DENSITY DIFFERENCES BETWEEN GENETIC MARKERS IN PNEUMOCOCCAL TRANSFORMING PRINCIPLE

By Ronald Rolfe* and Harriett Ephrussi-Taylor

Gates and Crellin Laboratories of Chemistry,† Norman W. Church Laboratory of Chemical Biology, California Institute of Technology, Pasadena, and Laboratoire de Génétique Physiologique du C.N.R.S., Gif-sur-Yvette, France

It is known that DNA carries genetic information, and it is commonly assumed that the identity and linear order of the genes are determined by the linear sequence of nucleotides in DNA. However, little is known about the detailed relationship (or "code") between the DNA and proteins of a given species. A linkage group of genes may correspond to a unique nucleotide sequence, or it is possible that a large number of different nucleotide sequences could in principle carry the same genetic information.

The DNA of a given bacterial species is restricted by some mechanism within narrow limits of nucleotide composition, relative to the observed range of compositions for the various species.[1, 2] To explain the wide differences in nucleotide composition between DNA's from species that presumably have considerable genetic information in common, codes have been proposed in which A is formally equivalent to C, and G is equivalent to T, so that the AT and GC pairs are interchangeable.[3, 4] Given such a code, mistakes in replication could lead to changes in nucleotide composition with no change in genetic information, and DNA molecules containing a given genetic marker would have a range of possible nucleotide compositions equal to that for the total DNA. Differences in nucleotide composition between molecules containing different genetic markers, and the nucleotide compositional heterogeneity of the DNA, should be consistent with the predictions of a random model for the occurrence of fluctuations in nucleotide composition around some equilibrium value.[5]

In the experiments to be described, the nucleotide compositions of DNA molecules containing specific genetic markers have been investigated, taking advantage of the fact that the buoyant density of a DNA molecule depends upon its nucleotide composition,[1, 2] and that the biological activity of transforming principle may be determined after isolation from a CsCl density gradient.

Materials and Methods.—Pneumococcal DNA containing the genetic markers streptomycin resistance (S^r), optochin resistance (Q^r), bryamycin resistance (B^r), and aminopterin resistance [A^r(136)] in various combinations was studied. A bacterial lysate containing about 20 μg of DNA was centrifuged in CsCl solution of density 1.70 g cm^{-3}, buffered at pH 8.5 with 0.01 M *tris*-(hydroxymethyl)-aminoethanol, and containing 0.001 M sodium versenate. After 48 hr at 35,000 rpm in the Sw39 head of the Spinco preparative ultracentrifuge, the CsCl solution was fractionated by making a small pinhole in the bottom of the lusteroid centrifuge tube and collect-

ing the effluent drops individually. The relationship between CsCl solution density and drop number was established by refractive index measurements[6] of individual drops (Fig. 1). The drops were assayed for biological activity as follows: One milliliter of a culture of receptor strain R36A, grown to maximal competence and chilled, was added to a cold test tube containing 0.01 ml of the CsCl drop to be assayed. The mixture was incubated 20 minutes at 37°C, chilled, and
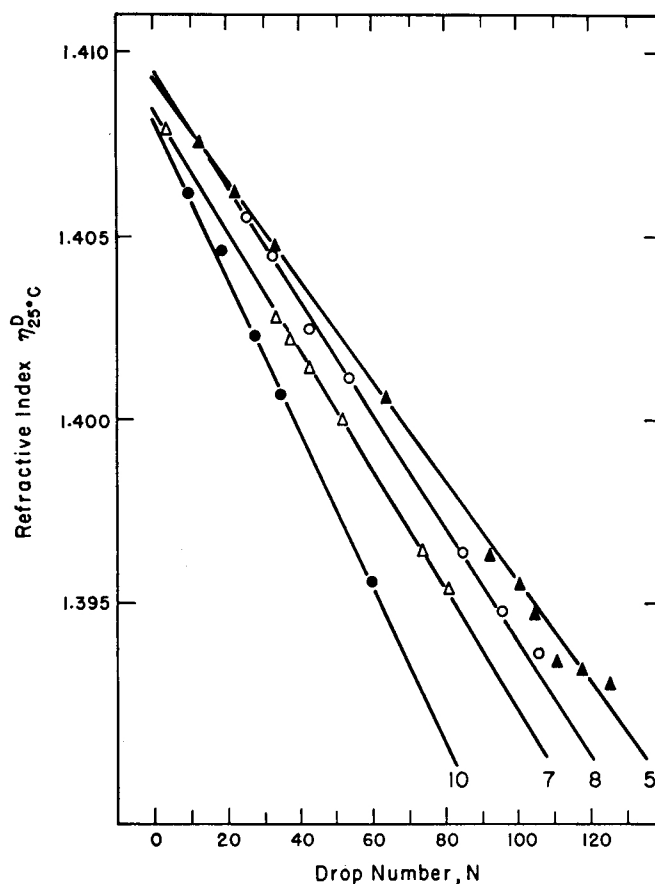

