A Stochastic Modelling of Mutations in Bacteria

CARL A. CRIBBS AND CARL A. BARROWS

Evolutionary theory depends upon certain processes such as an increase in complexity made possible (in part, at least) by changes in DNA, the genetic material. These mutations are considered by many to be mostly harmful, but still the most important source of change in evolution.[1,2] Recently, the importance of mutations in evolutionary processes has been questioned by Gould and others,[3,4] although few evolutionary theorists would claim that mutations don't play some role.

To help define the issue of whether mutations could serve as a source of new genetic traits, we decided to apply a mathematical model. The model assumes certain mutation frequencies in bacteria and follows the effect of cumulative mutations on subsequent generations. Ayala estimated frequency rates in animals to be one in 1,000,000 per gene per generation.[5] Others, such as Schute, have estimated a rate of one in 10^9 cells.[6]

Winchester[7] concluded from various studies including Mueller's experiments with induced mutations in fruit flies[8] that well over 99% of all mutations are harmful to some degree and reduce the fitness of the organism. According to Dobzhansky, most mutations are actually lethal.[9] Mutations that provide antibiotic resistance in bacteria actually are a genetic burden and disadvantage to those bacteria which possess the traits when selection pressures are absent.[10] We based our model upon a bacterial population in which one cell in every billion cells mutated. The probability of a mutation being harmful in some way was 99.99%. The remaining .01% was divided into neutral, reversion, and beneficial mutations. Here we were assuming that beneficial mutations could occur in our bacterial population. As our model demonstrates, within valid parameters, mutations result in an inevitable lethal genetic burden on the progeny of the bacterial cell in which the mutations occur. Offspring always reach the absorbing state of death, and the population as a whole suffers from "dirtying of the gene pool."

THEORY AND STATEMENT OF PROBLEM

Stochastic models are used when there are several events that are possible. An example would be the events in machining a component (using work stations A, B, and C to represent events). If the component is machined at one of the work stations, then machined at another station or reworked, a component might follow a pattern of steps such as: A to B to C; A to C to B to B to B; C to A to B to A to C; A to A to A to B to . . . etc., where A to A is staying for reworking. One may assume that the component does not have to follow a prescribed sequence of work stations, e.g., A to B to C. Based on past history of the machining patterns of the component, probabilities of going from one station to another (or staying) can be determined. These probabilities are represented by

P_{AB} = probability of going from station A to B

P_{AA} = probability of going from station A to A

P_{CB} = probability of going from station C to B

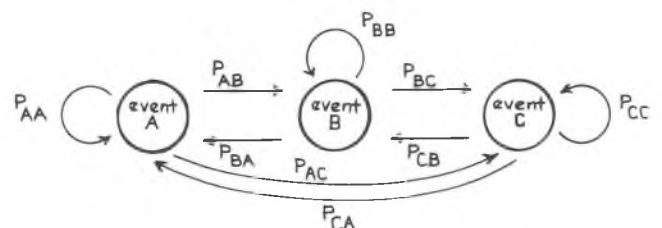The diagram of what can happen (with probabilities indicated) is shown in Figure 1.

Fig.1

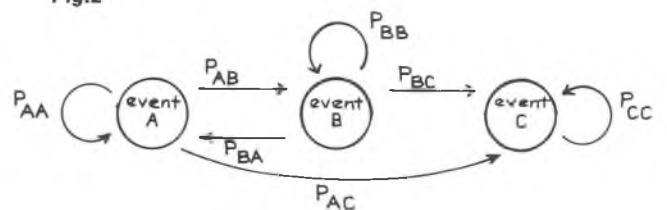Note: P_{AB} does not necessarily equal P_{BA}. Also:

$$P_{AA} + P_{AB} + P_{AC} = 1$$
$$P_{BA} + P_{BB} + P_{BC} = 1$$
$$P_{CA} + P_{CB} + P_{CC} = 1$$

Once the job is done, the component enters a station it does not return from. If station C is changed to the event of being stored in a warehouse, the diagram becomes as shown in Figure 2.

