[10] Morton, N. E., *Progress in Medical Genetics*, ed. A. G. Steinberg (New York: Grune and Stratton, 1961), vol. 1, p. 261.
[11] Bemiss, S. M., *Trans. Am. Med. Assoc.*, **11**, 319 (1858).
[12] Arner, G. B. L., *Columbia University Studies in History, Economics, and Public Law*, **31**, 1 (1908).
[13] Böök, J. A., *Ann. Human Genet.*, **21**, 191 (1957).
[14] Sutter, J., *Biol. méd.*, **47**, 563 (1958).
[15] Slatis, H. M., R. H. Reis, and R. E. Hoene, *Am. J. Human Genet.*, **10**, 446 (1958).
[16] Tanaka, K., *Proceedings of Israel Conference on Human Population Genetics* (in press).
[17] Masuda, M., and N. Fujiki, "Studies on consanguinity effects in isolated villages in Honshu, Japan," manuscript in preparation (1962).
[18] Yanase, T., "Studies on consanguinity effects in isolated villages in Kyushu, Japan," manuscript in preparation (1962).
[19] Hiraizumi, Y., and J. F. Crow, *Genetics*, **45**, 1071 (1960).
[20] Sugiyama, S., and W. J. Schull, *Monumenta Nipponica*, **15**, 126 (1960).
[21] Muller, H. J., *Am. J. Human Genet.*, **2**, 111 (1950).
[22] Neel, J. V., *ibid.*, **10**, 398 (1958).

ON THE GENETIC BASIS OF VARIATION AND HETEROGENEITY OF DNA BASE COMPOSITION

BY NOBORU SUEOKA

DEPARTMENT OF MICROBIOLOGY, UNIVERSITY OF ILLINOIS

Communicated by Sewall Wright, February 26, 1962

A large amount of information has been collected on the base composition of deoxyribonucleic acid (DNA) and a compilation of the available data has been recently presented.[1] Several features may be briefly noted. Among the bacteria there exists a wide variation in the mean DNA base composition.[2, 3] However, within each bacterial species the heterogeneity in the base composition is extremely small.[4, 5] In higher plants and animals the variation from species to species is considerably smaller than that found for the bacteria.[1, 6]

In the present paper, a unitary theory based on genetic and evolutionary considerations will be developed which attempts to account for the main characteristics of the distribution of DNA base composition in nature. A brief description of the theory has been presented in our previous paper.[1] The theory seems to be successful in the sense that it provides a consistent model which explains the stability of the mean base composition and the unimodal distribution of base composition among the DNA molecules from an organism.

As a preliminary to the description of the theory, certain pertinent facts relevant to the variation and heterogeneity of DNA base composition will be recapitulated here:

1. Among bacteria (possibly protozoa and algae also) the mean GC (guanine-cytosine) content of DNA varies approximately from 25 to 75 per cent, and this range extends over the range of the mean GC content of DNA of higher organisms.[2, 3]

2. Phylogenetic relations are reflected in the mean GC content. Thus, closely related organisms show similar base compositions.[2, 3, 6, 7]

3. Distribution of the GC content among the DNA molecules of an organism

is unimodal and the range of the heterogeneity is comparatively narrow. Consequently, if the mean GC contents of DNA of two species are different by 10 per cent in bacteria, there are few common DNA molecules of the same GC content between the two species.[4, 5]

4. When closely related organisms show a difference (usually small) in the mean GC content, the difference is rather uniform among the molecules.[7]

The Theory.—The cause for any change of the GC content of DNA must be found in conversions of base-pairs from AT (adenine-thymine) to GC or vice versa. For present purposes it is unnecessary to distinguish AT from TA and GC from CG. Consequently, AT and TA will be denoted as the α pair and, correspondingly, GC and CG as the γ pair. For the ideal case, we assume that each pair undergoes the conversion more or less independently. Further, the rates of the conversion per generation, u for $\gamma \rightarrow \alpha$ and v for $\alpha \rightarrow \gamma$, are assumed to be rather uniform for all base pairs in the DNA molecules of an organism. It is noted that in the present theory, this is the only major assumption. Diagrammatically we have then

$$\gamma \underset{v}{\overset{u}{\rightleftharpoons}} \alpha$$

The conversions are the cause of point mutations and therefore u and v are undoubtedly related to the mutation rates of genes. The actual rates which influence the DNA base composition, however, are the products of mutation and selection. In other words, only conversions which survive and establish themselves in the following generations contribute to the variation of the GC content.

