ISSN: 1984-7688

COMPLEX DYNAMICS ON HARDY-WEINBERG EQUILIBRIUM WITH MUTATION: A MATHEMATICAL MODEL

Marcos Teixeira Godinho*; **Sara Dalia Barbosa**

[1]Instituto de Ciências Biológicas, Centro Universitário Una, Belo Horizonte, MG, Brasil

*`godinho@prof.una.br`

Received: 23/02/2011 - Approved: 03/10/2011 - Available: 29/12/2011

ABSTRACT: This paper shows that a small modification of the Hardy-Weinberg law will lead to a completely different equilibrium. This new equilibrium is chaotic, in a mathematical sense. It is well known that dynamic difference or differential equation models, which are characterized by nonlinearity, can present chaotic properties over nontrivial ranges of values of their parameters. Chaotic orbit, besides its deterministic behavior, is undistinguishable from a random process, or rather, a process perturbed by a random shock. Despite its complexity, in a certain range, the model is stable, presenting a rich variety of behavior. The key point is that we do not need to attribute to a random walk or statistical model or to a drift as an external force to act upon the population.

KEYWORDS: Hardy-Weinberg Law, Chaotic Dynamic, Logistic Map

INTRODUCTION

A very important result and the starting point for any discussion in population genetics, the Hardy-Weinberg Law, tells us that in an indefinitely large population, the relative frequencies of allelomorphic genes remain constant over time, and the genotypic frequencies are related to allele frequencies if unaffected by disturbing factors, *e.g.*, mutation, migration or selection. The stability of the equilibrium, whatever the initial combination of the population, was proved by Hardy (1908), Weinberg (1908), and Robbins (1918), but that does not mean that equilibrium is attained instantaneously, but is rather approached asymptotically through an infinite number of generations (WRIGHT, 1930).

With this as base, one can incorporate the effects of any disturbing factors into the Hardy-Weinberg model, and see how these factors bring about genetic change in populations, and that has been accomplished throughout the years. Starting with the concept of genetic drift (WRIGHT, 1930), we have seen how fast the stochastic models flourished in the following decades. Gillespie (2000) introduced the concept of pseudo hichhiking, or genetic draft, which behaves as a stochastic force similar to the genetic drift, the difference laying in the dependence on recombination and selection.

Evolutionary forces are often divided into two sorts: stochastic and deterministic (WRIGHT, 1955, GILLESPIE, 2001). Genetic drift is considered to be the most important of the stochastic forces in the evolution of natural populations (GILLESPIE, 2001). That is why Gillespie's paper considered to what extent pseudo hitchhiking led to similar predictions for the change in allele frequency. Regarding the deterministic forces it was always not appealing, mainly due to its simplicity and not very profound conclusion as one have with the stochastic models.

This paper aims at showing that a deterministic model of population genetics, following Hardy-Weinberg Law, can generate complex trajectories mimicking the stochastic models. If it is so, it becomes too complicated to tell one force – deterministic – from the other – stochastic.

The next section shows how deterministic mutation is treated in the literature. The following section defines the new model. The full mathematical treatment of its dynamics is presented afterward. The last section discusses the mathematical model and includes some numerical examples.

Process that Change Allelic Frequencies – Mutation

This section is based on Tamarin (2001), Gillespie (2004) and Nei (1975).

Hardy-Weinberg Equilibrium

In 1908 Hardy and Weinberg discovered a rule that relates allelic and genotypic frequencies in a population of diploid, sexually reproducing individuals, if that population has random mating, large size, no mutation or migration, and no selection. The rule has three aspects: The allelic frequencies at an autosomal locus in a population will not change from one generation to the next; The genotypic frequencies are determined in a predictable way by the allelic frequencies; and, The equilibrium is neutral. In a population of individuals segregating the *A* and *a* alleles at the *A* locus, each individual will be one of three genotypes: *AA*, *Aa* or *aa*. If $p = f(A)$ and $q = f(a)$, then it is possible to predict the genotypic frequencies in the next generation. Thus,

$$\begin{aligned} &f(AA) = p^2 \\ &f(aa) = q^2 \\ &f(Aa) = 2pq \\ &p^2 + 2pq + q^2 = (p+q)^2 = 1 \end{aligned}$$

Using the above results it is possible to demonstrate that the allelic frequencies do not, change generation after generation.

