Discrimination Between Human T and B Lymphocytes and Monocytes by Computer Analysis of Digitized Data From Scanning Microphotometry. I. Chromatin Distribution Patterns

By B. G. M. Durie, L. Vaught, Y. P. Chen, G. B. Olson, S. E. Salmon, and P. H. Bartels

Digitized images of Feulgen-stained normal human B and T lymphocytes and monocytes were analyzed by computer and microphotometry. The algorithms assessed the chromatin distribution patterns. A substantial number of parameters descriptive of the chromatin distributions were found to show statistically highly significant differences between B and T cells and monocytes. Although the effects of a number of biologic variables, such as cell cycle phase and immune stimulation, need to be carefully evaluated, the preliminary results indicate potential for automatic classification and subclassification of human B and T lymphocytes in normal and disease states.

DISTRIBUTION PATTERNS of nuclear chromatin have been shown to be highly specific for cells of a given type or functional state.[1-3] Chromatin texture has been quantitatively analyzed to assess the proportions of euchromatin and heterochromatin.[4] Such studies have yielded a number of descriptors helpful for discrimination among leukocytes.[5]

Nuclear texture plays an important role in the computer discrimination of normal and malignant cells.[6] Recently it was shown that chromatin texture in lymphocytes was descriptive for differentiation of murine T and B cells.[7] These descriptors respond in a sensitive manner to small doses of x-irradiation.[8] Changes in the texture of nuclear chromatin were used to follow the effects of antineoplastic agents on peripheral blood lymphocytes in vivo.[9] During these studies it was found that computer assessment of the chromatin distribution patterns could render a clearcut discrimination between B and T cells in mouse thoracic duct lymphocytes.[10] Feasibility studies have shown that analogous differences can be found between human B and T lymphocytes.

It is the purpose of this study to describe the textural properties and the chromatin distribution patterns of Feulgen-stained human B and T lymphocytes.

MATERIALS AND METHODS

Human lymphocytes were obtained from seven normal adults, ages 21–43 yr. All individuals had prior lymphocyte studies in vitro documenting normal T and B rosette levels and responses

From the Section of Hematology and Oncology, Department of Internal Medicine, and the Optical Sciences Center and Department of Microbiology, University of Arizona, Tucson, Ariz.

Submitted February 14, 1977; accepted November 3, 1977.

Supported in part by Grants CA17094, CA14783, and CA15641 from the National Cancer Institute.

Dr. Durie is a Scholar of the Leukemia Society of America.

Address for reprint requests: B. G. M. Durie, M.D., Section of Hematology and Oncology, Dept. of Internal Medicine, University of Arizona, Tucson, Ariz. 85724.

 ISSN 0006-4971/78/5104-1001$02.00/0

Table 1. Results of Lymphocyte Studies In Vitro on the Normal Adults Used as Controls

Dose ($\mu g/10^6$ Cells)	Stimulation Indices* of Patient No.: 1	2	3	4	5	6	7
Phytohemagglutinin response							
1	184	160	64	170	102	280	171
5	154	140	50	145	52	260	164
50	41	30	9	35	9	84	44
Pokeweed mitogen response							
1	39	31	25	39	9	20	21
5	47	75	59	51	28	59	41
50	48	94	65	61	25	63	61
E rosettes (T)† (%)	83	76	84	71	77	71	75
EAC rosettes (B)† (%)	15	17	17	17	12	16	15

*Mean of duplicate or triplicate cultures (mitogen/resting culture).

†Mean of duplicates.

in vitro to phytohemagglutinin and pokeweed mitogen and mixed lymphocyte responses. The details of these control data are outlined in Table 1.

