MATHEMATICAL ANALYSIS OF DNA DISTRIBUTIONS DERIVED FROM FLOW MICROFLUOROMETRY

PHILLIP N. DEAN and JAMES H. JETT. From the Biophysics and Instrumentation Group, Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544

INTRODUCTION

The development of flow microfluorometric (FMF) techniques (Van Dilla et al., 1969; Kraemer et al., 1972) for the measurement of DNA content of single cells at high rates (>20,000/min) has permitted a statistical precision and sensitivity not obtainable previously. It has opened up new approaches to cell characterization in general and to life cycle analysis in particular. This paper presents a computer-based mathematical technique for calculating G_1, S, and G_2 + M fractions from FMF spectra of DNA distributions. Results for the S fraction of exponentially growing cells are confirmed by [^{3}H]-thymidine autoradiography.

MATERIALS AND METHODS

The distribution of DNA content of individual cells measured for an exponentially growing population of Chinese hamster cells (line CHO) is shown in Fig. 1 B. The cells were stained with a fluorescent Feulgen procedure employing acriflavine[1] and were measured with the FMF technique. The distribution has a shape that could be predicted on theoretical grounds. All cells in the G_1 portion of the life cycle would have one "unit" of DNA, those in G_2 + M would have two units, and those in S would have between one and two units. If the staining procedure were perfectly uniform with no cytoplasmic fluorescence and the cells were measured in a system that introduced no instrumental dispersion (broadening due to electronic noise, light scattering, etc.), the distribution of the cells in Fig. 1 B would appear as shown in Fig. 1 A. Both the G_1 and G_2 + M subpopulations would fall into single channels of the pulse-height analyzer a factor of two apart. The S subpopulation shape is determined by the rate of DNA synthesis and any synchrony present. Since this shape currently is unknown and may be variable, the S distribution is

[1] H. A. Crissman. 1973. Cellular DNA measurements by flow microfluorometry: preparative techniques. In preparation.