Therefore, the values u and v are taken as the rates of the conversion $\gamma \rightarrow \alpha$ and $\alpha \rightarrow \gamma$ respectively per generation per base pair which survives and is transmitted to the next generation. These are called the *effective base conversion rates.* Theoretical consequences of the above assumption will be similar to the basic principles of population genetics.[8]

If we put the GC (or γ) content (in terms of mole fraction) of DNA molecules at the nth generation as p_n and the AT (or α) content as $(1 - p_n)$, the change in the GC content (p) in one generation, Δp, becomes:

$$\Delta p = v(1 - p_n) - up_n = v - (u + v)p_n$$

or

$$p_{n+1} = p_n + v - (u + v)p_n \tag{1}$$

when we are dealing with a population of DNA molecules of various GC contents the above equation is also true for the mean value of p, or $\bar{p}$.

$$\Delta\bar{p} = v - (u + v)\bar{p}_n$$

or,

$$\bar{p}_{n+1} = \bar{p}_n + v - (u + v)\bar{p}_n \tag{2}$$

For the equilibrated state $\Delta\bar{p}$ is zero, so that the mean GC content at equilibrium ($\hat{p}$) is:

$$\hat{p} = \frac{v}{u + v} \tag{3}$$

The number of generations (n) necessary to change from an initial value ($\bar{p}_0$) to a certain transient value ($\bar{p}_n$) toward the final equilibrium value ($\hat{p}$) can be calculated from:

$$n = \frac{1}{u + v} \ln \left(\frac{\bar{p}_0 - \hat{p}}{\bar{p}_n - \hat{p}} \right) \tag{4}$$

The distribution function of the GC content of DNA molecules at equilibrium, $f(p)$, can also be derived, which gives an estimate of variance of the GC distribution of DNA at equilibrium. The function is a binomial distribution which is:

$$f_b(p) = \frac{b(b-1)\ldots(b-pb+1)}{(pb)!} (\hat{p})^{pb} (1-\hat{p})^{b(1-p)} \quad \text{(see appendix)} \tag{5}$$

and its variance is:

$$\sigma_b^2(p) = \frac{\hat{p}(1-\hat{p})}{b} \quad \text{(see appendix)} \tag{6}$$

where b is the number of base-pairs per DNA molecule which can be obtained by dividing the molecular weight of sodium salt of DNA by 660. Strictly speaking, in the presence of molecular weight heterogeneity, the number average of pair-numbers (b_N) should be used instead of b. However, for minimal estimation of the variance, the value b_W (weight average) can be used for b. It is important to note that because of the usual large values of the base pair-number (b) in DNA molecules (more than 10^3), the distribution function will be practically normal (therefore symmetrical) with the variance given in equation (6). Obviously equation (5) is true when the sequence of α and γ pairs in the DNA molecules is random. However, the equation will hold approximately true for cases where nonrandomness exists in short regions (for example, in triplets, quadruplets, etc.).

The consequences of these calculations may be summarized:

1. *Mean GC content:* The mean GC content of DNA ($\bar{p}$) of an organism will be determined by the relative magnitudes of effective base conversion rates u and v. In this system, the GC-content of DNA molecules can take any value between 0 and 100 per cent. The relation between fractional GC-content at equilibrium ($\hat{p}$) and the ratio u/v is as follows:

$\hat{p}$	0	0.25	0.4	0.5	0.6	0.75	1.0
u/v	∞	3.0	1.5	1.0	0.67	0.33	0

2. *Shifting of mean GC content:* The rate at which the mean GC content ($\bar{p}$) moves toward equilibrium is primarily dependent on the absolute values of u and v from equation (4). Estimation of u and v is not very easy because of the uncertainty in the estimation of the average mutation rate of genes and particularly in the estimation of the average extent of survival of the mutations. Just for an example, however, let us assume that spontaneous *gene* mutation rate is on the average 10^{-5} per generation, and that the two thirds of the point mutations are based on changes of

$\alpha \rightleftharpoons \gamma$ type. If we assume that each gene has 10^3 nucleotide pairs, and all conversions are transmitted to the offspring (although very unlikely), then the order of magnitude of u or v is 10^{-8} per pair per generation. This is probably a large overestimation, because any selection against mutants will lower the estimates of u and v.