Lymphocyte Separation

WHO/IARC guidelines were followed for lymphocyte preparation from heparinized whole blood by centrifugation over Ficoll-Hypague and B and T cell enumeration by rosette assays.[11] B lymphocyte-enriched preparations were obtained by centrifugation of E (T cell) rosettes over Ficoll-Hypaque at 4°C and collection of the free lymphocyte-rich layer.[12] Conversely, free T lymphocytes were obtained by enrichment of EAC (B cell) rosettes. Both T and B cell preparations were based on multiple enrichment procedures, since the sedimentation and removal of free lymphocytes from the interface were carried out three to six times. This procedure was slightly more successful with the more stable B lymphocyte rosettes than with the more readily disrupted T lymphocyte rosettes.

The purity of the free cell preparations was verified by E and EAC rosetting of cell suspensions. Thus the entire enrichment process was monitored by stepwise rosetting purification and further rosetting to check for cell-type enrichment. While there were a number of more powerful procedures for further purification of the separated lymphocytes available, no further purification was attempted, thus avoiding the danger of possible depletion of subpopulations or inducing artifacts.

The monocytes in the B cell preparations were identified by diffuse cytoplasmic staining for nonspecific esterase.[13] Owing to the overlap of spectral absorbance of the esterase stain and the Feulgen chromophor, the identification of monocytes had to be carried out as a two-stage process. It was employed primarily to check on the efficacy of the optical density (OD) histogram-derived discriminator for monocytes described in Results. Cytocentrifuge smears of the free cell suspensions were made using a Shandon cytocentrifuge and were stained according to the Feulgen procedure.[14]

Cell Scanning

Scanning microphotometry was used to record the microscopic images of cells as digitized data on magnetic tape. The digitized images of Feulgen-stained cells were recorded on a Leitz MPV II microscope photometer operated on line to a PDP 11/45 computer. A scanning spot size of 0.5 x 0.5 sq μm and light wavelength 560 μm were employed. All recorded cell images were edited immediately to eliminate erroneous OD points from the scanned area.

From each of the seven normal subjects, 100 cells each from the T and B cell preparations were scanned. Preliminary analysis of the data indicated a rather homogenous population of cells in the T cell preparations, with 5%-10% of the cells having features of the B cells. This contamination corresponded to percentages found by rosetting. The B cell preparations had a small percentage of cells (< 5%) with T cell features as well as a proportion of cells clearly having features of monocytes, identified by esterase staining. This latter group of cells had large nuclei

with light Feulgen staining. A composite descriptor formed from the cumulative frequency of occurrence of OD values in the range 0.02 (threshold) to 0.30 rendered an excellent objective discriminator for monocyte recognition.

Information Extraction

Image descriptors were computed for every cell using a series of feature extraction subroutines of the TICAS 11/45 program package. The descriptors computed and used were relative nuclear area, total optical density, average optical density, chromatin texture as expressed by the histogram of optical density values and the transition probabilities of a Markovian-dependence scheme,[15,16] measures of chromatin condensation,[17] and measures of chromatin granularity. These features are defined as follows:

The relative nuclear area was computed as the number of image points with an OD value in excess of 0.02. The total optical density was computed as the sum of all OD values recorded for that image and the average OD formed as the ratio of these first two descriptors. To assess the chromatin distribution patterns, the following features were computed: The histogram of OD values was formed. For the optical density range 0.02–1.80 (lower threshold and upper limit set by the photometric accuracy of the instrument) and a histogram interval of 0.10 OD units this resulted in 18 descriptors, each expressing the relative frequency of occurrence of values in an OD interval. Also computed was the variance of OD values in the nucleus. One feature extraction routine computed the averaged radial OD profile along the cell radius: the center of the cell was found, the digitized image was converted to polar coordinates, and the optical density along circles around the center of the cell was computed. This analysis was done at distances of 0.5, 1.0, 1.5, etc. μm from the cell center. Finally, the OD values on each circle were averaged and stored as a radial OD profile.

To assess the tendency of chromatin to be concentrated either near the center of the nucleus or near the periphery, the "second moment" of the OD values about the cell center was computed. Each OD value was squared and multiplied by its radial distance from the cell center. All of the resulting products were summed to form the descriptor value.