Mutation Pressure

At this point it is necessary to discuss some of the effects of violating, or relaxing, the assumptions of the Hardy-Weinberg equilibrium. Here we consider the effects of mutation on the model equilibrium. Thus, the following steps should be taken in order to solve for equilibrium in population genetics models. The steps are: Set up an mathematical model; Calculate allelic frequency in the next generation; Calculate change in allelic frequency between generations; Calculate equilibrium condition; and determine the stability of the equilibrium.

Mutation is any process that alters one allele to another and thus changing allelic and genotypic frequencies. Irreversible mutation of a gene at the rate *u* per generation change frequency *q* at the rate $\Delta q = -uq$. With reverse mutation at rate *v* the change in gene frequency is $\Delta q = -uq + v(1-q)$. In the absence of other pressures, equilibrium is reached between the two mutation rates when $\Delta q = 0$, giving $\hat{q} = \dfrac{v}{u+v}$. Once the Hardy-Weinberg equilibrium is a neutral one, it is necessary to prove if this equilibrium, $\hat{q}$, is a stable one. Thus, using the previous results, we can easily get $\Delta q = (u+v)(\hat{q} - q)$, knowing that $u > 0$ and $v > 0$, the sign of Δq depending only on $(\hat{q} - q)$. When $q < \hat{q}$, it implies that $\Delta q > 0$, and *q* must rise, and when $q > \hat{q}$, it implies that $\Delta q < 0$, and, *q* must decrease.

Hardy-Weinberg equilibrium with chaotic mutation dynamic

Consider a population of *N* individuals in which two alleles, *A* and *a,* exists. Every individual will be one of three genotypes: *AA*, *Aa* or *aa*. If $p = f(A)$ and $q = f(a)$, and *A* mutates to *a* at a rate of *u,* and *a* mutates back to *A* at a rate of *v,* then

$$f(A) = p = \frac{P_A}{P_N}$$
$$f(a) = q = \frac{P_a}{P_N}$$

If p_t and q_t are the frequencies of A and a in generation *t.* It is well known that the frequency in generation $t+1$, q_{t+1}, is given by

$$q_{t+1} = q_t + up_t - vq_t \qquad (1)$$

The change in allelic frequency between generations, $\Delta q = q_{t+1} - q_t$, is

$$\Delta q = u - (u+v)q_t \qquad (2)$$

If $\Delta q = 0$, we can compute the equilibrium point and obtain

$$\hat{q} = \frac{v}{u+v} \text{ and } \hat{p} = \frac{u}{u+v} \qquad (3)$$

Now, using the theory of first-order differential equations, it is possible to study the dynamic of the model, and one can obtain

$$q_t = \hat{q} + (q_0 - \hat{q})(1-u-v)^t \qquad (4)$$

and its solution is

$$q_t = \hat{q} + (q_0 - \hat{q})e^{-t(u+v)} \qquad (5)$$

where q_0 is the initial value, and q_t shows the behavior of q_t over time.

Equations (1)-(5) shows that the frequency is stable along the generations when u and v are fixed. If we assume that $P_A = 0$ or $P_a = 0$ cannot occur, then the model should incorporate some kind of inhibiting factor, φ, since *u* and *v* are no longer fixed for they are dependent on the population size and on how close P_A and P_a are to P_N. In the equations below we present a mathematical version of this relationship of u and v with P_A, P_a and P_N.

$$u_t = \varphi(P_{N,t} - P_{A,t}) \qquad (6)$$
$$v_t = \varphi(P_{N,t} - P_{a,t}) \qquad (7)$$

Now we must rewrite equation (1) in order to capture this new dynamic

$$q_{t+1} = q_t + u_t p_t - v_t q_t \qquad (8)$$

A little simplification on equation (8) could make things easier. The second and third terms can be captured by the mutation rates, u_t and v_t, that are not fixed as before. So we can rewritten (6), (7) and (8) to get

$$P_{A,t+1} = \varphi(P_{N,t} - P_{A,t})P_{A,t} \qquad (9)$$
$$P_{a,t+1} = \varphi(P_{N,t} - P_{a,t})P_{a,t} \qquad (10)$$

Equations (9) and (10) show the new law of movement for calculating the allelic frequency. Since there is an obvious analogy between equations (9) and (10), from now on we will consider only equation (10). Dividing (9) and (10) for $P_{N,t}$ and making the substitution $X_t = \frac{P_t}{P_{N,t}}$ we get

$$X_{a,t+1} = \varphi_t P_{a,t}\left[1 - X_t\right] \qquad (11)$$

If $\alpha_t = \varphi_t P_{N,t}$, and substituting this variable into equation (11), we obtain

$$X_{a,t+1} = \alpha_t X_{a,t}\left[1 - X_{a,t}\right] \qquad (12)$$

In this form, equation (12) can be recognized as the well known the logistic map. The dependence of the parameter α_t depends on $P_{N,t}$ shows that mutation rates are not constant along the generations. In fact, the mutation rate may change in response to environmental changes.