Suppose an organism which has a DNA complement of $\bar{p}_0 = 0.30$ is suddenly subjected to a different physiological condition where $u/v = 0.43$ (corresponding to $\hat{p} = 0.70$). Then we can ask how many generations are required for the average GC content ($\bar{p}$) to reach some intermediate point, say 0.40 ($\bar{p}_n$). Using equation (4) we obtain $n = 2 \times 10^7$ generations. For an organism of generation time one hour, this means two thousand years. The value obtained here may be considerably underestimated, because only a fraction of the mutations will be selected for and become established in the population. Consequently, in this model the mean value of the GC-content of DNA is seen to be an extremely stable characteristic, a feature which is consistent with the fact that taxonomically related organisms have similar base compositions.

3. *Nucleotide sequence:* In the idealized case, the sequence of the four kinds of nucleotides in DNA should be random. If the present theory is close to reality, we expect to find dinucleotide frequencies similar to those predicted from random sequence.

Experimental data on dinucleotide frequencies show that it is fairly close to random.[9] However, attention was brought to the author[10] that the frequency of each kind of dinucleotide pair actually has a tendency toward being slightly discrepant from the random frequency, namely more than random or less than random. This oriented discrepancy could be expected if the code is universal. In the universal code, certain dinucleotide sequence may be more than others if that sequence is found in sensible triplets (for nondegenerate case) or found in more frequent amino acids in the proteins of the organisms (for degenerate case).

4. *Variance of GC distribution:* At equilibrium, the distribution of GC content of the molecules will have a variance which depends upon the molecular weight and its heterogeneity. The calculated values from the theory are presented in Table 1. The values in the table are considerably smaller than the observed values listed in Table 2. However, this is not unexpected, since the theory is valid only for the ideal situation in which the following is assumed; (1) the distribution is in equilibrium; (2) the effective conversion rates (u and v) are the same for all DNA base pairs of the organism in question; (3) these rates are constant for a period during which the distribution has attained equilibrium.

TABLE 1

THEORETICALLY EXPECTED HETEROGENEITY OF DNA BASE COMPOSITION (2σ IN GC CONTENT, %) CALCULATED BY EQUATION (6)

Number average molecular weight ($\times 10^{-6}$)	$\hat{p}$ = .3 or 0.7	$\hat{p}$ = 0.4 or 0.6	$\hat{p}$ = 0.5
1	2.4 (%)	2.5 (%)	2.6 (%)
5	1.1	1.1	1.1
10	0.7	0.8	0.8
20	0.5	0.6	0.6

There is no way to be certain that all of these assumptions will rigorously be met, and any divergence from them will bring about the broadening of the distribution.

TABLE 2

HETEROGENEITY OF DNA BASE COMPOSITION ESTIMATED BY DENSITY GRADIENT CENTRIFUGATION*

	2σ (in GC content, %)	Reference
VERTEBRATES		
Human kidney cells	<9.6	
Mouse spleen	<7.6	1
Frog testis	<6.2	
Calf thymus	~9.6	22
BACTERIA		
Diplococcus pneumoniae	~3.9	22
Bacillus megaterium *Bacillus subtilis* *Escherichia coli* *Micrococcus lysodeikticus* *Micrococcus pyogenes aureus* *Serratia marcescens* *Shigella dysenteriae* *Sarcina lutea* *Pseudomonas aeruginosa*	<6.0	4

* The sign "<" indicates that the σ (standard deviation) has been calculated from the total variance (σ_T^2) of the DNA distribution in the density gradient field, which gives the maximum estimate of the heterogeneity. If we take the number-average molecular weight as 5×10^6, we will not be too far from reality. Then, corresponding "2σ expected," is about 1.0 (see Table 1).