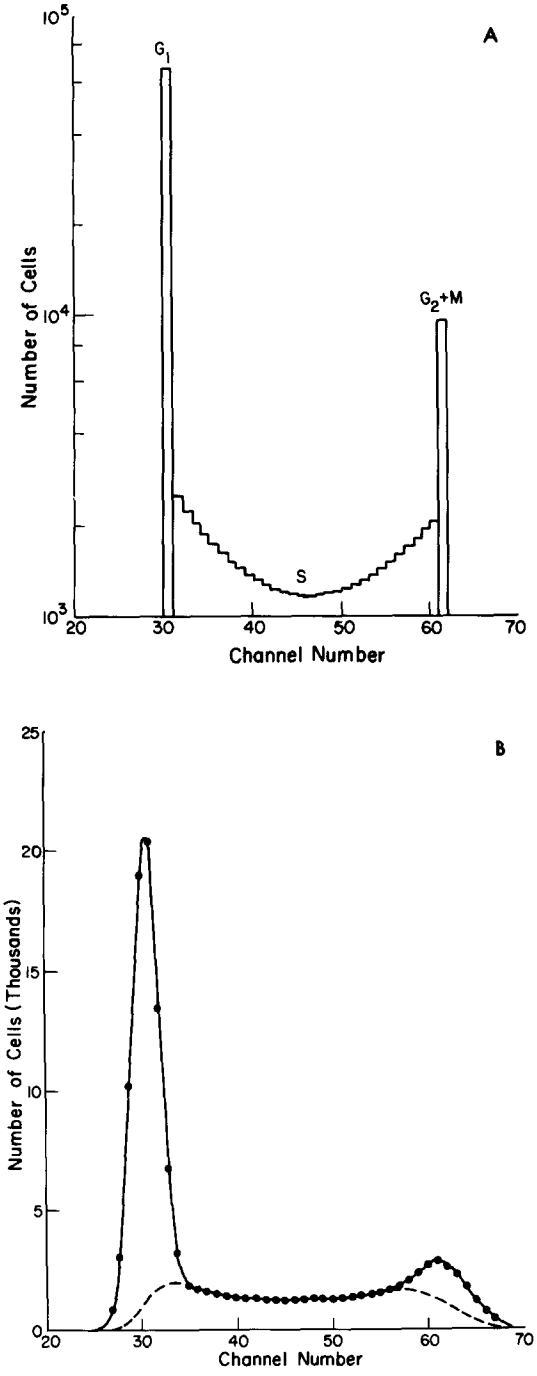

FIGURE 1 Cellular DNA distribution for exponentially growing CHO cells examined by FMF. The experimentally measured spectrum is shown in B. Solid circles are the data points, the solid line represents the computer fit to the data, and the dashed line is the broadened polynomial as described in the text. The unbroadened version of this spectrum is shown in A. The three subpopulations are identified by letter.

arbitrarily represented here as a second-degree polynomial, a function that seems to fit the data well to a first approximation. The difference between Fig. 1 A and 1 B, which are different kinds of plots of the same data, is due to the broadening effects mentioned above. If the broadening is random from cell to cell, then the G_1 and G_2 + M subpopulations could be described by normal curves. All points of the S subpopulation would also exhibit the same broadening. The solid line in Fig. 1 B is the result of fitting the data (solid circles) with two normal curves plus a second-degree polynomial by nonlinear least squares techniques. The parameters of the fit are shown in Table I. Although any parameter may be held fixed, in practice all nine parameters are allowed to vary. The complete function used is:

$$Y = \frac{A_1}{\sqrt{2\pi}\,\sigma_1} e^{-1/2\left(\frac{X-\bar{X}_1}{\sigma_1}\right)^2} + \frac{A_2}{\sqrt{2\pi}\,\sigma_2} e^{-1/2\left(\frac{X-\bar{X}_2}{\sigma_2}\right)^2} + P(X),$$

where Y is the number of cells in channel X, A_1 and A_2 are the numbers of cells in G_1 and G_2 + M, σ_1 and σ_2 are the standard deviations of the normal curves, $\bar{X}_1$ and $\bar{X}_2$ are the modes of the two peaks, and P(X) is the second-degree polynomial. The polynomial is broken into a histogram, existing only between $\bar{X}_1$ and $\bar{X}_2$, each channel of which is then replaced by a normal curve with the same coefficient of variation as the G_1 peak. The coefficient of variation of the G_1 peak is used, rather than an average of the G_1 and G_2 + M peaks, because the parameters of the G_1 peak, in general, are better determined statistically. Also, if all broadening is due to instrumental and preparational artifacts and not to true variation in the amount of DNA in individual cells, the coefficient of variation would be constant for all amounts of DNA. The shape of the broadened polynomial is shown in Fig. 1 B as a dashed line where it is different from the total fit.

RESULTS

The validity of the assumption of a normal curve for the G_1 and G_2 + M peaks and the accuracy of the method of resolving the FMF spectrum into its three components have been tested experimentally. In the experiments described, the cells were carefully prepared to avoid problems with cells clumping together. This must always be done for a proper analysis to be performed. Two G_1 cells clumped together would give the same result as one G_2 cell. In fact, this did not occur here as evidenced by the spectra described below. If there were clumping, one would expect to see cells at the 6C and 8C quantities. Peaks at these

TABLE I

Parameter Values Resulting from Mathematical Analysis of the Data shown in Figs. 1 B, 2 A, and 3