Texture has frequently been described by the transition probabilities of a Markovian-dependence scheme. A Markov system may exist in several different states; in a cell image each point may fall into one of several different OD ranges. Into which range a point falls depends with a certain probability on the OD range into which the point adjacent to it and preceding it along the line scan fell. A texture may be described by the probabilities with which values from a given OD range are followed by values in another OD range, i.e., by a set of transition probabilities.

The OD ranges were 0.20 OD units wide, and the transition probabilities were normalized by row vectors in the matrix. Only the matrix elements on the diagonal and two values above the diagonal were used in order to exclude excess descriptors carrying very little information. For the OD range 0.02–1.80 and the chosen interval width, this amounted to nine ranges and 24 matrix elements being used.

Three measures of chromatin condensation, CDN1, CDN2, and CDN3, were computed according to the definitions given by Vidal et al.[17] However, they were scaled so that the value of each fell into the range from zero to unity.

The feature CDN1 measured the peripheral tendency of the nuclear chromatin. If all chromatin was concentrated at a circular nuclear membrane, the feature CDN1 would assume its highest value, unity. For chromatin concentrated near the nucleus' center, CDN1 would assume a low value.

The feature CDN2 was computed as the ratio of circumference/area of chromatin granules having at least a certain optical density.

Feature CDN3 expressed the ratio of average optical density found in the highly condensed heterochromatin to the average optical density in the diffuse euchromatin of the nucleus.

Chromatin granularity was assessed by the feature extraction program CHRGRAN. Mean OD histograms for two selected files (e.g., B and T cells) were computed and the histograms displayed. Thresholds could then be selected in an effort to discriminate optimally between two different cell populations. The program had available 12 features for use in discrimination: the total number of granules (single points or groups of points), the number of points with an OD above the selected threshold OD, five features expressing the relative frequencies of granules in

Table 2. Summary of Basic Discriminators for T and B Lymphocytes and Monocytes

	B Lymphocytes	T Lymphocytes	Monocytes
Relative nuclear area*	126.6 ± 1.18	152.0 ± 1.21	188.5 ± 2.6
Total optical density* (arbitrary units)	4055 ± 23	3634 ± 24	3311 ± 52
Average optical density*	33.6 ± 1.8	21.6 ± 1.1	17.7 ± 2.4

Statistical methods described in Materials and Methods.

*Mean values: Sample size 700 cell images: 100 cell images (of each type) from each of seven control individuals. SEM based on error term mean square from the analysis of variance.

five size classes, and five features computed as the relative frequency of granule area divided by the granule perimeter length (scaled zero to unity).

Statistical Analyses

Analysis of variance was carried out using a mixed model factorial arrangement to evaluate differences between preparations from each normal individual as well as between the B and T lymphocytes and monocytes.[18] The SEM was calculated from the error term mean square.

To establish the statistical significance between values found in B and T cells for the 24 features derived from the transition probability matrix, a nonparametric statistic, the Kruskal-Wallis test was used.[19] This statistic makes no assumptions as to the manner of distribution of two sets of variables, but merely ranks their values, assigns rank order numbers, and tests whether or not the average rank from one set is significantly lower than that found for the other set.

Reference is made later in the text to a Bayesian decision boundary. It is drawn so that on one side the likelihood is greater that an observation came from one given distribution and on the other side that it came from another given distribution. The Bayesian boundary equalizes the risk of an erroneous assignment of an observation to either one of the two distributions.

RESULTS

Computation of relative nuclear area, total OD, and average OD showed a marked trend of values from the B lymphocytes to T lymphocytes and monocytes. B lymphocytes were small, and their chromatin stained densely. T lymphocytes were, on the average, approximately 20% larger in nuclear area and less densely stained. Monocytes were noticeably larger and characterized by very light Feulgen staining. Analysis of variance showed these differences in cell image features between cell types to be highly significant ($p = 0.001$). The mean square due to sample collection from seven normal individuals was found to be statistically not significant. The SEM given in Table 2 was based on the error term mean square.