Fig.2

Note:
$$P_{AA} + P_{AB} + P_{AC} = 1$$
$$P_{BA} + P_{BB} + P_{BC} = 1$$
$$P_{CC} = 1$$
$$P_{CA} = P_{CB} = 0 \text{ (cannot happen)}$$

Once event C happens, the job is done. This is called an absorbing state. A matrix is formed from this data called a probability matrix (the rows sum to 1.0). Using the second diagram, the matrix representation is

$$P = \begin{bmatrix} P_{AA} & P_{AB} & P_{AC} \\ P_{BA} & P_{BB} & P_{BC} \\ P_{CA} & P_{CB} & P_{CC} \end{bmatrix} = \begin{bmatrix} P_{AA} & P_{AB} & P_{AC} \\ P_{BA} & P_{BB} & P_{BC} \\ 0 & 0 & 1 \end{bmatrix}$$

With this matrix, calculations can be done to determine the probability of being in one of the stations in k number of steps, given the current station. For example, the probability can be determined for going from station B to station C (finished) in four steps (k = 4) following any legal path. Legal paths include:

B to B to A to C

B to A to B to C

B to C (once entered, finished)

etc.

but exclude:

B to C to A (must stop once C is entered)

or A to B to C (must start with station B)

This probability would be determined through taking the fourth power of P:

$$P \times P \times P \times P = P^4$$

so

$$P^4 = \begin{bmatrix} P_{AA} & P_{AB} & P_{AC} \\ P_{BA} & P_{BB} & P_{BC} \\ 0 & 0 & 1 \end{bmatrix}^4 = \begin{bmatrix} P_{AA}^{(4)} & P_{AB}^{(4)} & P_{AC}^{(4)} \\ P_{BA}^{(4)} & P_{BB}^{(4)} & P_{BC}^{(4)} \\ 0 & 0 & 1 \end{bmatrix}$$

where $P_{BC}^{(4)}$ is not P_{BC} to the fourth power, but $P_{BC}^{(4)}$ is the probability of going from station (or state) B to station C in four steps. A more technical approach is now presented.

A Markov chain is defined formally as "for each k, the conditional probability that the system is in a given state after k steps, knowing the states of the system at all preceding steps, is the same as the conditional probability knowing only the state at the immediately preceding step."[11] That is, the only past information needed for the current step is what happened in the previous step. So, a probability matrix is:

$$P = \begin{bmatrix} P_{00} & P_{01} & \cdots & P_{0n} \\ P_{10} & P_{11} & \cdots & P_{1n} \\ \vdots & \vdots & & \vdots \\ P_{n0} & P_{n1} & \cdots & P_{nn} \end{bmatrix}$$

where the rows sum to one, and

P_{00} is the probability of going from state 0 to state 0 in one step

P_{10} is the probability of going from state 1 to state 0 in one step

$\vdots$ $\qquad$ $\vdots$

P_{ij} is the probability of going from state i (the previous state) to state j, $0 \leqslant P_{ij} \leqslant 1$; $i,j = 0, 1, \ldots, n$, where a state can be an event, condition, change, etc. Therefore, the probability of going from event i to event j in n steps is $P_{ij}^{(n)}$. $P_{ij}^{(n)}$ is a probability as described above, but $P_{ij}^{(n)}$ is the result of multiplying matrix P times itself n times (P^n).

Using the probability matrix P, the following matrix is established for traits (capabilities) for a bacteria:

$$P = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \\ P_{30} & P_{31} & P_{32} & P_{33} \end{bmatrix}$$
$$= \begin{bmatrix} .00009 & .00000999 & .00000001 & .9999 \\ .00008 & .0001 & .00002 & .9998 \\ .0000799 & .00002 & .0000001 & .9999 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where state 0 is a neutral mutation

state 1 is a beneficial mutation

state 2 is a reverting mutation

state 3 is a lethal mutation or harmful mutation

The probability of a neutral mutation must be greater than or equal to the probability of a reversion since (in the long term) the number of reversions cannot exceed the number of neutral mutations.

Assumptions for a beneficial mutation:

1. increase the probability of another beneficial mutation
2. increase the probability of a reversion mutation
3. decrease the probability of a lethal mutation

A lethal or harmful mutation ends all (an absorbing state). Mutation probabilities in the top row are based on mutation data given earlier. The .01% is divided up as follows:

.009 neutral

.00000999 beneficial (never really observed)

.00000001 reverting (conservative, favoring the evolutionist)

The probability matrices for 2, 3, 5, 10, 15, and 20 stages of mutations are shown in Figure 3. The computations were done on a TRS 80 Model I computer with 16 significant digits and accuracy to 10^{-38}.

