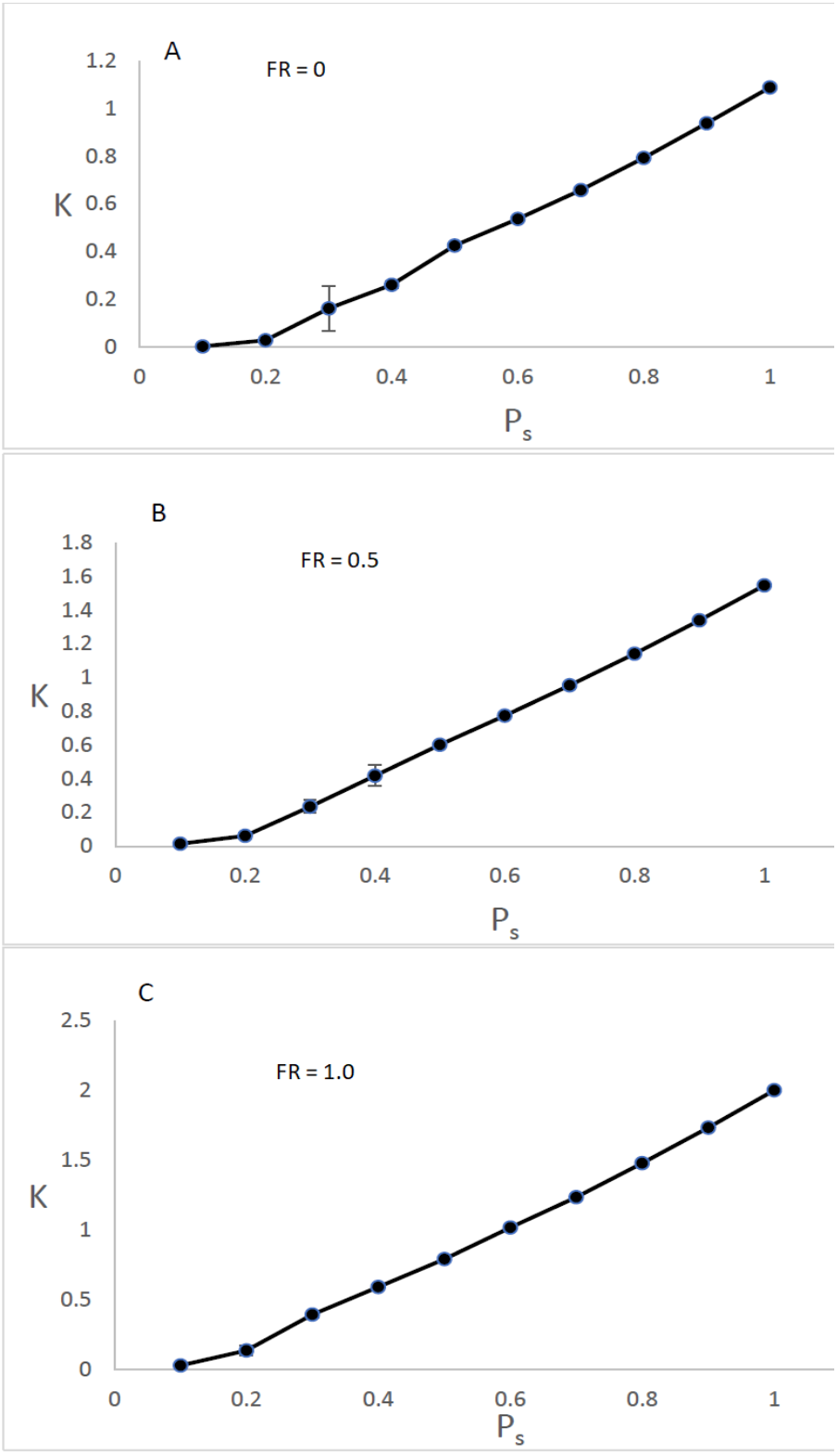
A
FR = 0
K
1.2
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
Ps
B
FR = 0.5
K
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
Ps
C
FR = 1.0
K
2.5
2
1.5
1
0.5
0
0
0.2
0.4
0.6
0.8
1
Ps