The OD histograms showed a similar trend, as seen in Fig. 1. By adding up the relative frequency of occurrence of OD values in the first three histogram intervals we obtained a cell image descriptor that allowed a ready discrimination between B lymphocytes (B) and monocytes (M). For a mixed sample of B lymphocytes and monocytes computation of the values of this composite cell image descriptor resulted in a bimodal distribution, as seen in Fig. 2.

B lymphocytes have denser granules than T lymphocytes and monocytes. This difference is expressed not only in the OD histogram but also in the variance of nuclear OD values. The mean variance computed for the sample of 700 B lymphocytes was 395.2, as compared to values of 214.0 and 95.6 for T cells and monocytes, respectively. The distribution of chromatin in monocytes and T cells is reflected in the radial OD profiles, as seen in Fig. 3. The more uniform distribution of Feulgen chromophor in T cells and monocytes, as

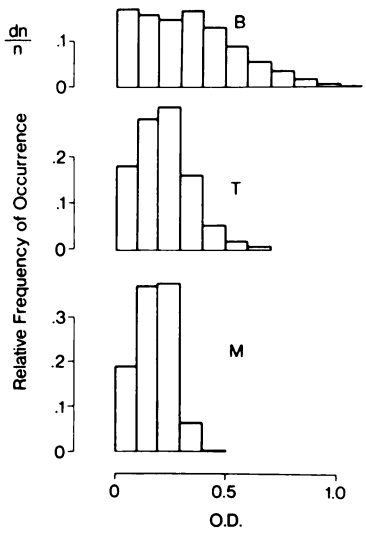

Fig. 1. Histogram of OD values for B and T lymphocytes and monocytes (M). Relative frequency of occurrence of OD values in the digitized images plotted against OD. Each of the B and the T lymphocyte histograms based on 700 cell images. Histogram for monocytes based on 75 cell images.

opposed to B cells, made it possible to find transition probability features with good potential for the discrimination between cells of these cell types. The Kruskal-Wallis test showed that for numerous elements of the transition probability matrix statistically significant differences existed. These textural features could be expected to render good discrimination. The transition probability profiles, averaged over the entire cell sample of B cells, the entire sample of T cells, and over the monocytes, are shown in Fig. 4.

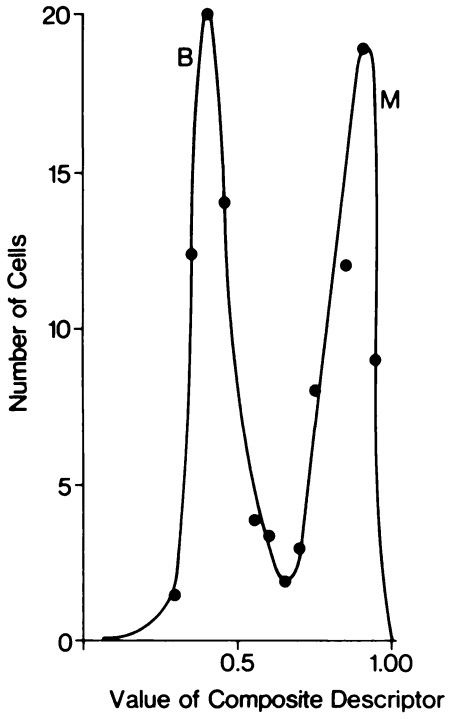

Fig. 2. Distribution of composite descriptor formed by summing relative frequencies of occurrence of OD values in the first three histogram bins for a set of B cells and monocytes (M). Relative number of cells, ordinate; value of composite descriptor in relative units, abscissa. Composite descriptor lends itself as an efficient discriminator between B lymphocytes and monocytes.