Fig. 1.—The relationship between CsCl drop number and refractive index, for the various experiments. Refractive indices for sodium D light at 25°C. $\eta_D^{25°C}$ for effluent CsCl drops are plotted against drop number. The values are correct to within an additive amount which was constant for a given experiment. Data from experiments 5, 7, 8, and 10 are presented.

plated on blood agar after appropriate dilution. The plates were incubated 2 hr at 37°C, to allow expression of phenotype, overlayered with an equal volume of antibiotic-containing agar, to select the transformants, and colonies were counted 24 to 36 hr later. Antibiotics used were dihydro-streptomycin (200 μg/ml), bryamycin[7] (1.5 μg/ml), optochin (5 μg/ml), or aminopterin (0.75 × 10^{-5} M). Reaction of the cells with the DNA is restricted to 10 min, which is the duration of competence of the receptor strain under the culture conditions used.

The number of transformants for a given marker in the assay system was proportional to the DNA concentration, in the range employed in these experiments. Ratios of numbers of transformants for the different markers found in replicate cultures were constant for a given medium and DNA preparation, and were independent of DNA concentration. However, these ratios were not completely independent of CsCl and other ion concentrations in the medium and it was

found necessary to employ a standard amount of the CsCl solution in the assay culture. Differences in CsCl concentration of the drops from opposite ends of the density gradient were insufficient to modify the transformation ratios.

Density-gradient distributions of transforming activity for the different genetic markers have been plotted according to the following expression:

$$F(N) = T(N)/\sum_N T(N) \tag{1}$$

where $F(N)$ is the frequency of transformation for the Nth drop, and $T(N)$ is the absolute number of transformants for the given marker obtainable from the Nth drop. Mean drop number $\bar{N}$ and variance σ^2 were determined for each density-gradient distribution using the formulae

$$\bar{N} = \sum_N NF(N) \tag{2}$$

$$\sigma^2 = \sum_N (N - \bar{N})^2 F(N) \tag{3}$$

Variances expressed in drops² were converted to variances in gm² cm⁻⁶ using the relationship

$$\bar{\sigma}^2 = \left(\frac{d\rho}{dN}\right)^2 \sigma^2 \, \text{gm}^2 \, \text{cm}^{-6} \tag{4}$$

where $d\rho/dN$ is the density gradient. In units of gm² cm⁻⁶, variances are characteristic for a given DNA sample, independent of the density gradient.

Results.—DNA molecules containing different genetic markers occupy different portions of the density-gradient distribution of the total DNA of pneumococcus. This can be seen in Figure 2, which shows the density-gradient analysis of a sample of genetically marked DNA, simultaneously labeled with P³². Comparing the distributions of P³² and of biological activity for streptomycin resistance, we observe that the marker S^r is associated with DNA molecules which have an average buoyant density greater than that of the bulk of the DNA. In addition, the density-gradient distributions of the S^r and B^r marked molecules differ from each other, and are considerably narrower than the distribution for the total DNA. Density-gradient distributions for pairs of unlinked genetic markers were compared in several experiments (Figs. 2–5), and significant differences in mean drop number $\bar{N}$ were detected for all pairs except (S^r, A^r). Density differences (Table 1) between pairs

TABLE 1

TABULATED PARAMETERS* FOR DENSITY-GRADIENT DISTRIBUTIONS
OF GENETICALLY MARKED PNEUMOCOCCAL DNA

Expt.	$d\rho/dN$ gm cm⁻³ drop⁻¹ × 10²	Marker	N_m drop number ±0.01	$\bar{\sigma}_m^2$ gm² cm⁻⁶ × 10⁶	$\dfrac{d \log T_{sr}/T_m}{d\rho}$	$(N_m - N_{sr}) \times d\rho/dN$	$-\dfrac{d \ln T_{sr}/T_m}{d\rho}\bar{\sigma}_m^2$
						$-\Delta\rho(S^r, m)$ gm cm⁻³ × 10⁴	
5	1.37	S^r	76.706	3.1	...	...	...
		Q^r	77.086	2.9	80	5.2	5.5
10	2.10	S^r	42.818	3.1	...	...	...
		Q^r	42.994	3.3	57	3.8	4.2
		A^r	42.809	3.1	57	4.0	4.2
8	1.54	S^r	70.866	3.7	...	...	...
		B^r	71.512	4.7	123	10	10.6
		P³²	71.906	5.5	...	...	...
7	1.63	S^r	58.153	7.2	...	...	...
		B^r	58.905	6.3	141	12.2	22