In this connection, the "hot spot" and wide range of reversion rates in the case of the bacteriophage T4 rII locus[11, 12] may be cited. It is possible that the nonuniformity of the mutation rates, which may be attributed to an effect of the neighboring base sequence,[11, 12] is associated with various u/v ratios. At the genic level also, a wide variation in mutation rates among different genes is observed in various organisms, which could also be accompanied by variation in the u/v ratio. It has been pointed out that when there are repetitions of the same sequence of the nucleotides, the variance is greater than that expected from random sequence.[13] Another possible factor tending to broaden the distribution is the incorporation of heterologous DNA into the genome of an organism. This includes a wide range of phenomena, from amphidiploidy to bacterial transformation. In this connection, several exceptions for the unimodality of DNA base composition in higher organisms[1,14–16] and episomic transduction[17] among different species of enterobacteriaceae of different DNA GC contents suggest the possibility of incorporation of heterologous DNA.[7, 18] The incorporated DNA will be subjected to the mutation pressure which tends to push the GC content toward the equilibrium, thus decreasing the introduced heterogeneity. The rate of the compositional shift of the heterologous DNA will depend on the degree of resistance of the base pairs by selection, to the mutation pressure.

The factors mentioned above tend to broaden the distribution of base composition whereas others tend to narrow it. Thus, the DNA molecules of a given organism are in a similar metabolic environment and mutagenic effect by chemical mutagens or ionizing radiation will affect the DNA molecules rather uniformly, and this may tend to narrow the distribution by a rather uniform u/v ratio.

The effective conversion rates, u and v, require more explanation. Each mutagenic effect should have its own average effective conversion rates, u_i and v_i, and, consequently, a unique ratio, u_i/v_i. Then, the net conversion rates (u and v) are sums of these component rates.

$$u = \Sigma u_i$$

$$v = \Sigma v_i$$

There is accumulating evidence that mutagenic effects of various agents, so far found, are not equally effective with regards to $\alpha \rightarrow \gamma$ and $\gamma \rightarrow \alpha$.[19-21]

The main assumption, in the present theory, is that each mutagenic factor acts rather uniformly on base pairs of DNA without regard to their location in the chromosome or in the DNA molecules.

The system described by the present theory takes a unique position in interpreting the observed phenomena of variation and heterogeneity of the DNA base composition and leads naturally to several predictions of the coding problem in general. The essence of the theory places special emphasis on mutations caused by base pair conversion as the primary factor in influencing the base composition of DNA. Since selection operates mainly through phenotype or more specifically on the functional efficiency of protein, it has been taken as an indirect and relatively weak effect in the sense that it cannot impose a uniform effect on the whole genotype. With these considerations in mind, it is difficult to conceive how selection pressure could impose a uniform change in the GC content of DNA. Consequently, the over-all effect of selection pressures is to increase heterogeneity rather than decrease it. In this connection, the greater heterogeneity of DNA base composition found in the more highly differentiated organisms may be interpreted as the result of a resistance of local base compositions to the leveling effect of mutation pressure by selection. It is also predicted that the base composition is rather uniform (closer to random) not only among DNA molecules of one organism but also within the molecule. Although experimental data on intramolecular heterogeneity are rather scarce, the available seem in general to agree fairly well with this prediction.[22]

Several implications of the theory on the biological coding problem may be noted. In the first place, if there is information transfer from DNA to RNA through a direct complementary copying process, we expect from the theory a rather perfect correlation between the base compositions of the two nucleic acids, even if the copying does not extend over the whole of the DNA. Evidence is accumulating that such a complementary RNA fraction exists in various organisms.[23-29] Although the compositional heterogeneity of such RNA has not been examined, the over-all picture fits the average composition of DNA. It is noted that the base composition of the bulk RNA (soluble and ribosomal) does not usually mimic the DNA base composition. This may be interpreted either that the ribosomal and/or soluble RNA is copied from a portion of DNA (or a locus) and the portion occupies only a small fraction of the whole DNA complement, or that the two types of RNA are not copied from DNA but replicate themselves. If the former alternative is right, we expect to find one locus or a few loci responsible for the structure of ribosomal and of soluble RNA.

Consequence of the theory for the amino acid composition of protein may be summarized as follows:

1. When the GC content of the DNA of two organisms differs appreciably, it is unlikely that a protein found in one will be similar in primary structure to any found in the other. This also applies to enzymes of identical function with the exception

that the active site may be similar but the dispensable parts of the molecule will be quite different.[5] In this connection, partial dispensability of parts of molecules in papain[30] and ribonuclease[31] for their enzymatic activity should be remembered. It is also interesting to note that a part of B-cistron of phage T4 rII locus can be dispensed with either as a deletion[32] or as variously mutated forms[33] without losing the function of the B-cistron.