ANALYSIS

Using probabilities designed to give values optimistic to the evolutionary viewpoint, the stochastic (series of random events) models show that **"with time"** (an important element for evolutionary theory) the tendency is toward a loss of capabilities. The probability of a harmful mutation increases with each mutation. Even though the matrix represents one

```
M ^ 1
   9.000000000000000D-05      9.990000000000000D-06      1.000000000000000D-08      9.999000000000000D-01
   8.000000000000000D-05      1.000000000000000D-04      2.000000000000000D-05      9.998000000000000D-01
   7.930000000000000D-05      2.000000000000000D-05      1.000000000000000D-07      9.999000000000000D-01
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00


m ^ 2
   8.899999000000000D-09      1.898300000000000D-09      2.007010000000000D-10      9.999999890010000D-01
   1.679800000000000D-08      1.119920000000000D-08      2.002800000000000D-09      9.999999700000000D-01
   8.798990000000000D-09      2.800201000000000D-09      4.008090000000000D-10      9.999999880000000D-01
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00


M ^ 3
   9.688999199000000D-13      2.827550100010000D-13      3.807507009000000D-14      9.999999999987100D-01
   2.567779972000000D-12      1.327788020000000D-12      2.243522600000000D-13      9.999999999958800D-01
   1.047943819100000D-12      3.759381901000000D-13      5.613209080000000D-14      9.999999999985200D-01
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00


M ^ 5
   1.370794913732930D-20      5.112510475391250D-21      7.760217476395810D-22      1.000000000000000D+00
   4.713149353923977D-20      2.037281125886430D-20      3.264572209018060D-21      1.000000000000000D+00
   1.613566776496830D-20      6.356711256955570D-21      9.857518309569260D-22      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00


M ^10
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00


M ^15
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00


M ^20
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
   0.000000000000000D+00      0.000000000000000D+00      0.000000000000000D+00      1.000000000000000D+00
```

Fig. 3 Computer print-out of the probability matrices for 1, 2, 3, 5, 10, 15 and 20 stages of mutations.

system of n mutations, the best a descendant of this bacterium can attain is a probability matrix of P. Basically, it would lie between P and P^n, unless a mutation occurs at birth (resulting in a matrix between P and P^{n+1}, which is worse). This tendency to go toward a harmful mutation with each new mutation (step) is due to the fact that a harmful mutation is an absorbing state (a state once entered that cannot be left).

Theorem. In an absorbing Markov chain, the probability that the process will be absorbed is 1.[12]

Therefore, the presence of the ability to have a lethal or harmful mutation (no matter how small the probability), causes the probability of the next lethal or harmful mutation to increase.

CONCLUSIONS

This is only a statistical model, but it puts a ceiling on mutations. What exists now is optimal, the ability to have mutations causes a loss of capabilities (a downward trend toward a higher probability to have a harmful or lethal mutation). The organism would be better off not to have the ability to mutate (and definite beneficial mutations have not been observed) since the probability of a harmful mutation is so high. Even if we assume a lower probability that a certain mutation would be harmful to a bacterial cell, we still can show with this model that a critical genetic burden results. This stochastic model shows a real tendency toward decline rather than new traits that would improve the organism's chances of survival. So we conclude that mutations always result in a net disadvantage for the organism even when a new trait is introduced.

REFERENCES

1. Dobzhansky, T. **Mankind Evolving**. New Haven: Yale University, 1962, p. 140.
2. Mayr, E. **Populations, Species and Evolution**. Massachusetts: Harvard Univ. Press, 1970, p. 102.
3. Gould, J. J. "In Evolutionary Theory Under Fire," **Science**, Vol. 210 (Nov., 1980), pp. 883–887.
4. Salisbury, F. "Doubts About the Modern Synthetic Theory of Evolution." **The American Biology Teacher** (Sept., 1971), pp. 335–338. Vol. 33, No. 6.
5. Ayala, F. "Teleological Explanations in Evolutionary Biology," **Philosophy of Science**, Vol. 37 (March, 1970), p. 3.
6. Schute, E. **Flaws in the Theory of Evolution**. New Jersey: Craig Publishing Co., 1961, p. 31.
7. Winchester, A. M. **Genetics**. Dallas: Houghton Mifflin, 1966, p. 405.
8. Mueller, H. J. "Radiation Damage to the Genetic Material," **American Scientist** (Jan, 1950), pp. 33–59.
9. Dobzansky, T. **Evolution, Genetics, and Man**. New York: John Wiley and Sons, 1955, pp. 98–103.
10. Koch, A. L. "Evolution of Antibiotic Resistance Gene Function," **Bacteriological Review**, Vol. 45 (1981), pp. 355–378.
11. Clarke, A. Bruce, and Disney, Ralph L. **Probability and Random Processes for Engineers and Scientists**. New York: John Wiley and Sons, 1970, p. 218.
12. Kemeny, John; Schleifer, Arthur, Jr.; Snell, J. Laurie; and Thompson, Gerald L. **Finite Mathematics with Business Applications**. Englewood Clifts, N. J.: Prentice-Hall. Inc., 1962, p. 283.