* The following are tabulated: $d\rho/dN$ gm cm⁻³ drop⁻¹: the gradient of CsCl-solution density ρ with drop number N; N_m: mean drop number for the distribution of biological activity for the genetic marker m; $\bar{\sigma}_m^2$ gm² cm⁻⁶: variance of the distribution of biological activity for the marker m; T_{sr}/T_m (N): ratio of the number of transformants for the marker S^r to the number of transformants for the marker m obtainable from the Nth drop; $\Delta\rho(S^r, m)$ gm cm⁻³: the difference in mean buoyant density between DNA molecules containing the marker m and those containing the marker S^r.
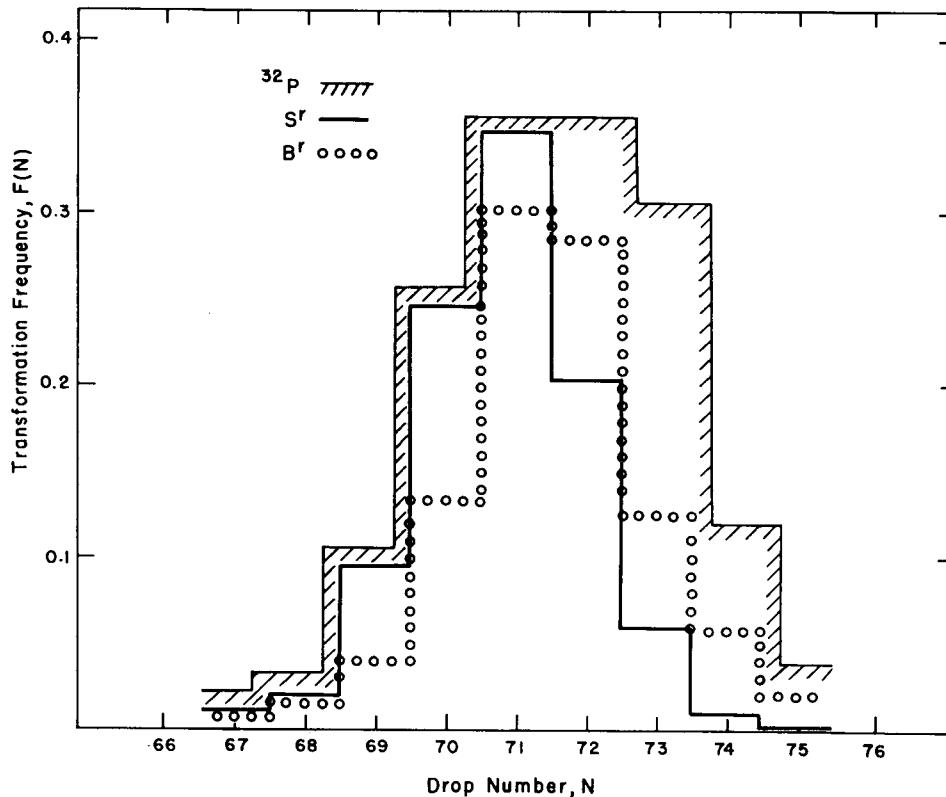
FIG. 2.—Equilibrium density gradient distributions for genetically marked, ^{32}P containing pneumococcal DNA. *D. pneumoniae* carrying the genetic markers for streptomycin resistance S^r, and bryamycin resistant B^r, was grown in a medium containing ^{32}P. Genetically marked DNA labeled to the extent of one ^{32}P atom per 10 molecules was isolated from these organisms and studied in the CsCl density gradient. Distributions of radioactivity and of biological activity for each marker were determined. For the markers S^r and B^r transformation frequency $F(N)$ is plotted against effluent drop number N. For ^{32}P activity, the fraction of total counts multiplied by the factor 1.55 is plotted against N.

of markers $\Delta\rho(A,B)$ were determined from differences in mean drop number $\Delta\bar{N}(A,B)$ and the density gradient $d\rho/dN$ using the relationship

$$\Delta\rho(A,B) = \Delta N(A,B) \left(\frac{d\rho}{dN}\right) \tag{5}$$

For the markers S^r, Q^r and A^r, the density gradient distributions of biological activity are unimodal and not grossly asymmetrical.[8] Due to differences in drop size, the values of $d\rho/dN$ and the width of the marker distributions measured in drops differed in different experiments. However, as can be seen in Table 1, the variances of the distributions for the different markers in a single experiment, and for the same marker in different experiments, were closely similar when expressed in gm² cm⁻⁶.[9]