Mathematical analysis of equation (12)

It is well known that dynamic difference or differential equation models, which are characterized by nonlinearity, can present chaotic properties over nontrivial ranges of values of their parameters. Chaotic orbit, besides its deterministic behavior, is undistinguishable from a random process, or rather, a process perturbed by a random shock. Despite its simplicity, equation (12), which is called the logistic map, displays remarkably rich variety of behavior as the parameter α varies. As the parameter leaves its stable range the model will move from stability to instability. In this region oscillatory behavior persists, but the time series generated by it will fail to replicate itself perfectly.

Fixed Points

Using equation (12) as the law of motion of a discrete dynamical system, it is possible to investigate the long-run behavior of all its possible orbits. But first, it is necessary to put forth the concept of fixed point. In a linear system, convergent orbits always converge to a fixed point. This property holds in general. If a continuous function f has a convergent orbit $x, f(x), \cdots, f^n(x)$ and L is its limit, then

$$f(L) = f\left(\lim_{x\to\infty} f^n(x)\right) = \lim_{x\to\infty} f^{n+1}(x) = L$$

So that L is a fixed point of f. In words, if the system converges to any finite limit, that limit must be a fixed point.

Dynamic behavior

The equation (9) that is repeated it here for convenience

$$X_{a,t+1} = \alpha X_{a,t}\left[1 - X_{a,t}\right]$$

Can be rewritten in the following way

$$X_{a,t+1} = f(\alpha, X_{a,t}), \qquad X_{a,t} \in [0,1], \alpha \in [0,4]$$

The function f is a unimodal curve in the $\left(X_{a,t+1}, X_{a,t}\right)$ plane, since $\frac{\partial f}{\partial X_{a,t}} = 0$ for $X_{a,t} = 1/2$, and $\frac{\partial^2 f}{\partial X_{a,t}^2} = -2\alpha < 0$. When $X_{a,t} = \frac{1}{2}$, we have $f = \frac{\alpha}{4}$, which shows that the maximum shifts upwards when α increases in its interval of definition. Equation (9) has two fixed points $X_{a,t} = 0$, the trivial solution, and the other at $\kappa(\alpha) = \frac{\alpha - 1}{\alpha}$.

The main types of dynamics can be classified according to the values of the parameter α:

Case 1: $0 < \alpha \leq 3$

In this region $X_{a,t+1}$ will converge, sometimes oscillating ($\alpha > 2$), to the fixed point, $\kappa(\alpha)$, and this point is stable. This is shown in Fig 1 in the supplementary material.

Case 2: $3 < \alpha \leq \psi$

As α increases the stable 2-period cycle loses stability, especially when it reaches the 3.44949, where a new stable 4-period cycle appears. The orbits created repeats itself to an infinite sequence of bifurcations and period-doubling, (Fig 2 in the supplementary material). The accumulation point is the Feigenbaum constant, ψ, which is defined as

$$\psi = \lim_{x\to\infty} \frac{\alpha_x - \alpha_{x-1}}{\alpha_{n+1} - \alpha_n}$$

Case 3: $\psi < \alpha \leq 4$

The 4-period cycle behavior is stable for some range of α until a value is reached at which the 4-period cycle bifurcates, and this will lead to another stable cycle, which is called the 8-period cycle, and as α continues to increase the orbit of $X_{a,t}$ will bifurcates again and again. But it will reach eventually a finite value for α, the Feigenbaum constant, ψ, at which the limit cycle becomes infinitely long, that never repeats itself. Despite all that, the orbit generated by this process is not yet considered as chaotic.

As α increases the system reaches the range chaotic behavior. That will take place as soon as α becomes greater than ψ, as illustrated by Fig. 3 and 4 (in the supplementary material).