2. If the code is universal, there may be some correlation between GC content of DNA and amino acid content of protein. The exact feature of the correlation will depend on the nature of the code. For example, on the assumption of the universal code among RNA viruses, consistent coding models have been constructed.[34, 35] The result on the total protein of various species of bacteria and *Tetrahymena* has been reported elsewhere, which shows positive, negative, and no correlations in different amino acids.[36] The existence of amino acids, e.g., threonine and leucine, which have practically no correlation with DNA base composition between 25 and 72 per cent GC, indicates that the simplest code (triplet, nonoverlapping, nondegenerate code)[37, 38] needs some modification. Furthermore, the shape of the regression curves of amino acid content on the GC-content of DNA was best interpreted as evidence for degeneracy.[39] From the above reasons, the correlation data were taken as supporting evidence for the universality and degeneracy of the code among bacteria and possibly protozoa.[39] Therefore, there is no inconsistency between the wide variation of DNA composition and over-all amino acid contents.

Concerning the phylogenetic aspect of DNA base composition, it is noted that two lines of thought are inseparably mixed in the present consideration. One is that the similarity of DNA base composition among closely related organisms comes from their common origin and from the stable nature of the base composition. The other is that organisms with similar internal environment have a similar u/v ratio, thus making their base compositions alike. Information on similarity in the DNA base sequence between organisms with similar GC content of DNA will clarify the picture. In this connection, DNA hybridization experiments[40] seem promising. The DNA base composition of invertebrates, vertebrates, and higher plants does not seem to supply much instructive information, since there is little variation in the mean GC content among them.[1, 6] Bacteria have the greatest variation, but unfortunately the taxonomic relationship from other criteria is somewhat arbitrary. The algae and possibly fungi may provide more information, because their range of DNA base composition seems to be wide and the natural classification is certainly better established than in bacteria.

Whether the unimodal DNA distribution which we generally observe comes from wider (converging) or from narrower (diverging) distributions, we do not know. The latter possibility, however, fits the idea that the evolution of the gene complement may arise by duplication of the pre-existing genes and mutations of the duplicated genes, thus providing differentiation of genomes.[41] In this connection, the author was reminded[42] that the duplication and polyploidy will relax the extra loci from selection and thus allow them to convert more readily by mutation pressures.

It may be worth while here to discuss the alternative theoretical possibilities to account for the observed facts of DNA base composition.

Nonuniversal code model: If the coding mechanism is not universal among dif-

ferent organisms, each coding system may come to a unique optimal DNA base composition. Such a possibility is not inconsistent with the adaptor hypothesis[43] and the discovery of the amino acid-transfer RNA molecules[44] which serves as a step in the transfer of information from DNA to the protein synthesizing apparatus. The view that the code could be rather flexible may take the shift of DNA GC content as the transition from one coding system to another.[45] At present, evidence for universality[36, 46-48] outweighs that for possible nonuniversality.[49, 50] No general conclusion, however, can yet be drawn especially for cases like ambivalent mutation of *E. coli*-phage T4 system[49] and intergenic suppressor mutations which affect the primary structure of a polypeptide chain of tryptophan synthetase of *E. coli*.[50]

It is noted, however, that universality or nonuniversality does not basically change the theory described in this paper. For the nonuniversal case, the main factor which determines u and v is now selection pressure rather than the mutation pressure.

Two-symbol code model: The two-symbol code of the 6 keto-6 amino type[51] needs special attention, because it provides the possibility of coding the same information with different GC contents of DNA. In this model, cytosine and adenine, and thymine and guanine act as the same unit letter, and are thus interchangeable without changing the information. This model does not contradict the present theory, but so far there is no supporting evidence for this possibility. Moreover, if the recent findings of Nirenberg and his colleagues[52, 53] and Ochoa and his group[54, 55] about the stimulation of amino acid polymerization with artificial ribonucleotide polymers are the true picture *in vivo*, the two-symbol code loses its validity.

Nonprotein-coding-DNA model: This model consists of a set of rather unlikely possibilities.

In the first place, it assumes there is a region in each DNA molecule where the GC content is close to 50 per cent. In order to have the over-all GC-contents different from 50 per cent, there must be accumulations of α or γ pairs in different parts of the molecules. Secondly, it is assumed that the region with approximately 50 per cent GC is genetically significant and the other part of the molecule is nonsense. In this model, the genetically important part of the DNA molecules are similar in different organisms.

