Istituto di Matematica Guido Castelnuovo and Istituto di Calcolo delle Probabilità,
University of Rome, Italy

GENE DECAY

*II. Analytic Simulation of Gene Decay**

CARLA ROSSI, RICCARDO SANGERMANO

SUMMARY

The models of gene decay, previously worked out on the basis of Gedda-Brenci's *Ergon/Chronon System*, now undergo an analytic simulation, in order to verify how the mathematical model fits the stochastic one.

INTRODUCTION

On the basis of the parameters of gene stability, suggested by Gedda and Brenci (1969) within a complete model referred to as *Ergon/Chronon System*, Rossi (1972) has worked out a mathematical and a stochastic model of gene decay. The object of the present paper is to verify, by analytic simulation, how the mathematical model fits the stochastic one. After having verified the efficiency of the mathematical model, the function of gene decay, $E[t(\alpha, \beta)]$, shall be calculated.

Reference should first be made to some previous results (Rossi 1972) and, more particularly, to the expression of the function $E[t(\alpha, \beta)]$;

$$E[t(\alpha, \beta)] = \frac{K}{\mu(\alpha)} + \frac{1}{\mu(\alpha)} [\beta - E(\beta)] \qquad [1]$$

where,

$\mu(\alpha) = qe^{-\lambda\alpha}$ is the velocity of the mutation process, i.e., the expected number of mutations per time unit (1 year);

α is the index of stability of a DNA molecule, i.e., the number of G-C pairs as third in a triplet, the probability distribution of which is normal with parameters $m = 250$ and $\sigma^2 = 125$;

β is the redundancy, i.e., the number of efficient molecules in a gene at the time $t = 0$, the probability distribution of which is a binomial one:

$$p(\beta = h) = p_h = \binom{20}{h} p^h (1 - p)^{20-h};$$

* The stochastic scheme and the logical basis of simulation are due to C. Rossi (Istituto di Matematica Guido Castelnuovo, but presently at Istituto di Calcolo delle Probabilità); the automatic procedures and the analysis of the results, to R. Sangermano (Istituto di Calcolo delle Probabilità). Both authors thank Professor B. de Finetti for his useful suggestions and criticism.

https://doi.org/10.1017/S1120962300010908 Published online by Cambridge University Press

K is the constant that we find in the expression of $E[t(\alpha, \beta)]$:

$$K = \frac{L+1}{20} \binom{20}{L+1} p^{L+1} \sum_{i=0}^{20-(L+1)} (A_i/\underline{B}_i^2) \text{ (Rossi 1972);}$$

λ, q are parameters that can be estimated knowing two values of $\mu(\alpha)$, e.g., the expectation and the maximum value;

p is the probability that a molecule is efficient at the time $t = 0$;

L is the lower limit of the number of efficient molecules, below which the gene is not active.

ANALYTIC SIMULATION OF THE PROCESS OF GENE DECAY

In order to simulate the process of gene decay, we need some important data, i.e.: the maximum value and the expectation of $\mu(\alpha)$ (Asimov 1962, Watson 1965, Burch 1968), the lower limit L (Szilard 1959, Burch 1968), and the probability p (Szilard 1959).

The expectation of the number of mutations per translation, and the upper limit of it, can be explicitly found. Multiplying this by the number of translations per year, the expectation [1] and the upper limit of $\mu(\alpha)$ are obtained:

$$E[\mu(\alpha)] = 0.42; \qquad U.L.[\mu(\alpha)] = 2.1.$$

Using the latter, q and λ can be calculated:

$$q = 2.1; \qquad \lambda = \frac{1.61}{250}.$$

The lower limit, $L \leq 5$, and the probability, $p = 7/8$, can be explicitly found respectively in Szilard and Burch, and in Szilard.

Using the value of p, $E(\beta)$ can be calculated: $E(\beta) = 17.5$.

In order to calculate K we use the expression, $K = 20 \sum_{i=6}^{17} i^{-1}$, which is equivalent [2] to the expression considered in the introduction, but makes it much easier to calculate it: $K = 23.12$.