DISCUSSION

A key property shared by all the mutation models is the assumption that the mutation rate is constant, although it has been known that mutation rate vary within a certain range (BAER, 2008; GILLESPERE, 2004, DRAKE, 1998). Evidence has shown that the mutation rate is correlated with fitness. Thus, natural selection will favor mutation rates, which are low under normal conditions but higher under stressful conditions. Evidence from *E. coli* has shed some light on this subject, showing that a stress induced mutation rate differs consistently with certain ecological circumstances (BAER, 2008; BJEDOV *et al,* 2003). The understanding of how natural selection changes allele frequencies or how it works when genotypes have different fitness is beyond the scope of this paper.

The interaction of natural selection and genetic drift has been studied in the literature. The conclusions always go to the direction that genetic drift influences the fate of rare alleles, even in very large populations. In another word, the fate of most advantageous mutations is extinction, which leads to the conclusion that evolution is essentially a stochastic process that is not repeatable or reversible (GILLESPERE, 2004). According to equation (12) it is not necessarily true. For the value the parameter α will determine the kind of trajectory the model present through the generations. It is possible to have a non repeatable and irreversible trajectory without taking into consideration a random shock.

The natural selection, fitness, can influence the trajectory of equation (12) through equation (7). While equation (10) was used to model how the environment stress will determine the mutation rate. The parameter α can be redefined using the original variables of the model. With some algebraic manipulation we get

$$\alpha_t = \frac{u_t}{q_t}$$

Or

$$\alpha_t = \frac{u_t P_{N,t}}{P_{a,t}} \qquad (13)$$

Which is a much more appealing definition for α_t. Equation (13) shows that the dynamic of the population throughout the generations will assume different trajectories depending on how the natural selection, or fitness, will change the allele frequencies from one generation to the next.

REFERENCES

BAER, C.F. Does mutation rate depend on itself? PLoS Biology v.6, v.2, p.e52, 2008. doi:10.1371/journal.pbio.0060052

BJEDOV, I.; TENAILLON, O.; GÉRARD, B.; SOUZA, V.; DENAMUR, E.; RADMAN, M.; TADDEI, F.; MATIC, I. Stress-induced mutagenesis in bacteria. Science n.300, p.1404-1409, 2003.

DRAKE, J.W.; CHARLESWORTH, B.; CHARLESWORTH, D.; CROW, J.F. Rates of spontaneous mutation. Genetics n.148, p.1667-1686, 1998.

GILLESPIE, J.H. Genetics drift in an infinite population: the pseudohitchhikig model. Genetics n.155, p.909-919, 2000.

GILLESPIE, J.H. Is the population size of a species relevant to its evolution? Evolution v.55, n.11, p.2161-2169, 2001.

GILLESPIE, J.H. Population genetics: A concise guide. The John Hopkins University press, Baltimore and London. 217pp, 2004.

HARDY, G.H. Mendelian proportion in a mixed population. Science n.28, p.49-50, 1908.

NEI, M. Molecular population genetics and evolution. North-Holland publishing company, Amsterdam. 288pp, 1975.

ROBBINS, R.B. Some applications of mathematics to breeding problems. Genetics n.3, p.375-389, 1918.

TAMARIN, R.H. Principles of genetics. The McGraw-Hill companies. 686 pp, 2001.