If the DNA molecules containing a given genetic marker are homogeneous in density and molecular weight, their density-gradient distribution will be gaussian. Gaussian character cannot be judged directly from the few observed points for
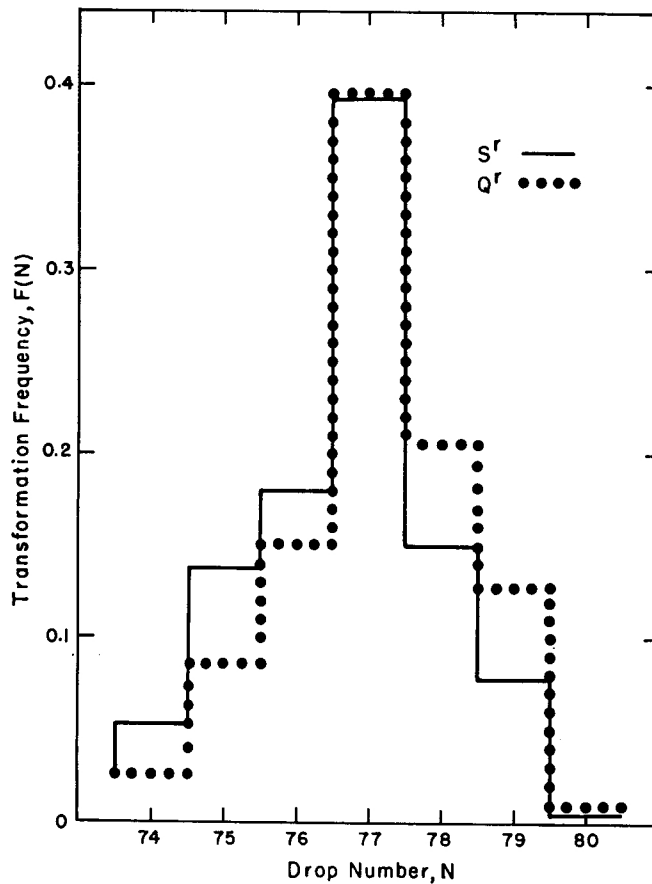
Fig. 3.

each marker distribution.   However, a study of ratios of numbers of transformants for pairs of genetic markers (T_A/T_B), as a function of drop number N, suggests that the density-gradient distributions may be gaussian.

Assume that the density-gradient distributions for the genetic markers A,B are gaussian with equal variances.   If ρ is solution density, $C_i(\rho)$ is the concentration at ρ of molecules carrying the marker i, and ρ_i is the buoyant density of these molecules, we can write

$$C_i(\rho) = C_i(\rho_i) \exp\left(- \frac{(\rho - \rho_i)^2}{2\bar{\sigma}_i^2} \right)$$ (6)

$$\bar{\sigma}_i^2 = \bar{\sigma}^2; \quad i = A,B$$

Then the logarithm of the ratio C_A/C_B should be a linear function of solution density ρ

$$\ln \frac{C_A}{C_B} (\rho) = \alpha\rho + \beta;$$
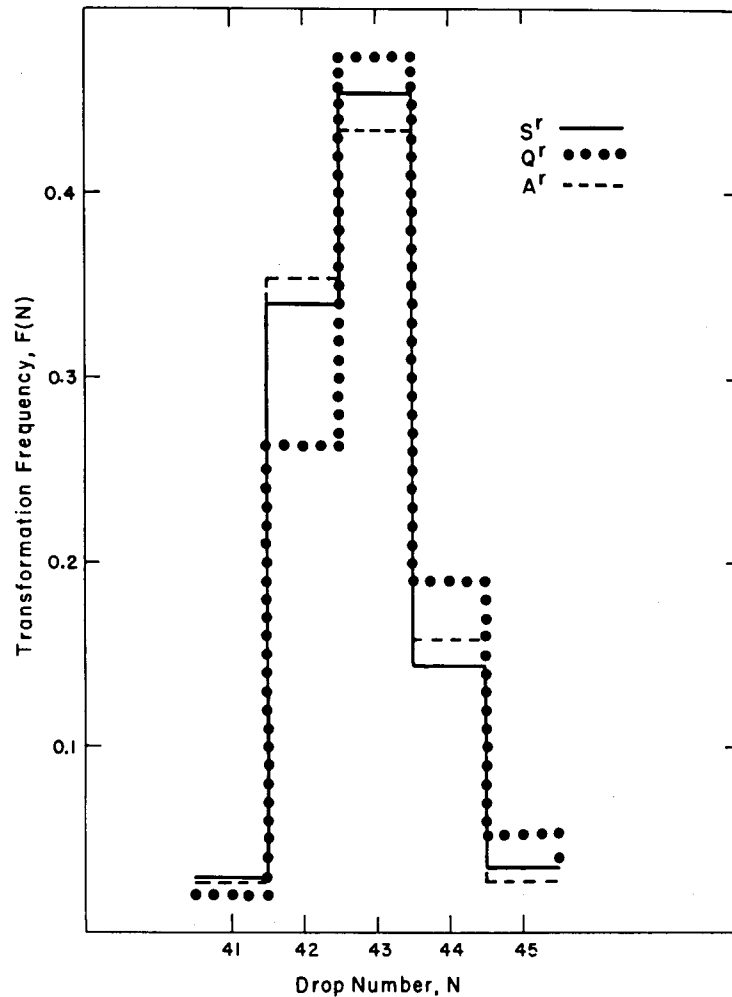
$$\alpha = \frac{\Delta\rho(A,B)}{\bar{\sigma}^2},$$ (7)