Then, we obtain the expressions:

$$\mu(\alpha) = 2.1 \exp\left(-\frac{1.61}{250}\alpha\right) \qquad [2]$$

$$E[t(\alpha, \beta)] = \frac{1}{\mu(\alpha)}(5.62 + \beta). \qquad [3]$$

[1] The expression of $\mu(\alpha)$ can be obtained by multiplication, the process being a stochastic one with independent increments.

[2] For the same reason referred to in note 1, the expectation of the time of gene decay can be obtained by summing the expectations of the time of the single errors; and, considering $\beta = E(\beta)$, we obtain:

$$E[t(\alpha, \beta)] = \frac{K}{\mu(\alpha)} \simeq \frac{20}{\mu(\alpha)} \sum_{1-6}^{17} i^{-1}.$$

https://doi.org/10.1017/S1120962300010908 Published online by Cambridge University Press

It is useful to consider something about the probability distributions of α and β (Fig. 1, Table I).

The probability distribution of α is a normal one with expectation $m = 250$ and variance $\sigma^2 = 125$ ($\sigma = 11.2$).

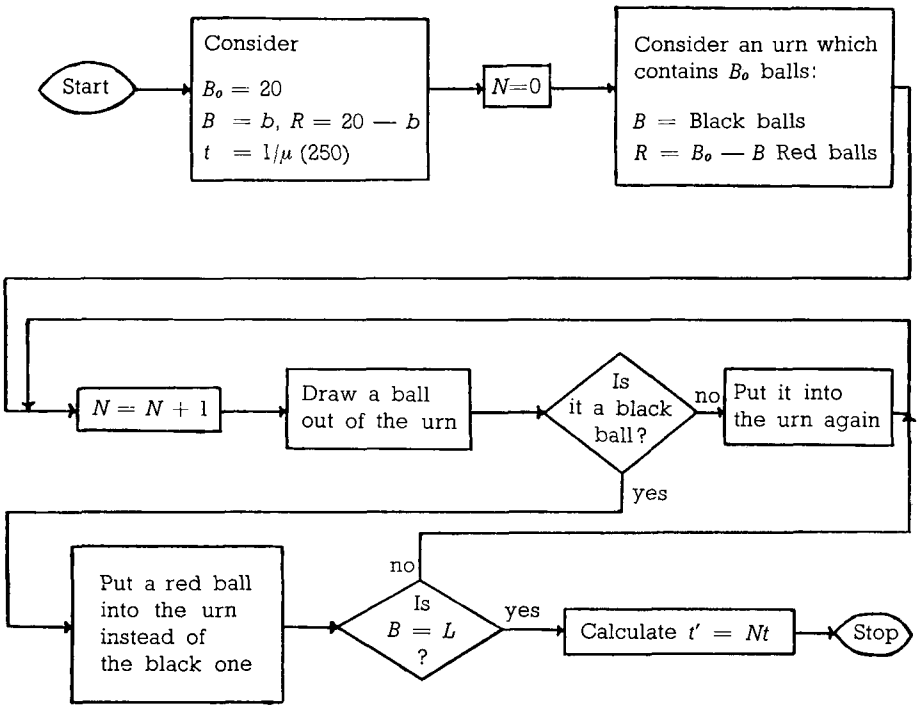

FIG. 1. *Flow chart for the analytic simulation of gene decay*

TABLE I

PROBABILITY DISTRIBUTION OF THE PARAMETER β: $p(\beta = h) = \binom{20}{h} p^L (1-p)^{20-h}$

h	$p(\beta = h)$	h	$p(\beta = h)$
0	0.00000000000000000	11	0.00028806106831780
1	0.00000000000000012	12	0.00151232060670736
2	0.00000000000000807	13	0.00651461184861545
3	0.00000000000042840	14	0.02280114147288259
4	0.00000000001008988	15	0.06384319608332589
5	0.00000000022601341	16	0.13965699146501719
6	0.00000000395523479	17	0.20555722196586 43
7	0.00000005537328705	18	0.26836049358826130
8	0.00000062987114013	19	0.19773931091185659
9	0.00000587879730673	20	0.06920875882497057
10	0.00004522263779449		

https://doi.org/10.1017/S1120962300010908 Published online by Cambridge University Press

Consider the Chebyshev inequality (de Finetti 1970); it is possible to determine an interval, in which α can vary, in such a way that the probability of α assuming a value out of the interval be less than ε (assuming $\varepsilon = 0.01$). The interval I_α that must now be considered is $I_\alpha \equiv [200,300]$: we call such an interval " critical zone " for α.