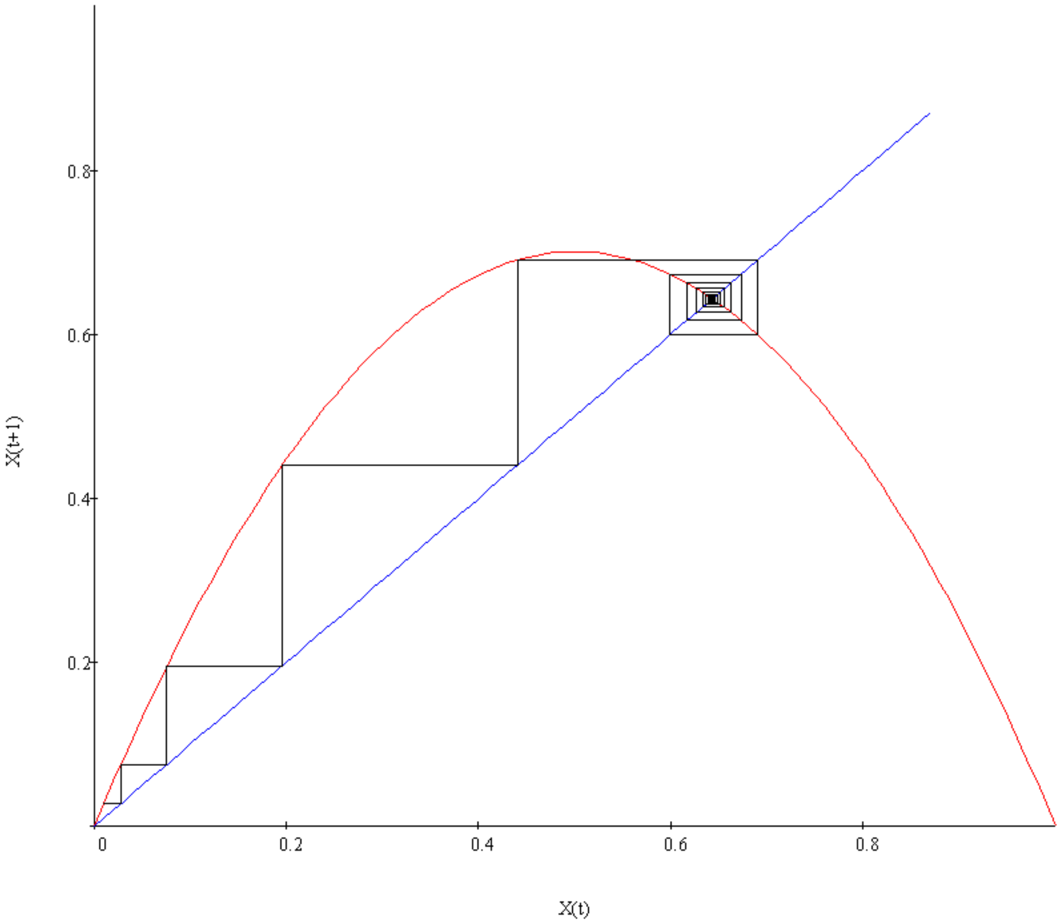
WEINBERG, W. On the demonstration of heredity in man. In: Boyer SH (ed) Papers on human Genetics. Prentice-Hall, Englewood Cliffs, pp4-15, 1963.

WRIGHT, S. Evolution in mendelian populations. Genetics n.16, p.97-158, 1930.

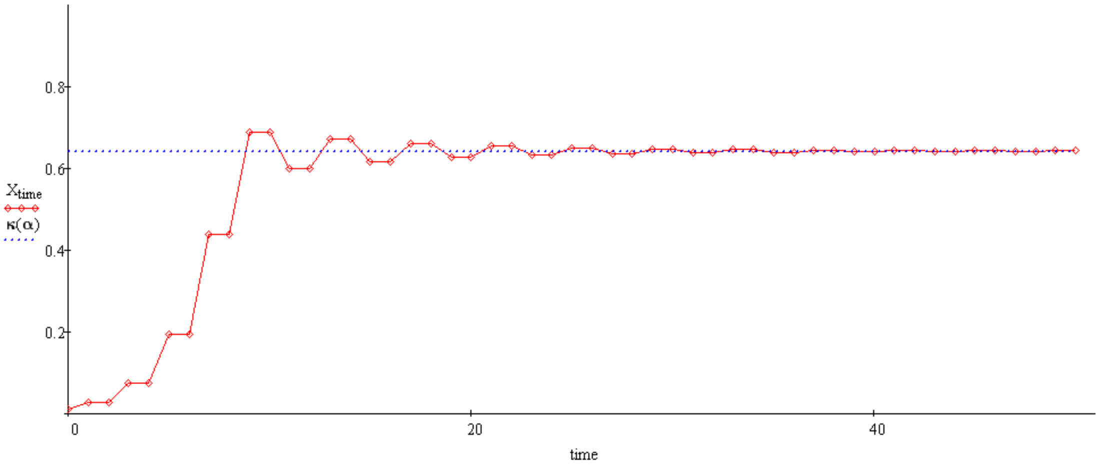
WRIGHT, S. Classification of the factors of evolution. Cold Spring Harbor Symposia on Quantitative Biology n.20, p.16-24D, 1955.