FIG. 4.

$$\beta = \frac{\rho_B{}^2 - \rho_A{}^2}{\bar{\sigma}^2} + \ln \frac{C_A(\rho_A)}{C_B(\rho_B)}$$

Plots of log (T_A/T_B) versus drop number N for the various pairs of genetic markers are fairly linear (Figs. 6–8). Since we may write

$$\ln (C_A/C_B) = 2.3 \log (T_A/T_B) + \text{constant} \tag{8}$$

the density difference $\Delta\rho(A,B)$ may be calculated from the slope $[d \log(T_A/T_B)]/dN$, the density gradient $d\rho/dN$, and the variance $\bar{\sigma}^2$ according to the expression

$$\Delta\rho(A,B) = \frac{2.3 \dfrac{d \log (T_A/T_B)}{dN}}{d\rho/dN} \bar{\sigma}^2 \tag{9}$$

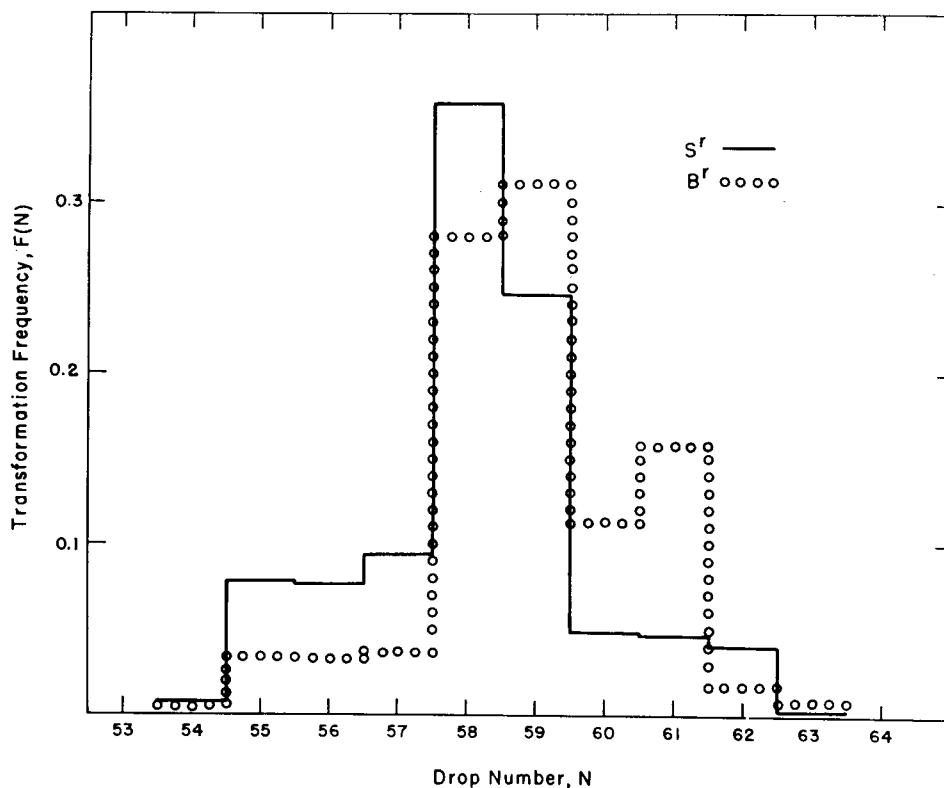Density differences between markers calculated on the assumption that the distri-

FIG. 5.

FIGS. 3–5.—Equilibrium density gradient distributions of transforming activity for the genetic markers S^r, Q^r, B^r. For the different genetic markers in experiments 5, 10, and 7, transformation frequency $F(N)$ is plotted against drop number N. In experiment 10, a mixture of two kinds of DNA was studied: (1) A purified DNA marked with A^r and (2) a DNA marked with S^r and Q^r isolated from a bacterial lysate.

butions are gaussian (Table 1) agree well with differences calculated according to equation (5).

It has been shown that the DNA of pneumococcus contains classes of DNA molecules associated with different genetic markers. These classes differ in mean buoyant density, but are nearly identical with respect to the variance of their density gradient distributions $\bar{\sigma}^2$ (marker). If we assume that the DNA of pneumococcus consists *entirely* of such classes of molecules, we may use the variance $\bar{\sigma}^2$ (marker) to calculate lower bounds for the number average molecular weight[10] M_n and for the density heterogeneity[11] σ_D of the DNA. The true values of these parameters will exceed the calculated lower bounds only to the extent that molecules containing a specific genetic marker are themselves heterogeneous in density.