Parameter	Fig. 1 B	Fig. 2 A	Fig. 3
A_1	$6.74 \pm 0.14 \times 10^4$	$9.95 \pm 0.10 \times 10^4$	$8.20 \pm 0.72 \times 10^3$
σ_1	1.41 ± 0.02	1.96 ± 0.02	1.39 ± 0.07
$\bar{X}_1$	30.64 ± 0.03	33.67 ± 0.02	24.13 ± 0.10
A_2	$9.91 \pm 1.50 \times 10^3$	$2.82 \pm 1.08 \times 10^3$	$6.22 \pm 0.99 \times 10^3$
σ_2	2.60 ± 0.26	2.73 ± 0.30	2.07 ± 0.21
$\bar{X}_2$	62.07 ± 0.27	65.78 ± 0.29	47.39 ± 0.20
α	$1.11 \pm 0.23 \times 10^3$	$6.59 \pm 1.08 \times 10^3$	$6.67 \pm 2.41 \times 10^3$
β	$-4.23 \pm 1.03 \times 10^2$	$-2.60 \pm 0.44 \times 10^2$	$-2,90 \pm 1.38 \times 10^2$
γ	4.51 ± 1.14	2.66 ± 0.44	3.86 ± 1.96
W.V.*	4.55	5.33	4.16

The $\pm$ numbers are one standard deviation as estimated from the least squares fit to the data. The latter three parameters are the coefficients of the polynomial used to fit the S population, $P(X) = \alpha + \beta X + \gamma X^2$.

* W. V. is the weighted variance as defined in the text.

levels were not observed. To determine the shape of the peaks, two populations of CHO cells were used. One batch of cells was maintained in suspension culture until it grew to confluency and the total number of cells stabilized. A sample of these cells was measured by FMF. The spectrum obtained is shown in Fig. 2 A. Almost all cells are in G_1, and the small number of cells in other phases can have little influence on the shape of the peak which is fit very well with the normal curve. Another population of cells was synchronized by mitotic selection (Tobey et al., 1967). They were measured about 15 min after selection at which time about one-third of the cells had divided. The spectrum is shown in Fig. 2 B. Virtually all of these cells are in either M or G_1. As in Fig. 2 A, the few cells in S can have little influence on the shapes of the peaks. In this case, both peaks are fit very well with normal curves. In addition, the mean of the M peak is twice the mean of the G_1 peak as it should be, and the coefficient of variation (standard deviation/mean) is approximately the same. An additional experiment was performed using L5178Y cells in exponential growth. These cells were selected because of their short G_1 and long S phases. The spectrum is shown in Fig. 3. In this case, there are significantly more cells in S than in the CHO cell spectra. Even with this large S population, the data are fit very well with the function used. The parameters resulting from the analysis of all spectra are shown in Table I. For all spectra, the "weighted variance"[2] ranged from 2–10 with an average of about 5, indicating that the mathematical function used describes the data well.

Autoradiographic techniques were used to validate the calculation of the fraction of cells in the S phase of the life cycle. The cells used for Figs. 1 A and 3 were also pulse-labeled with [^{3}H]thymidine prior to staining. In addition, a population of WI-38 cells was also pulse labeled and analyzed by FMF. Autoradiographs were then made, and the fraction of cells in S was obtained by visual scoring. The results of these measurements are shown in Table II. The agreement between the methods is clearly excellent. To test reproducibility, a batch of exponentially growing CHO cells was separated into 10 samples, each of which was pulse labeled, stained, autoradiographed, and measured by FMF. The aver-

[2] The "weighted variance" is the sum of squares of the weighted deviations divided by the number of degrees of freedom. The weighting factor used here is the reciprocal of the statistical error (square root of the number of counts). If the function and the weighting factor fit the data perfectly, the weighted variance would be exactly one. Since we are probably not including all variances in the weighting factor and there is no reason to believe the polynomial truly represents the shape of the S distribution, the weighted variance in our calculations is not one but averages about five. In our experience, a "good" fit would have a weighted variance in the range of four to six.