Supplementary Material

(fig. 1a) Dynamic of Equation (9) for alfa = 2.8

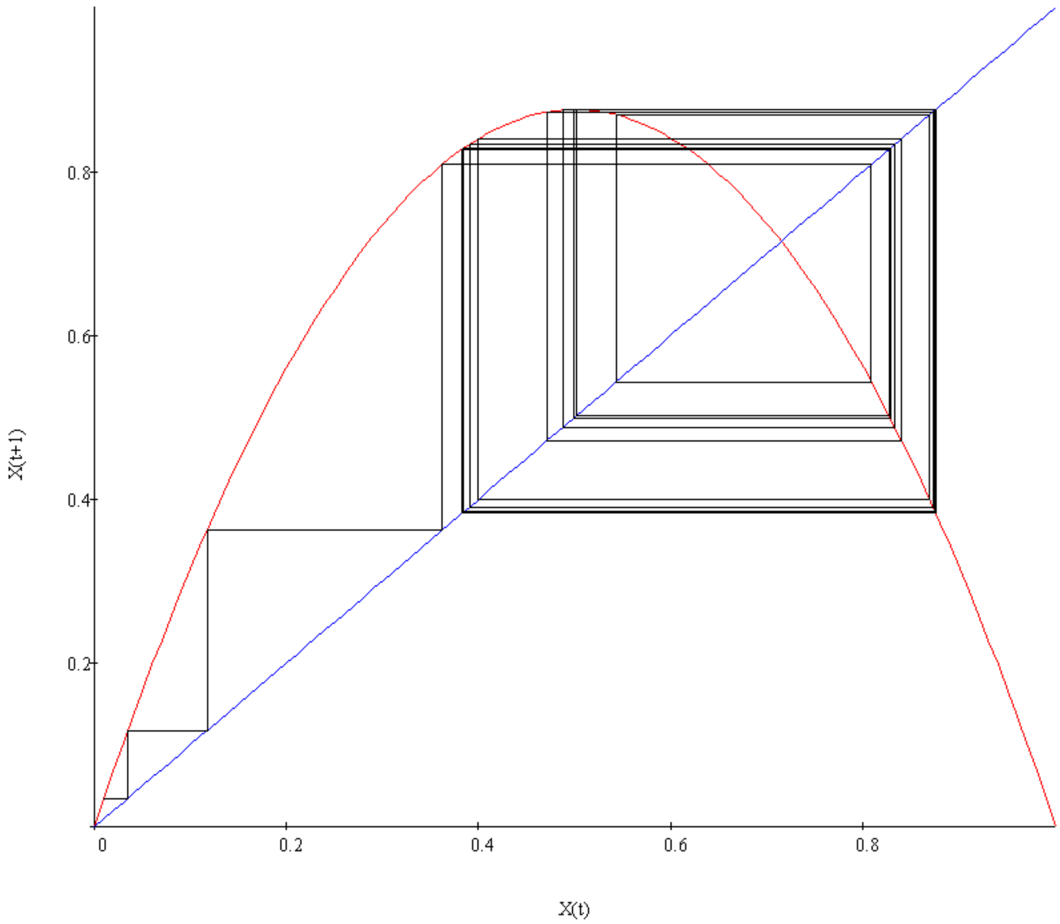


(fig 1b) Time path of mutation through the generations



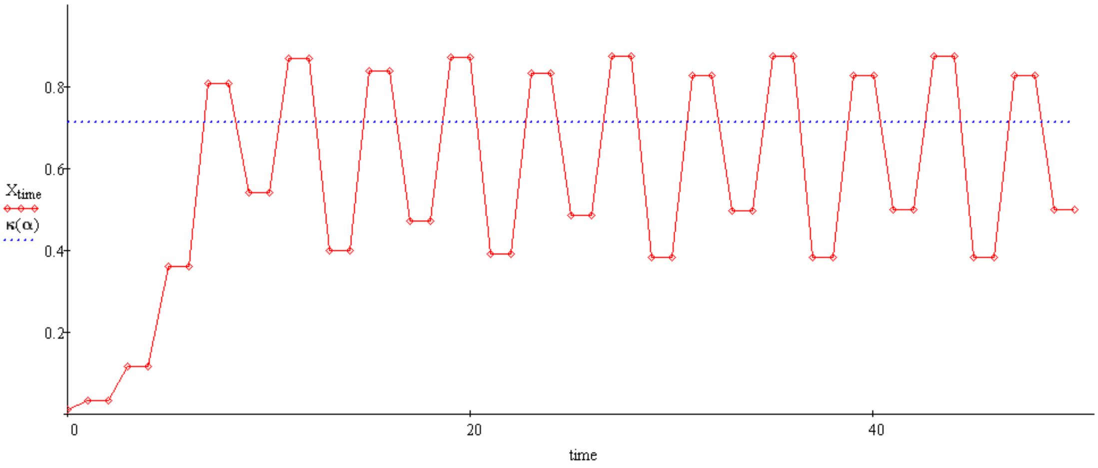
Supplementary Material