This model does contradict our theory. Although this model does not seem very likely,[5, 56] more data on intramolecular heterogeneity of GC content should be accumulated before final conclusions can be drawn.

Summary.—An idealized theory is presented to account for the main features of base composition distribution of DNA, wide variation and small heterogeneity. The theory seems to account for the main features of DNA base composition distribution in nature. It is based on rather uniform mutation and selection pressures affecting the base pair conversion.

Appendix.—There are several ways to derive equations (5) and (6). As an example a derivation based on a discussion on the birth and death process[57] will be presented.

The number of GC pairs in a molecule is pb where the total number of pairs is b. The probability of a change GC $\rightarrow$ AT in time dt is $\mu_{pb}\,dt$ where $\mu_{pb} = up$, and for AT $\rightarrow$ GC in time dt is $\lambda_{pb}dt$ where $\lambda_{pb} = v(1 - p)$. Since dt is very small, only single changes occur in this time. Then

as Feller shows, equation (5.2) in his book,[57] for $pb = 1$,

$$\frac{df_{pb}(t)}{dt} = (\lambda_{pb} + \mu_{pb})f_{pb}(t) + \lambda_{pb-1}f_{pb-1}(t) + \mu_{pb+1}f_{pb+1}(t) \qquad (1')$$

and for $pb = 0$

$$\frac{df_0(t)}{dt} = -\lambda_0 f_0(t) + \mu_1 f_1(t) \qquad (2')$$

where $f_{pb}(t)$, $f_1(t)$, and $f_0(t)$ are the probabilities that there are pb, 1, and 0 GC pairs respectively at time t. To obtain the equilibrium state, we put the derivatives equal to zero. From (2') this gives:

$$f_1 = \frac{\lambda_0}{\mu_1} f_0 = \frac{bv}{u} f_0$$

Putting $pb = 1$ in (1') we get:

$$f_2 = \frac{b(b-1)}{2}\left(\frac{v}{u}\right)^2 f_0$$

Knowing that

$$f_0 = \left(\frac{u}{u+v}\right)^b = (1-\hat{p})^b,$$

$$f_2 = \frac{b(b-1)}{2}\left(\frac{v}{u+v}\right)^2\left(\frac{u}{u+v}\right)^{b-2} = \frac{b(b-1)}{2}\hat{p}^2(1-\hat{p})^{b-2}$$

Obtaining $f_3, f_4, \ldots$, we get the equilibrium distribution as a binomial distribution,

$$f_{pb} = \frac{b(b-1)\ldots(b-pb+1)}{(pb)!}\hat{p}^{pb}(1-\hat{p})^{b(1-p)}$$

Therefore the variance for this size of molecule is

$$\sigma_b^2(p) = \frac{\hat{p}(1-\hat{p})}{b}$$

When there is the molecular weight heterogeneity, we can get an average value of the variance as follows. Let the distribution function for base number per molecule be $B(b)$, where $B(b)$ if expressed in the weight concentration. Then,

$$\sigma^2(p) = \frac{\int_R \sigma_b^2(p)B(b)db}{\int_R B(p)db} = \hat{p}(1-\hat{p})\frac{\int_R \frac{B(b)}{b}db}{\int_R B(b)db}$$

Here $\int_R B(b)db / \int_R \frac{B(b)}{b}db$ is the number average of b, which is written as b_N in the text. R indicates that the integrations cover the whole range of the distribution.

An essentially same treatment as presented here had been formulated two years ago by the author. Since then, he has received criticisms by a number of people. The author is especially grateful to Drs. K. C. Atwood, S. Benzer, W. G. Cochran, M. Delbrück, P. Doty, M. Kimura, R. P. Levine, C. Levinthal, S. E. Luria, S. Spiegelmen, and S. Wright. For correcting the manuscript, the author owes much to Dr. S. Spiegelman and also to Mrs. M. E. McCarty. This work was supported by a grant from the National Science Foundation (G-15080).

[1] Sueoka, N., *J. Mol. Biol.*, **3**, 31 (1961).

[2] Lee, K. Y., R. Wahl, and E. Barbu, *Ann. Inst. Pasteur*, **91**, 212 (1956).

[3] Belozersky, A. N., and A. S. Spirin, *Nature*, **182**, 111 (1958).

[4] Rolfe, R., and M. Meselson, these PROCEEDINGS, **45**, 1039 (1959).