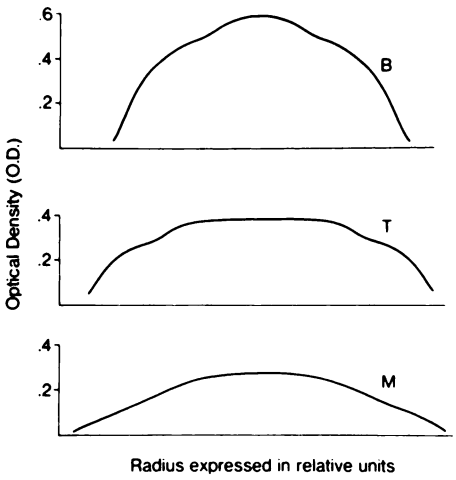

Fig. 3. Radial profiles of OD for Feulgen-stained images of B and T lymphocytes and monocytes (M). OD ordinate; cell radius in relative units, abscissa.

Chromatin condensation was evaluated by the three texture descriptors CDN1, CDN2, and CDN3 for B cells, T cells, and monocytes. The results are illustrated in profile form in Fig. 5. It can be seen that the relative values are similar for the B and T groups, but those for monocytes are significantly different. By way of example, the "receiver operating characteristic curve" for discrimination between T lymphocytes and monocytes on the basis of the texture descriptor CDN1 is shown in Fig. 6. The high value found for the measure of detectability indicates the good discrimination potential of this measure of chromatin condensation.

As a further illustration, the texture features CDN1 and CDN2 are plotted in Fig. 7 as bivariate distributions. Shown are the 95% confidence regions for the bivariate means and the 50% tolerance regions. The confidence regions for the B and T cell populations overlap almost completely, but a clear statistical difference exists for the monocyte cell population. CDN1 and CDN2 thus have

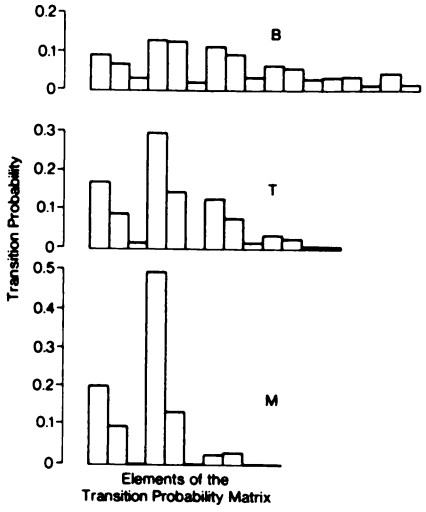

Fig. 4. Profiles of transition probabilities between OD values in the digitized images of B and T lymphocytes and monocytes (M). Transition probability in relative units, ordinate; linear arrangement of the elements of the transition probability matrix, abscissa. Since only the diagonal elements of the matrix and two elements above the diagonal are used, the sequence of matrix elements is as follows: elements 1,1;1,2; 1,3; 2,2; 2,4; 2,5; 3,3; 3,4; 3,5; etc.

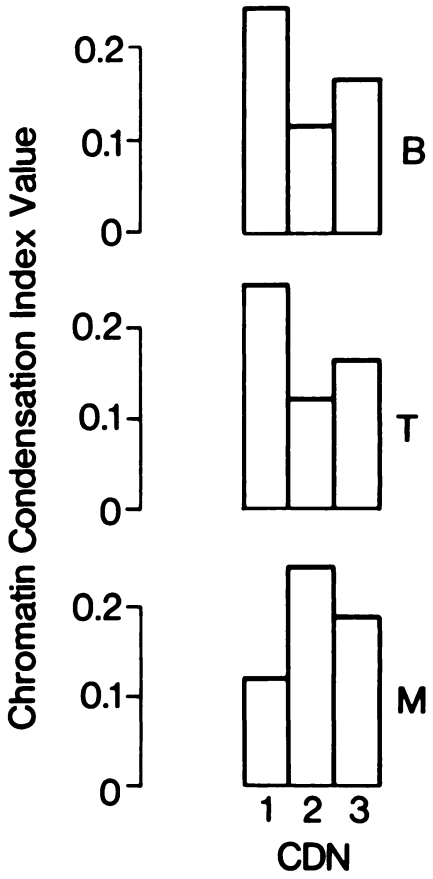

Fig. 5. Profiles of the chromatin condensation indices CDN1, CDN2, and CDN3 for B and T lymphocytes and monocytes (M). Index plotted in relative units on the ordinate; for each of the three indices possible range of values extends from zero to unity.