(fig. 2a) Dynamic of Equation (9) for alfa = 3.5



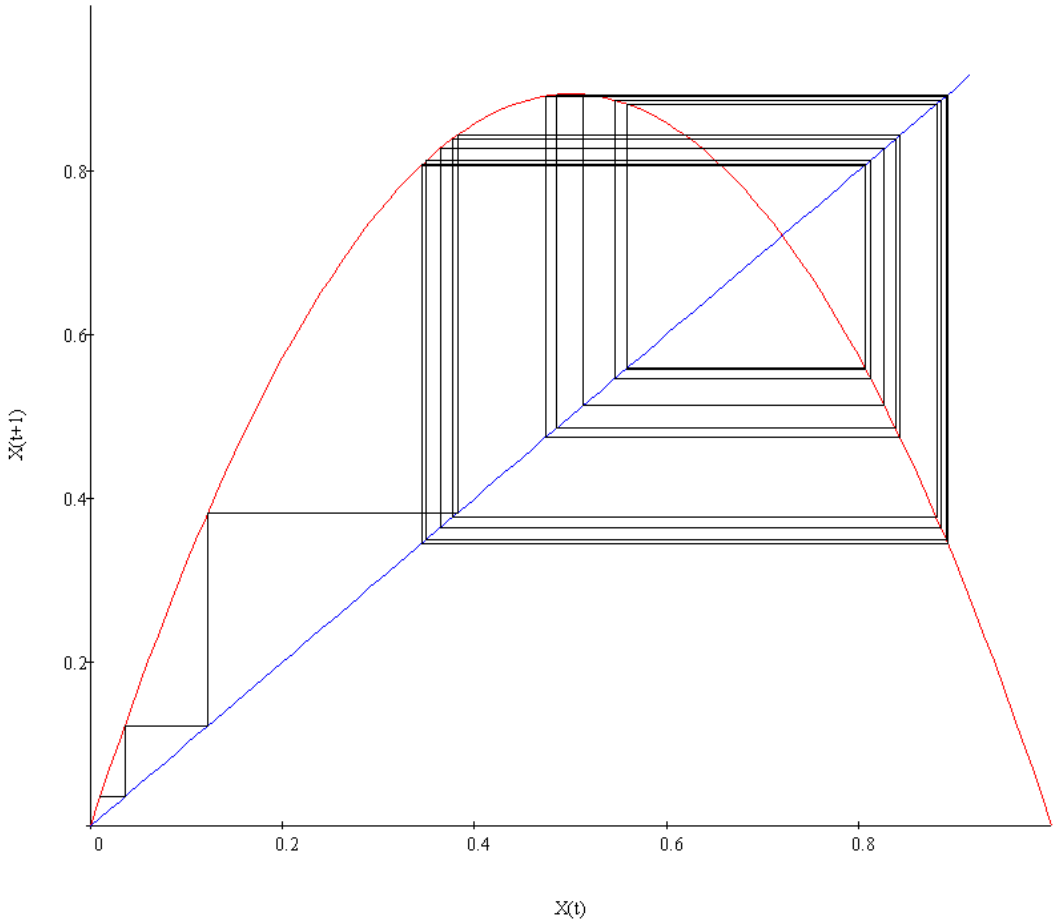
Supplementary Material

(fig 2b) Time path of mutation through the generations

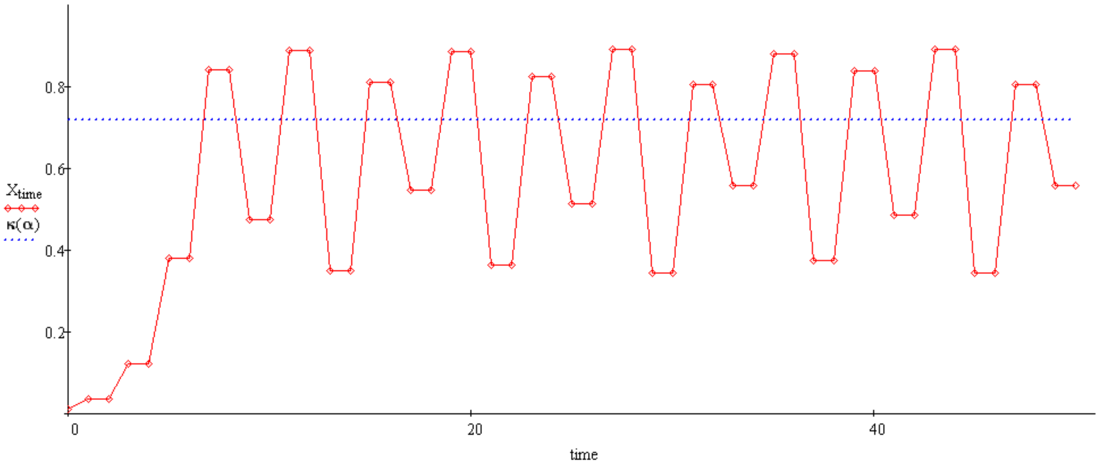


Supplementary Material

(fig. 3a) Dynamic of Equation (9) for alfa = Feigenbaum Constant