Figure 3

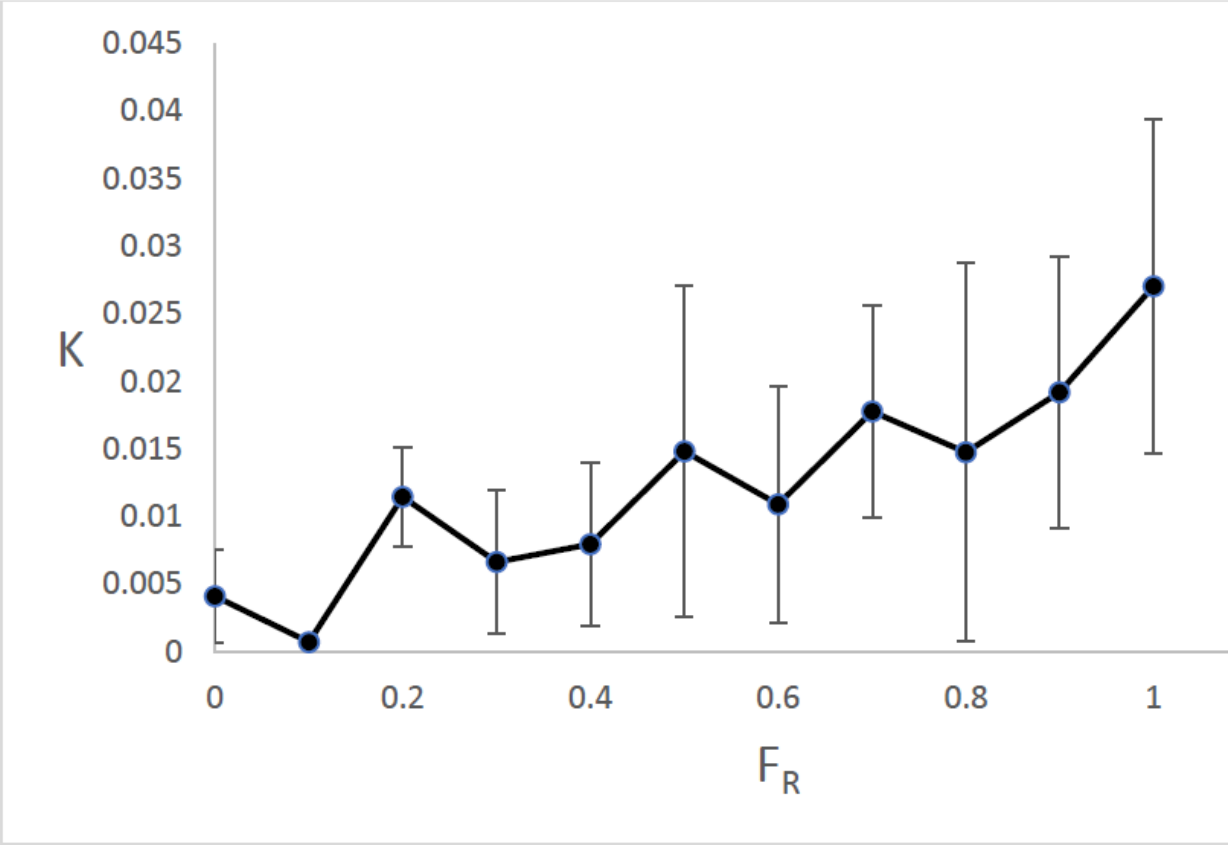

Figure 4

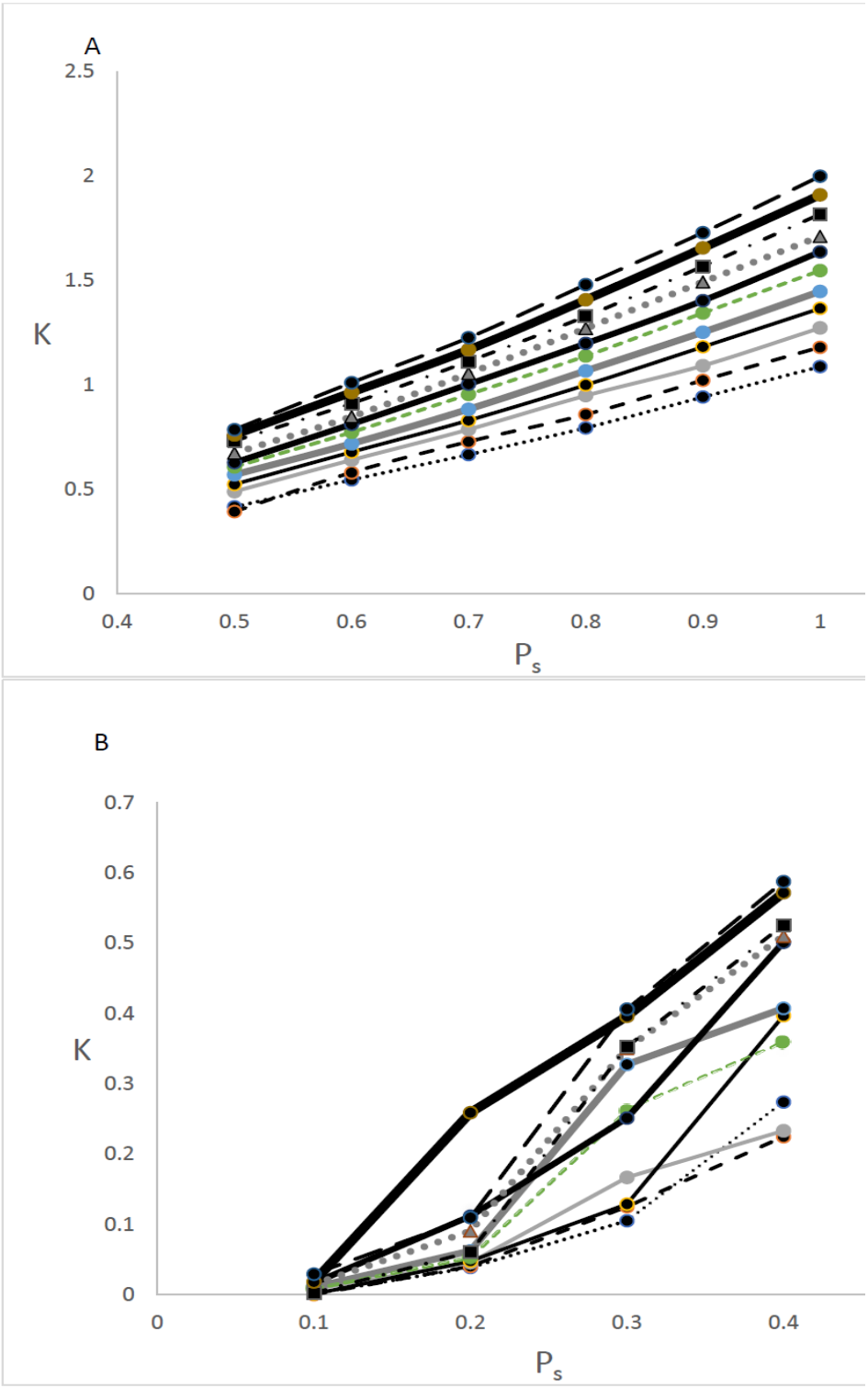
A
2.5
2
1.5
K
1
0.5
0
0.4
0.5
0.6
0.7
0.8
0.9
1
Ps
B
0.7
0.6
0.5
0.4
K
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
Ps