Thus we may write

$$M_n \text{ (total DNA)} = M_n \text{ (marker molecules)} \geqslant \frac{35}{\bar{\sigma}^2 (\text{marker})} \qquad (10)$$

$$\bar{\sigma}^2_T > \bar{\sigma}^2_D \geqslant \bar{\sigma}^2_T - \bar{\sigma}^2 \text{ (marker) gm}^2 \text{ cm}^{-6} \qquad (11)$$
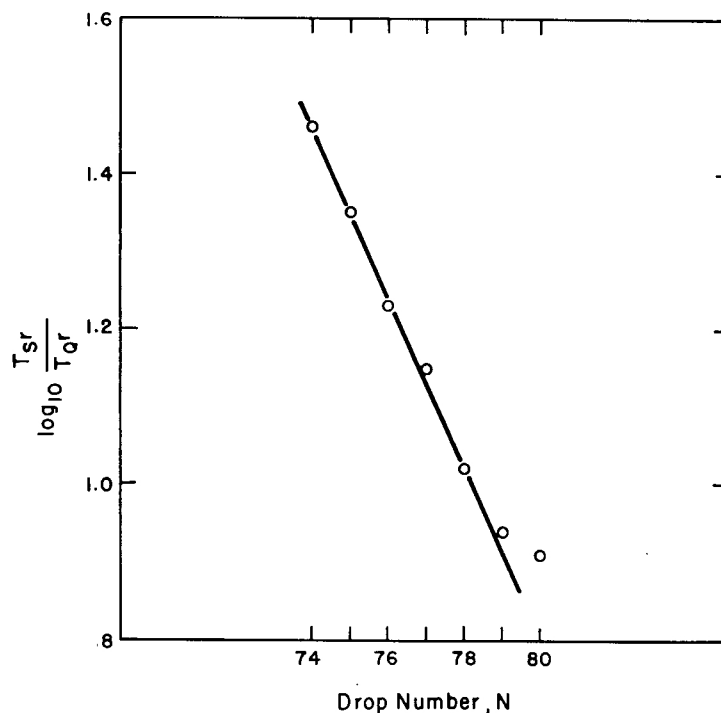
FIG. 6 (EXPT. 5).

where $\bar{\sigma}^2{}_D$ is the contribution of density heterogeneity to the total variance of the DNA $\bar{\sigma}^2{}_T$. Taking the observed value $\bar{\sigma}^2$ (marker) $= 3 \times 10^{-6}$ gm^2 cm^{-6} we obtain

$$M_n \geqslant 9 \times 10^6$$

corrected to the sodium salt of DNA, but neglecting solvation effects.[12] Using values of $\bar{\sigma}^2{}_T$ (P[32]) and $\bar{\sigma}^2$ (S^r) from experiment 8 (Table 1) and substituting in formula 11, we obtain

$$2.35 \times 10^{-3} > \bar{\sigma}_D \geqslant 1.34 \times 10^{-3} \text{ gm cm}^{-3}$$

Taking the density differences between DNA molecules containing different genetic markers as entirely due to differences in nucleotide composition, we use the relationship[2]

$$\rho_{25^\circ C} = 1.658 + 0.100 \, GC \text{ gm cm}^{-3} \tag{12}$$

where $\rho_{25^\circ C}$ is the buoyant density in CsCl at 25°C and GC is the mole fraction of guanine-cytosine, to calculate the compositional differences between molecules containing the markers S^r, B^r, A^r, and Q^r. Similarly, we estimate the nucleotide compositional heterogeneity of the DNA from its heterogeneity in density.

Assuming a molecular weight of 9×10^6, molecules containing the marker S^r have 12,000 nucleotide pairs; they have 50 more GC pairs than molecules marked with Q^r, 120 more than those marked with B^r, and differ in nucleotide composition from molecules containing A^r by less than 10 GC pairs. For randomly constructed
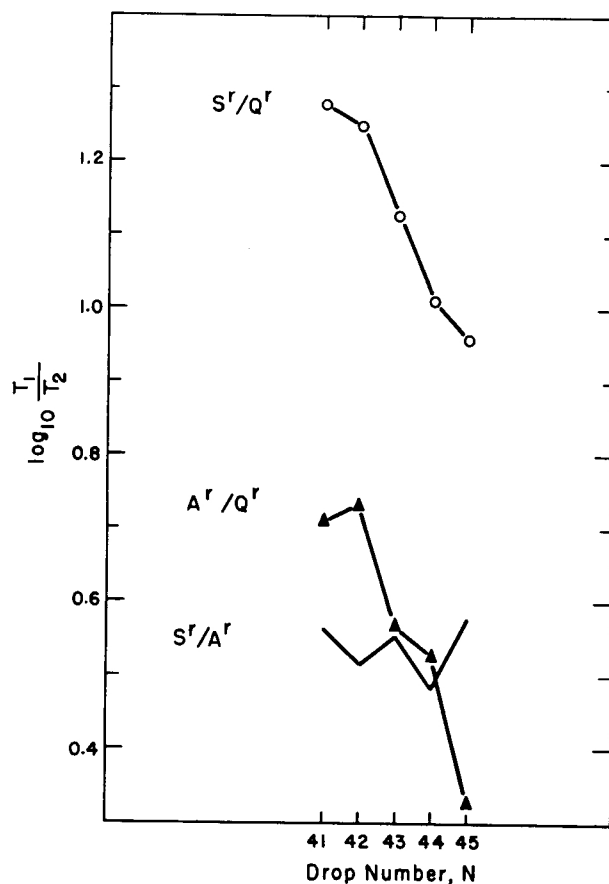
FIG. 7 (EXPT. 10).