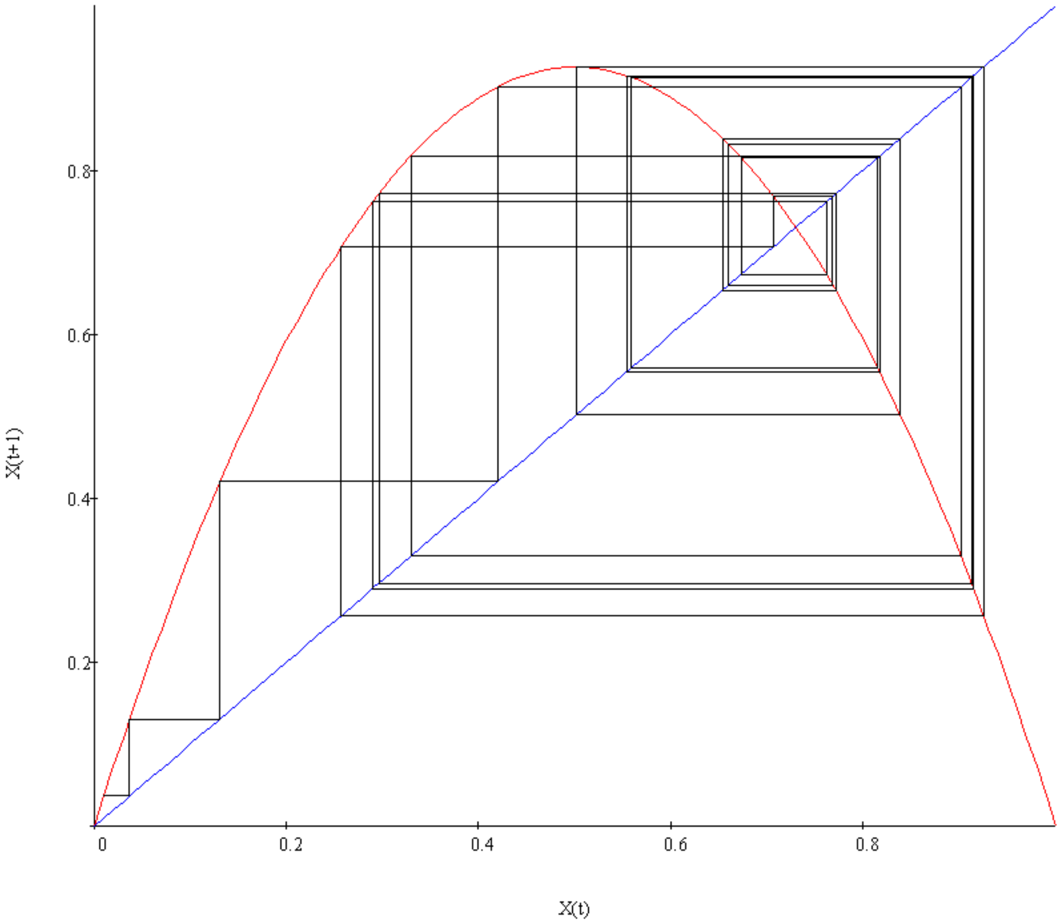


(fig 3b) Time path of mutation through the generations



Supplementary Material

(fig. 4a) Dynamic of Equation (9) for alfa = 3.7



(fig 4b) Time path of mutation through the generations