Figure 5

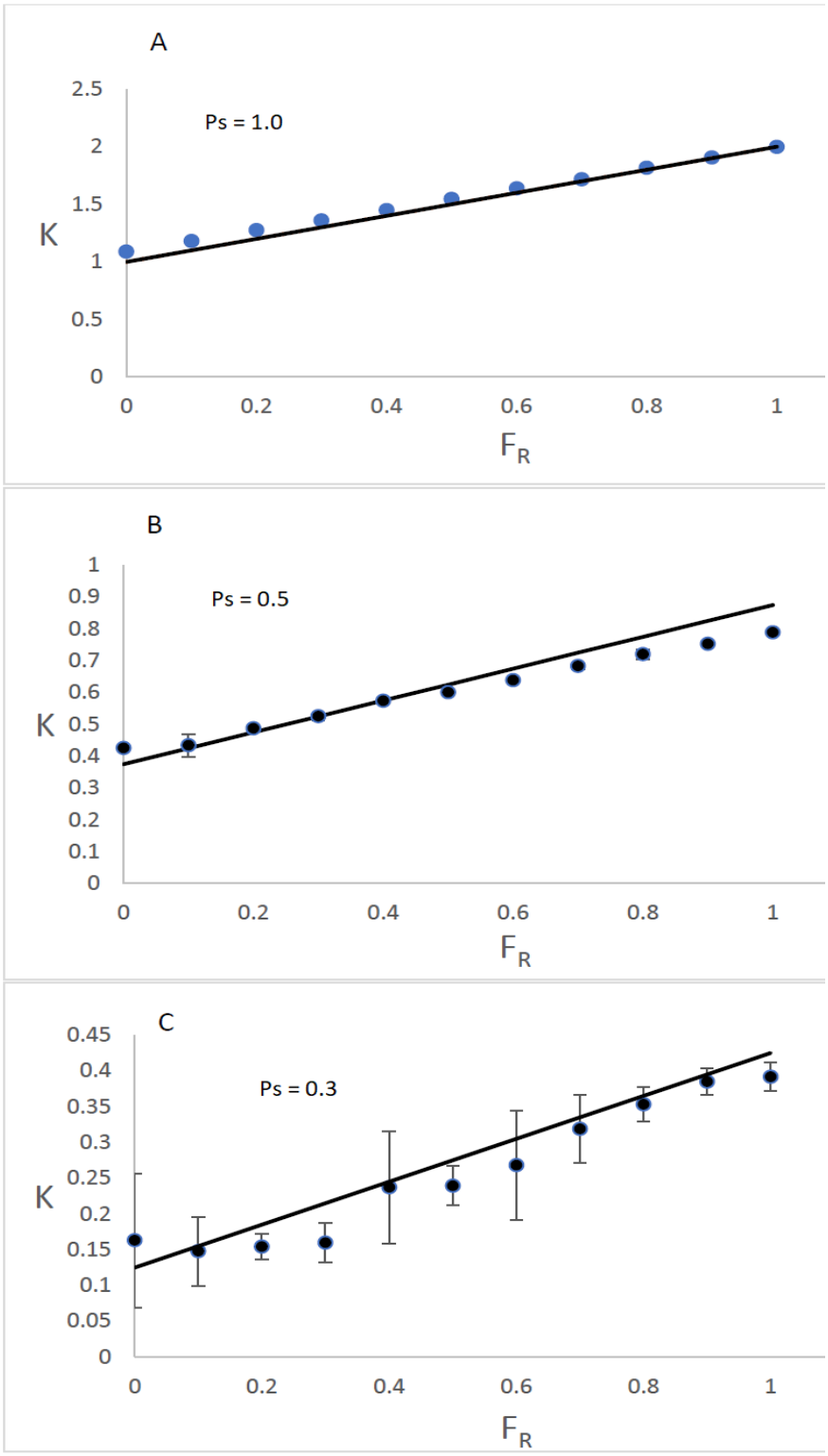
A
Ps = 1.0
K
2.5
2
1.5
1
0.5
0
0
0.2
0.4
0.6
0.8
1
F_R
B
Ps = 0.5
K
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.2
0.4
0.6
0.8
1
F_R
C
Ps = 0.3
K
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.2
0.4
0.6
0.8
1
F_R