molecules of this size (similar results also hold if M_n is greater than 9×10^6) let us write[13]

$$n = 12{,}000 \text{ nucleotide pairs}$$
$$\text{probability of } GC = 1 - (\text{probability of } AT) = 0.4 = p$$
$$\text{standard deviation} = [np \, (1 - p)]^{1/2} = 53 \; GC \text{ pairs/molecule}$$

The standard deviation of 53 GC pairs for the random model is much less than a lower bound for the standard deviation calculated from the observed lower bound for the density heterogeneity of the DNA

$$2.34 \text{ per cent } GC > \text{S.D. } (GC) = 10 \; \sigma_D$$
$$\geqslant 1.34 \text{ per cent } GC \text{ or } 160 \; GC \text{ pairs per molecule.}$$

With respect to their own mean, the nucleotide compositions found for the markers studied have a standard deviation of only 64 guanine-cytosine pairs; however it is interesting that the observed compositions for the markers all lie on the guanine-cytosine-rich side of the total DNA distribution, with a mean composition differing from that of the total DNA by 150 GC pairs, or 1.3 per cent
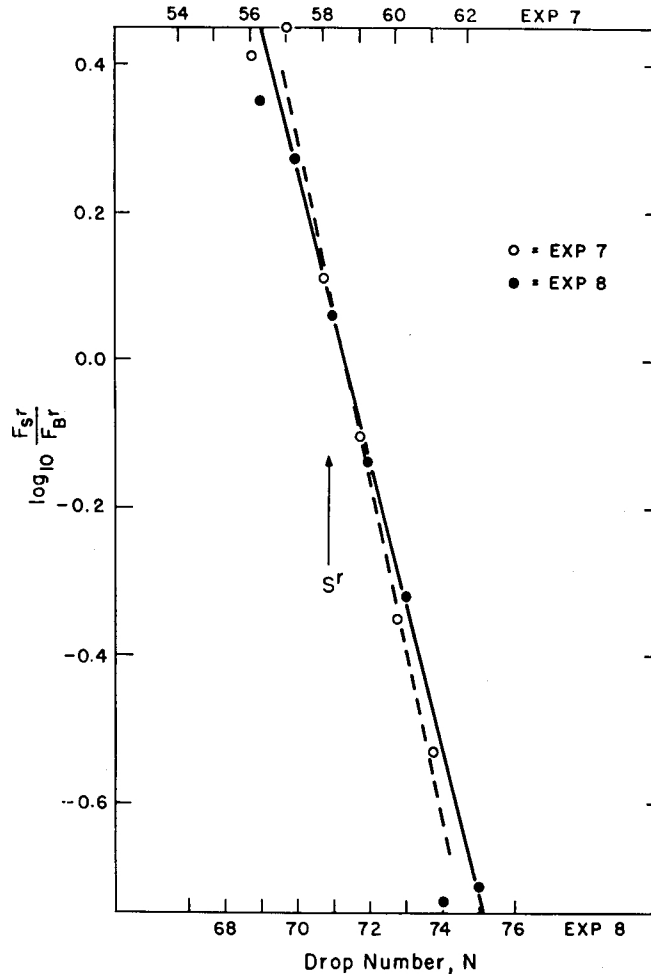
FIG. 8 (EXPTS. 7 AND 8).