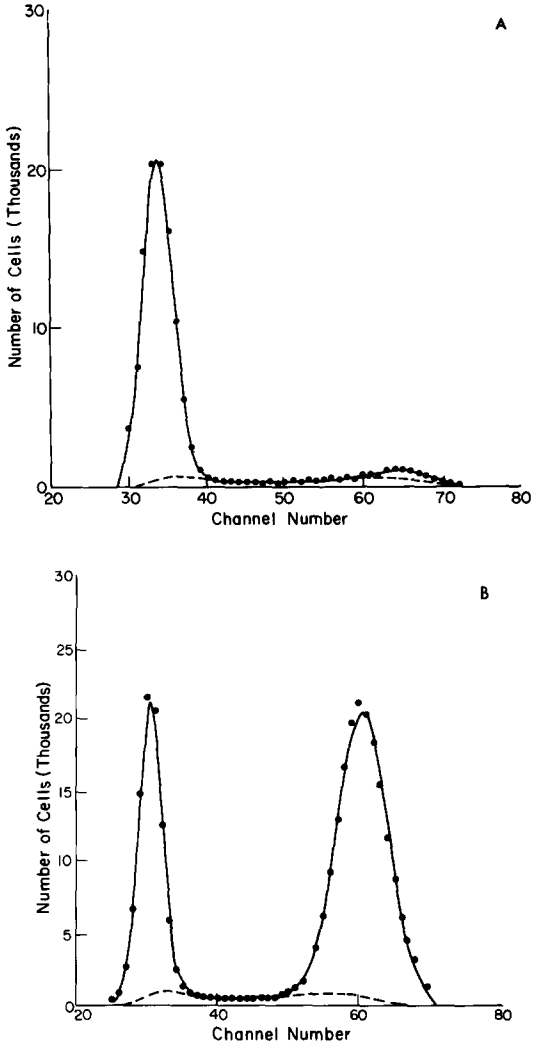

FIGURE 2 DNA distributions for two different populations of CHO cells. A population of cells was allowed to grow to confluency such that the total number of cells was not changing with time. A sample of these cells was removed from the culture, and the spectrum shown in A was measured with virtually all cells in G_1. Another population of cells was synchronized by mitotic selection. The spectrum in B was measured 15 min after selection when the cells were partly in G_1 (the peak at channel 30) and partly in M (peak at channel 60). In both spectra, the solid circles are the data, the solid line is the fit to the data, and the dashed line is the broadened polynomial.

age percent of the cells in S was 37.4 ± 1.1 by FMF and 36.7 ± 2.1 by autoradiography. One of the autoradiography results was not used because it differed from the average by more than five standard deviations. The FMF analysis also

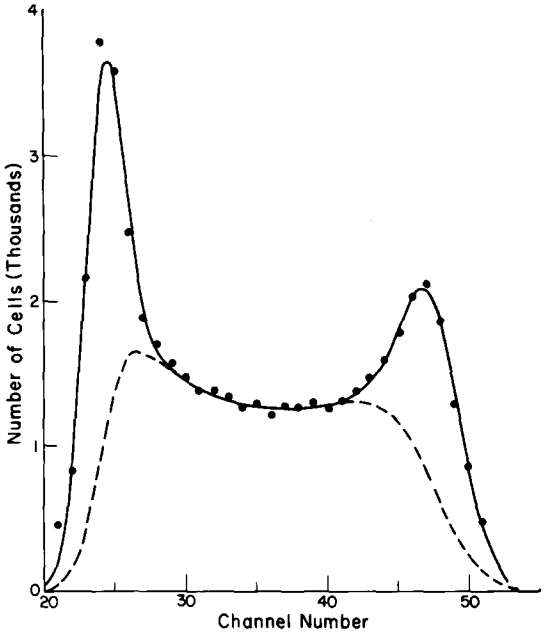

FIGURE 3 DNA distribution for L5178Y S-S (radiation-sensitive mutant) cells in exponential growth. The solid circles are the data, the solid line is the fit to the data, and the dashed line is the broadened polynomial.

TABLE II

Percent Cells in S by Autoradiography and FMF and Percent Cells in G_1 and G_2 + M by FMF

	S			
Cell Population*	Autoradiograph	FMF	G_1	G_2 + M
CHO (1 A)	37.8	38.9	53.6	7.5
L5178Y S-S (3)	63.0	64.3	20.1	15.5
WI-38	36.3	38.9	53.5	7.6
L5178Y	51.8	54.8	38.6	6.6

* Figure number of spectrum is in parentheses.

yielded 54.4 ± 1.3% cells in G_1 and 8.3 ± 1.0% cells in G_2 + M.