Figure 6

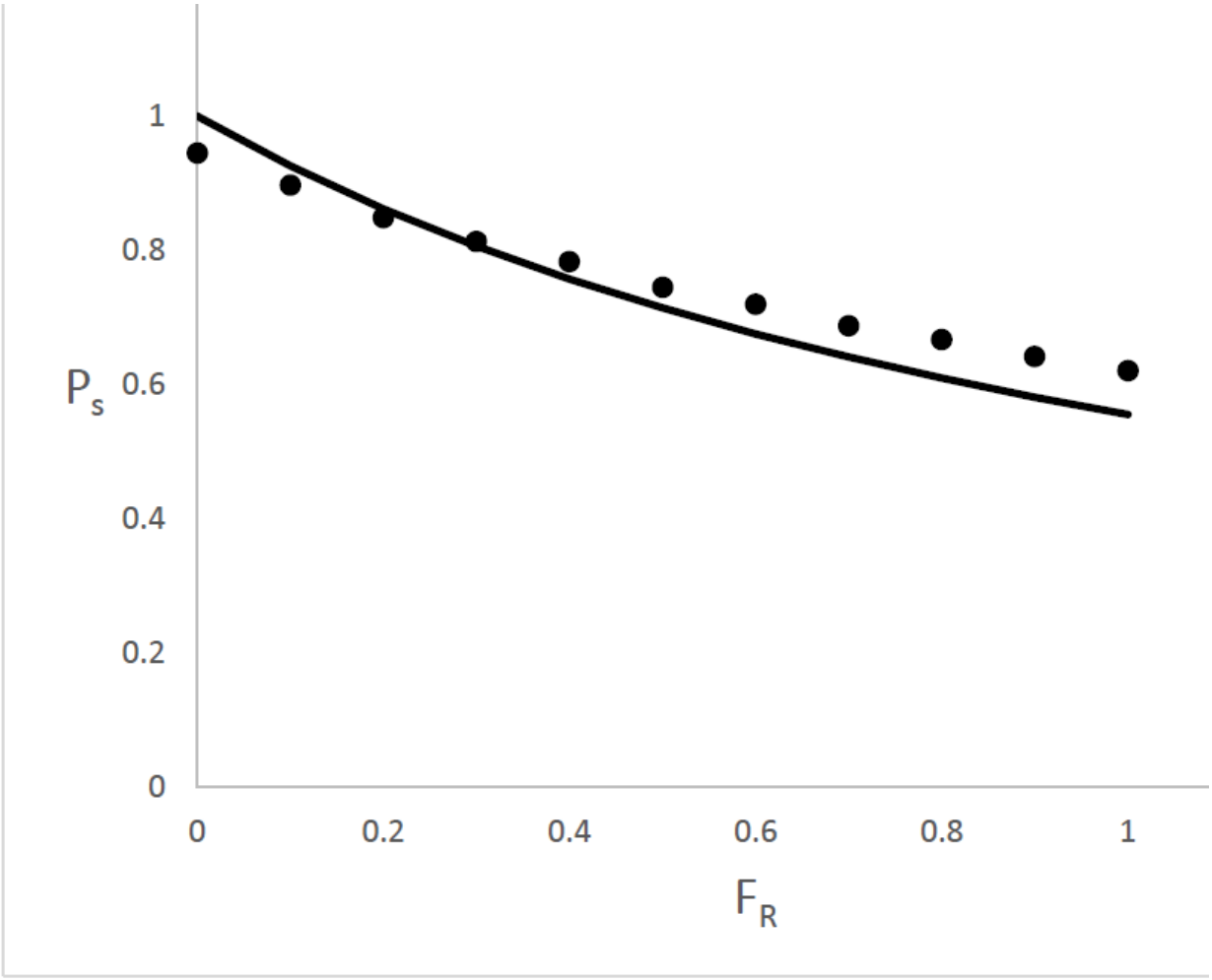

Figure 7