[5] Sueoka, N., J. Marmur, and P. Doty, *Nature*, **183**, 1429 (1959).

[6] Chargaff, E., *The Nucleic Acids*, ed. E. Chargaff and J. N. Davidson (New York: Academic Press, 1955), vol. 1, p. 307.

[7] Sueoka, N., and D. Nanney (in preparation).

[8] Wright, S., *Ann. Eugenics*, **15**, 323 (1951). Also see C. C. Li, *Population Genetics* (The University of Chicago Press, 1955).

[9] Josse, J., A. D. Kaiser, and A. Kornberg, *J. Biol. Chem.*, **236**, 864 (1961).

[10] Kaiser, A. D., personal communication (1961).

[11] Benzer, S., *The Chemical Basis of Heredity*, ed. W. D. McElroy and B. Glass (The Johns Hopkins Press, 1957), pp. 70–93.

[12] Benzer, S., these PROCEEDINGS, **47**, 403 (1961).

[13] Kimura, M., *Genet. Res.*, **2**, 127 (1961).

[14] Szybalski, W., personal communication (1961).

[15] Kit, S., *The Molecular Basis of Neoplasia*, The Fifth Annual Symposium on Fundamental Cancer Research, ed. S. Kit (Houston: Texas University Press, 1961), in press.

[16] Jacob, F., P. Schaeffer, and E. L. Wollman, in *Microbial Genetics*, Tenth Symposium of the Society for General Microbiology ed., W. Hayes and R. C. Clowes (Cambridge University Press, 1960), p. 67.

[17] Carey, W. F., W. M. Spilman, and L. S. Baron, *Bacteriol. Proc.*, **60**, 73 (1960).

[18] Marmur, J., R. Rownd, S. Falkow, L. S. Baron, C. Schildkraut, and P. Doty, these PROCEEDINGS, **47**, 972 (1961).

[19] Nakaya, R., A. Nakamura, and Y. Murata, *Biochem. Biophys. Res. Comm.*, **3**, 654 (1960).

[20] Freese, E., E. Bautz, and E. Bautz Freese, these PROCEEDINGS, **47**, 845 (1961).

[21] Bautz, E., and E. Freese, these PROCEEDINGS, **46**, 1585 (1960).

[22] Sueoka, N., these PROCEEDINGS, **45**, 1480 (1959).

[23] Volkin, E., and L. Astrachan, *Virology*, **2**, 149 (1956)

[24] Nomura, M., B. Hall, and S. Spiegelman, *J. Mol. Biol.*, **2**, 306 (1960).

[25] Yčas, M., and W. S. Vincent, these PROCEEDINGS, **46**, 804 (1960).

[26] Astrachan, L., and T. M. Fisher, *Fed. Proc.*, **20**, 359 (1961).

[27] Hall, B. D., and S. Spiegelman, these PROCEEDINGS, **47**, 137 (1961).

[28] Brenner, S., F. Jacob, and M. Meselson, *Nature*, **183**, 1608 (1961).

[29] Gros, F., H. Hiatt, W. Gilbert, C. G. Kurland, R. W. Risebough, and J. D. Watson, *Nature* **190**, 581 (1961).

[30] Hill, R. L., and E. L. Smith, *Biochim. Biophys. Acta*, **19**, 376 (1956).

[31] Kalmitsky, G., and W. I. Rogers, *Biochim. Biophys. Acta*, **20**, 378 (1956).

[32] Champe, S. P., and S. Benzer, *J. Mol. Biol.* (1962), in press.

[33] Crick, F. H. C., L. Barnett, S. Brenner, and R. L. Watts-Tobin, *Nature*, **192**, 1227 (1961).

[34] Yčas, M., *Nature*, **188**, 209 (1960).

[35] Woese, C. R., *Biochem. Biophys. Res. Comm.*, **5**, 88 (1961).

[36] Sueoka, N., these PROCEEDINGS, **47**, 1141 (1961).

[37] Crick, F. H. C., J. S. Griffith, and L. E. Orgel, these PROCEEDINGS, **43**, 416 (1957).

[38] Golomb, S. W., L. R. Welch, and M. Delbrück, *Biol. Medd. Dan. Vid. Selsk.*, **23**, 9 (1958).

[39] Sueoka, N., in *Cellular Regulatory Mechanisms*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 26 (1961), in press, p. 35.