In the same way one can consider the probability distribution of β, which is a binomial one with expectation $m = 17.5$ and variance $\sigma^2 = 2.1875$ ($\sigma = 1.479$): the " critical zone " for β is then $I_\alpha \equiv [13,20]$.

For our purposes, we can consider every function of the parameters α and β in the critical zones only.

[Automatic procedures have been carried out by IBM 1130 computer (FORTRAN program)].

SIMULATION RESULTS

Eleven simulation tests have been obtained, using a scale 1 : 100[3], with β assuming in the critical zone values $\beta = 13$, $\beta = 15$, $\beta = 17$, $\beta = 20$.

The Monte Carlo method has been used by the flow chart obtained in the stochastic model of gene decay (Fig. 2).

Let $\alpha = E(\alpha) = 250$, since only $\mu(250)$ is a value that may be obtained with good precision.[4]

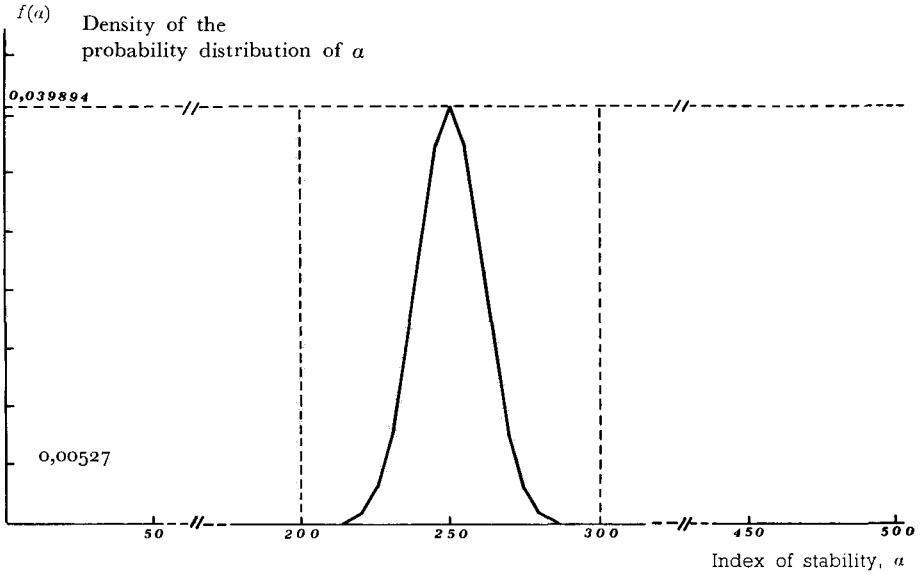

FIG. 2. *Probability distribution of the parameter* α

[3] The error of this method may thus be reduced below a well-defined limit.

[4] The other values of $\mu(\alpha)$ may be obtained by an approximating function that cannot grant absolute precision.

https://doi.org/10.1017/S1120962300010908 Published online by Cambridge University Press