DISCUSSION

The mathematical algorithm selected for analyzing DNA distributions derived by FMF techniques has been demonstrated to yield very accurate results for the S fraction for cells in exponential growth. In a measurement time of 1–5 min and an analysis time of a few seconds on a CDC-6600 computer,[3] a life cycle analysis can be

[3] Control Data Corp., Minneapolis, Minn.

performed from a single sample. The calculated fractions of cells in the different phases of the life cycle can be converted easily to fractions of the generation time assuming the cells are in exponential growth. If the generation time is known, then the phase durations can also be computed. The data of Figs. 1–3 and Table II show that the technique is valid for a wide range of S, G_1, and G_2 + M subpopulation fractions. The autoradiographic experiments confirm the calculation of only the S fraction of cells. Complete biochemical analysis of the life cycle was not performed for the specific cells used in the experiments described. Detailed experiments (Puck et al., 1964) performed on CHO cells with a generation time of 12.4 h yielded G_1, S, and G_2 + M fractions of 0.46, 0.317, and 0.223, respectively. The generation time of cells used in the current experiments was 16 h. Recent data[4] have shown that the durations of S and G_1 depend on cell concentrations and culture conditions and can change daily. Therefore, there can be no direct comparison between results reported here and in previous experiments. Indeed, difficulties in making the biochemical measurements result in large errors in the fractions obtained. The much simpler FMF analysis may become the standard method of life cycle analysis. We believe that it is advisable to perform a life cycle analysis each time cells are used in an experiment, an easy task with the techniques described in this paper.

As mentioned earlier, the shape of the S distribution is determined by the rate of DNA synthesis across the S phase of the life cycle. In principle, the shape of the polynomial used to represent the S phase then should yield information on the rate of DNA synthesis. A method of analysis analogous to the Collins-Richmond method for determining rate of volume growth (Collins and Richmond, 1962; Anderson et al., 1969) is now being attempted, and the results will be reported at a later date.

[4] H. A. Crissman, J. H. Jett, and D. F. Petersen. 1973. Life cycle analysis of mammalian cells by flow microfluorometry and labeled DNA precursors. In preparation.

We wish to acknowledge the valuable assistance of Dr. H. A. Crissman in making the FMF measurements and Dr. D. F. Petersen and Mrs. P. C. Sanders for performing the autoradiography.

This work was performed under the auspices of the United States Atomic Energy Commission.

Received for publication 30 July 1973, and in revised form 12 November 1973.

REFERENCES

ANDERSON, E. C., G. I. BELL, D. F. PETERSEN, and R. A. TOBEY. 1969. Cell growth and division. IV. Determination of volume growth rate and division probability. *Biophys. J.* **9**:246.

COLLINS, J. F., and M. H. RICHMOND. 1962. Rate of growth of *Bacillus cereus* between divisions. *J. Gen. Microbiol.* **28**:15.

KRAEMER, P. M., L. L. DEAVEN, H. A. CRISSMAN, and M. A. VAN DILLA. 1972. DNA constancy despite variability in chromosome number. *In* Advances in Cell and Molecular Biology. E. J. DuPraw, editor. Academic Press, Inc., New York. **2**:47.

PUCK, T. T., P. SANDERS, and D. F. PETERSEN. 1964. Life cycle analysis of mammalian cells. II. Cells from the Chinese hamster ovary grown in suspension culture. *Biophys. J.* **4**:441.

TOBEY, R. A., E. C. ANDERSON, and D. F. PETERSEN. 1967. Properties of mitotic cells prepared by mechanically shaking monolayer cultures of Chinese hamster cells. *J. Cell. Physiol.* **70**:63.

VAN DILLA, M. A., T. T. TRUJILLO, P. F. MULLANEY, and J. R. COULTER. 1969. Cell microfluorometry: a method for rapid fluorescence measurement. *Science (Wash. D.C.).* **163**:1213.