[40] Schildkraut, C., J. Marmur, and P. Doty, *J. Mol. Biol.*, **3**, 595 (1961).

[41] Lewis, E. B., in *Genes and Mutations*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 16, 159 (1951).

[42] Atwood, K. C., personal communication (1962).

[43] Crick, F. H. C., in *The Biological Replication of Macromolecules*, Symposia for the Society for Experimental Biology, No. 12, ed. F. K. Sanders (London: Cambridge University Press, 1959), p. 138.

[44] Hoagland, M. B., in *Structure and Function of Genetic Elements*, Brookhaven Symposia in Biology, No. 12, 40 (1959).

[45] Levinthal, C., personal communication (1959).

[46] Jacob, F., and J. Monod, in *Cellular Regulatory Mechanisms*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 26 (1961), in press.

[47] Signer, E. R., A. Torriani, and C. Levinthal in *Cellular Regulatory Mechanisms*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 26 (1961), in press.

[48] Ehrenstein, G. v., and F. Lipman, these PROCEEDINGS, **47**, 941 (1961).

[49] Benzer, S., and S. P. Champ, these PROCEEDINGS, **47**, 1025 (1961).

[50] Yanofsky, C., D. R. Helinski, and B. D. Maling, in *Cellular Regulatory Mechanisms*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 26 (1961), in press.

[51] Sinsheimer, R. L., *J. Mol. Biol.*, **1**, 218 (1959).

[52] Nirenberg, M. W., and J. H. Matthaei, these PROCEEDINGS, **47**, 1588 (1961).

[53] Matthaei, J. H., and M. W. Nirenberg, *Biochem. Biophys. Res. Comm.*, **4**, 404 (1961).

[54] Lengyel, P., J. F. Speyer, and S. Ochoa, these PROCEEDINGS, **47**, 1936 (1961).

[55] Speyer, J. F., P. Lengyel, C. Basilio, and S. Ochoa, these PROCEEDINGS, **48**, 63 (1962).

[56] Crick, F. H. C., in *Structure and Function of Genetic Elements*, Brookhaven Symposia in Biology, No. 12, 35 (1959).

[57] Feller, W., *An Introduction to Probability Theory and its Applications* (New York: John Wiley and Sons, 1957), vol. 1, p. 407.

---

*EFFECT OF THE RIGIDITY OF THE INNER CORE ON THE FUNDAMENTAL OSCILLATION OF THE EARTH**

BY C. L. PEKERIS, Z. ALTERMAN, AND H. JAROSCH

DEPARTMENT OF APPLIED MATHEMATICS, THE WEIZMANN INSTITUTE, REHOVOT, ISRAEL

Communicated January 23, 1962

The Bullen B and Gutenberg models of the earth which we studied previously gave periods for the spheroidal oscillation $n = 2$ of 53.70 and 53.52 min respectively, as against the average observed seismic and gravimetric values of 53.9 min. In order to explain this discrepancy, we have studied the effect on the period of an assumed rigidity in the inner core ($r < 1{,}250$ km). It is found that the period of the "core" oscillation of about 101 min diminishes rapidly with increasing rigidity μ of the inner core, reaching an asymptotic value of about 53.8 min at large μ, while simultaneously the amplitude spreads into the mantle, eventually assuming the pattern of a normal oscillation at the asymptotic period. The observed period of 53.9 min is reached at a value of μ of about $^1/_2 \times 10^{12}$ dyne/cm^2, and could be fitted within the observational error into the range of 1.5×10^{12} to 4×10^{12} for μ inferred by Bullen on the basis of seismic data.

1. *Introduction.*—In the interpretation of the spectrum of the earth which was observed gravimetrically[1] and seismically[2] on the records of the great Chilean earthquake of May 22, 1960, we compared the observed periods with theoretical ones evaluated for Bullen's model B and Gutenberg's model of the earth.[3] For the fundamental spheroidal oscillation $n = 2$, these periods came out 53.70 and 53.52 respectively, compared with the average of 53.89 min for the observed gravimetric[1] doublet of 54.98 and 52.80, and 53.9 min for the center of the observed seismic[2] doublet of 54.7 and 53.1 min. The discrepancy of 0.2 min between theory and observation for Bullen's model and of about 0.4 min for Gutenberg's model is noteworthy in view of the better agreement between theoretical and observed periods that was found for the higher modes, especially for the Gutenberg model, and the