FIGS. 6–8.—Logarithmic plots of ratios of numbers of transformants for pairs of genetic markers versus drop number for experiments 5, 10, 7, and 8.

guanine-cytosine. Let us assume that (1) the DNA molecules of pneumococcus that carry genetic information have the same mean composition as the bulk of the DNA and (2) the four markers studied represent a random sample of the total marker population. Using these assumptions, let us calculate P, the probability of observing the compositions found for the various genetic markers. There are two alternative ways of making the calculation, corresponding to two alternative ways of expressing the deviations of observed compositions for the markers from the mean composition for the DNA. The deviations (Table 2) may be expressed in terms of (1) the standard deviation observed for the distribution of nucleotide compositions among the DNA molecules (160 guanine-cytosine pairs per molecule), or in terms of (2) the calculated standard deviation for randomly constructed molecules (53 guanine-cytosine pairs per molecule). In each case P is the proba-

TABLE 2

NUCLEOTIDE COMPOSITIONS OF MOLECULES CONTAINING SPECIFIC GENETIC MARKERS IN PNEUMO-
COCCAL DNA; DEVIATIONS FROM MEAN COMPOSITION OF TOTAL PNEUMOCOCCAL DNA

Genetic marker	Mole fraction, guanine-cytosine	Guanine-cytosine, pairs per molecule	Deviation from Mean DNA Composition in Units of Standard Deviation for the Distribution of Nucleotide Compositions	
			Observed in the total DNA	Calculated from a random model
S^r	+0.016	+190	1.2	3.6
A^r	+0.016	+190	1.2	3.6
B^r	+0.006	+ 70	0.45	1.3
Q^r	+0.012	+140	0.9	2.6
Mean	+0.013	+150	...	...

bility that all of the marker compositions exceed the population mean by at least X standard deviations; thus in case 1, $X = 0.45$ and P is less than 0.013 and in case 2, $X = 1.3$ and P is less than 0.0001.

Discussion.—These experiments demonstrate that the DNA molecules carrying a specific genetic marker have the following properties.

(1) They are more homogeneous in buoyant density than the bulk of the DNA.

(2) They have a mean density characteristic for the marker;[14] it differs from that for other unlinked markers and from that for the bulk of the DNA.

(3) The variances of their density gradient distribution are essentially independent of the marker in question. The molecular weights obtained from the variances are in the range of values reported for the best preparations of pneumococcal DNA, as determined by sedimentation velocity studies.[11]

The experimental observations are inconsistent with a model for DNA composition in which (1) the genetic markers of pneumococcus are evenly distributed among its DNA molecules and (2) the nucleotide composition of a particular molecule is a random variable with a fixed, species-determined probability of occurrence for the guanine-cytosine pair.

It appears likely that the nucleotide composition of a DNA molecule containing a specific genetic marker is uniquely determined by the compositions of the genes of a definite linkage group comprising this molecule. This hypothesis will require modification if the existence of density heterogeneity in molecules containing the bryamycin marker is confirmed, and if in addition it can be shown that unlinked genes for the B^r phenotype are not present in the DNA. Finally it would be interesting to examine the nucleotide compositions of molecules containing other genetic markers of pneumococcus, to investigate the possibility that genetic information is confined to a special, guanine-cytosine rich fraction of the total DNA.

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

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

[3] Sinsheimer, R. L., *J. Molec. Biol.*, **1**, 218 (1959).

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

[5] Sueoka, N., *J. Molec. Biol.*, **3**, 31 (1961).

[6] Weigle, J., Meselson, M., and Paigen, K., *J. Molec. Biol.*, **1**, 379 (1959).

[7] Kindly supplied by the Bristol Drug Company.

[8] Distributions for the marker B^r differed importantly from those for the other markers. The variance of the B^r distribution was significantly larger than that for S^r in experiment 8, and in

experiment 7, the distribution for B^r was bimodal. These observations may result from gross differences in drop size, but if they are not experimental artifacts, they suggest that molecules containing the B^r marker are not homogeneous in buoyant density.

[9] A partially purified pneumococcal DNA preparation was used instead of a bacterial lysate in experiment 7. The large values of the variances for the marker distributions may be due to a small quantity of low molecular weight DNA in the preparation. The agreement between the values of $\dfrac{d \log T_S{}^r/T_B{}^r}{d\rho}$ in experiments 7 and 8 indicates that the bulk of the DNA was of a size comparable to that of the DNA studied in the other experiments. The slightly increased value for $\sigma^2(S^\lambda)$ in experiment 8 may be due to molecular scissions associated with P^{32} decay.

[10] Meselson, M., Stahl, F., and Vinograd, J., these PROCEEDINGS, **43**, 581 (1957).

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

[12] Hearst, J., Thesis, California Institute of Technology, 1961.

[13] The nucleotide composition of the DNA from the strain of pneumococcus used is 0.40 mole fraction guanine-cytosine, both by chromatographic analysis (R. Litman personal communication) and buoyant density ($\rho_{25°C} = 1.698$ gm cm^{-3}).

[14] A given genetic marker may be introduced into different strains of pneumococcus by transformation. Experiments in progress in these laboratories suggest that (1) the characteristic density of a marker is the same in different strains of pneumococcus, and (2) independent occurrences of a given marker in different strains of pneumococcus have the same characteristic density.