TABLE II

RESULTS OF THE ANALYTIC SIMULATION

$\beta = 13$	$\beta = 15$	$\beta = 17$	$\beta = 20$
$T(1) = 42.1$	$T(1) = 48.1$	$T(1) = 53.4$	$T(1) = 61.0$
$T(2) = 43.3$	$T(2) = 51.0$	$T(2) = 56.4$	$T(2) = 62.5$
$T(3) = 42.6$	$T(3) = 48.8$	$T(3) = 54.1$	$T(3) = 60.7$
$T(4) = 42.0$	$T(4) = 47.7$	$T(4) = 53.4$	$T(4) = 61.5$
$T(5) = 44.1$	$T(5) = 50.3$	$T(5) = 56.2$	$T(5) = 63.5$
$T(6) = 41.7$	$T(6) = 48.2$	$T(6) = 54.0$	$T(6) = 63.9$
$T(7) = 42.8$	$T(7) = 48.7$	$T(7) = 54.3$	$T(7) = 61.2$
$T(8) = 41.3$	$T(8) = 47.4$	$T(8) = 53.5$	$T(8) = 61.6$
$T(9) = 41.5$	$T(9) = 47.9$	$T(9) = 53.4$	$T(9) = 60.3$
$T(10) = 43.3$	$T(10) = 49.2$	$T(10) = 55.2$	$T(10) = 62.8$
$T(11) = 43.2$	$T(11) = 49.2$	$T(11) = 54.0$	$T(11) = 61.3$
$E(T) = 42.5$	$E(T) = 48.8$	$E(T) = 54.4$	$E(T) = 61.8$
$Var(T) = 0.81$	$Var(T) = 1.03$	$Var(T) = 1.19$	$Var(T) = 1.38$

TABLE III

EXPECTATION OF THE TIME OF GENE DECAY AS A FUNCTION OF α AND β

α	$\mu(\alpha)$	$E[t(\alpha, 7.5)]$	$E[t(\alpha, 10.)]$	$E[t(\alpha, 12.5)]$	$E[t(\alpha, 15.)]$	$E[t(\alpha, 17.5)]$	$E[t(\alpha, 20.)]$
0	2.100	6.25	7.44	8.63	9.82	11.01	12.20
50	1.522	8.62	10.26	11.91	13.55	15.19	16.83
100	1.103	11.90	14.16	16.43	18.69	20.96	23.23
150	0.799	15.39	19.54	22.67	27.80	28.93	32.05
200	0.579	22.65	26.97	31.28	35.60	39.92	44.23
210	0.543	24.16	28.76	33.36	37.97	42.57	47.17
220	0.509	25.76	30.67	35.58	40.49	45.40	50.31
230	0.477	27.48	32.71	37.95	43.19	48.42	53.66
240	0.448	29.31	34.89	40.47	46.06	51.64	57.23
250	0.420	31.26	37.21	43.16	49.12	55.08	61.03
260	0.394	33.33	39.69	46.04	52.39	58.74	65.09
270	0.369	35.55	42.33	49.10	55.87	62.65	69.42
280	0.346	37.92	45.14	52.37	59.59	66.82	74.04
290	0.324	40.44	48.14	55.85	63.55	71.26	78.97
300	0.304	43.13	51.35	59.57	67.78	77.00	84.22
350	0.220	59.51	70.85	82.19	93.53	104.87	116.21
400	0.160	82.12	97.77	113.42	121.01	124.71	160.36
450	0.116	113.32	134.91	156.50	166.99	199.69	221.28
500	0.084	156.37	186.16	215.96	245.75	275.55	305.34

https://doi.org/10.1017/S1120962300010908 Published online by Cambridge University Press

Table II shows the results obtained by the simulation tests and the sample expectation and variance that could be estimated through them.

In order to compare the mathematical with the stochastic model, we need some values calculated by the mathematical model; putting $\alpha = 250$, we obtain:

Values of the parameter β:	13	15	17	20
Stochastic model:	44.00	49.12	53.80	61.08
Mathematical model:	43.16	49.12	55.08	61.03

It may be noted that the mathematical model essentially fits the stochastic one. As a result, the values of $E[t(\alpha, \beta)]$ may now be completely calculated through [1]. The results are shown in Table III.

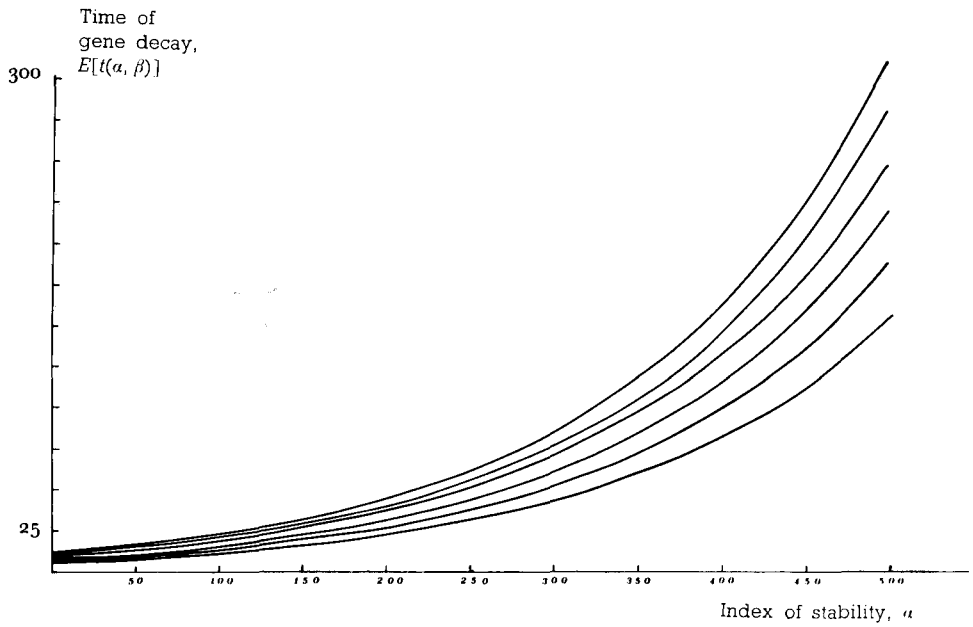

FIG. 3. *Expectation of time of gene decay as a function of α and β*

A nonintegral step for β is chosen: it takes the expectation of β into account, but it is clear that nonintegral values of β have not a biological meaning. In order to have a biological meaning, only the results should be considered, putting $\beta = [\beta]$ (integer part of β).

Figure 3 shows the graphs, obtained by computer, of the expectation of the time of gene decay as a function of α and β, using β as parameter.

https://doi.org/10.1017/S1120962300010908 Published online by Cambridge University Press

REFERENCES

Asimov I. 1962. The Genetic Code. New American Library of World Literature Inc., N.Y.

Burch P.R.S. 1968. Growth, Disease and Aging. Oliver & Boyd, Edinburgh.

Buslenko N.P., Golenko D.I. 1967. The Monte Carlo Method. Pergamon Press, London.

De Finetti B. 1970. Teoria delle Probabilità. Einaudi, Torino.

Feller W. 1957. An Introduction to Probability Theory and Its Application. Wiley, New York.

Gedda L., Brenci G. 1969. Biology of the gene: the Ergon/Chronon System. Acta Genet. Med. Gemellol., 18: 329-379.

Rossi C. 1972. Gene decay. I. Stochastic model of gene decay. Acta Genet. Med. Gemellol., 21: 191-196.

Szilard L. 1959. On the nature of the aging process. Proc. Natl. Acad. Sci. USA, 45: 30-45.

Watson J.D. 1965. Molecular Biology of the Gene. W.A. Benjamin Inc.

RIASSUNTO

I modelli sul decadimento del gene, elaborati in precedenza sulla base del *Sistema Ergon/Chronon* di Gedda-Brenci, vengono ora sottoposti a simulazione analitica al fine di verificare l'aderenza del modello matematico con quello stocastico. Viene quindi calcolato il tempo atteso di decadimento del gene.

RÉSUMÉ

Les modèles de l'épuisement du gène, précédemment élaborés sur la base du *Système Ergon/Chronon* de Gedda-Brenci, sont maintenant vérifiés par une simulation analytique controlant si le modèle mathématique correspond au modèle stochastique. Le temps théorique de l'épuisement du gène est ensuite calculé.

ZUSAMMENFASSUNG

Die zuvor auf Grund des *Ergon/Chronon-System* von Gedda-Brenci ausgearbeiteten Modelle des Genverfalls werden nun einer analytischen Simulation unterzogen, um zu kontrollieren, ob das mathematische mit dem stochastischen Modell übereinstimmt. Sodann wird die theoretische Verfallszeit des Gens errechnet.

Dr.ssa C. Rossi, Istituto di Calcolo delle Probabilità, Facoltà di Scienze Statistiche dell'Università, Rome, Italy.

https://doi.org/10.1017/S1120962300010908 Published online by Cambridge University Press