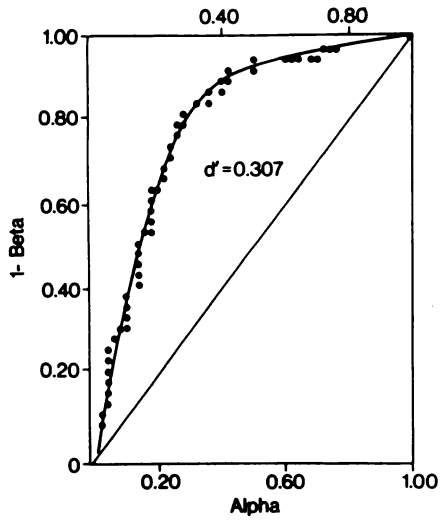

Fig. 6. Receiver operating characteristic (ROC curve) for discrimination between T lymphocytes and monocytes by means of the chromatin condensation measure CDN1. "Power of the test" $1 - \beta$, ordinate; chances for error of the first kind (α), abscissa. Detection measure $d' = 0.307$; its values can span a range from zero (indicating no discrimination) to 0.50 (complete discrimination).[24]

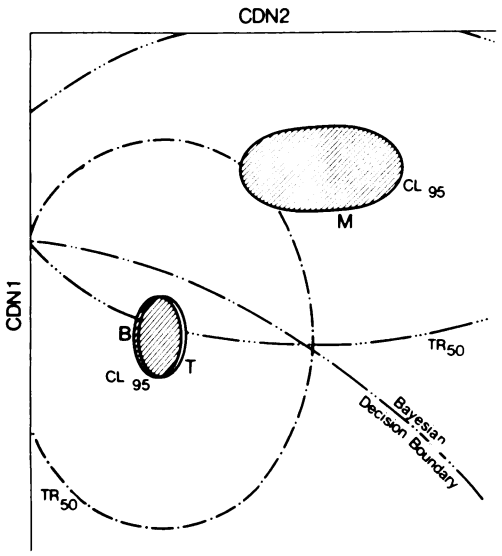

Fig. 7. Bivariate distributions of the two chromatin condensation measures CDN1 and CDN2 for B lymphocytes, T lymphocytes, and monocytes (M). Bivariate mean for B and T lymphocytes and the 95% confidence regions (CL_{95}) for these means overlap practically completely. Bivariate mean for monocytes is significantly different. Also indicated are the 50% tolerance regions (TR_{50}) and the Bayesian decision boundary. Latter illustrates that even though the bivariate mean for B and T lymphocytes on one hand and monocytes on the other are significantly different, classification using a Bayesian rule would lead to a substantial number of misclassifications if the two discriminators CDN1 and CDN2 only were employed.

potential for the discrimination of monocytes from B and T cells. The two features CDN1 and CDN2 are uncorrelated, a desirable property in features used in discrimination (correlation coefficient $r_{CDN1,CDN2}$ = 0.04 for T cells, −0.01 for monocytes).

Finally, the assessment of chromatin granularity in B cells, T cells, and monocytes rendered image descriptors that showed good potential for a machine recognition of these cell types. The program CHRGRAN permitted the user to set a threshold in the OD histogram, and only OD values above this threshold were considered for the computation of the granularity descriptors. This threshold may be chosen so that differences between cell types are accentuated. In this case, the threshold was set at OD = 0.40 to optimize discrimination between T cells and monocytes. The profiles shown in Fig. 8 represent the values obtained for each of the twelve features computed by the program, as explained in Materials and Methods.

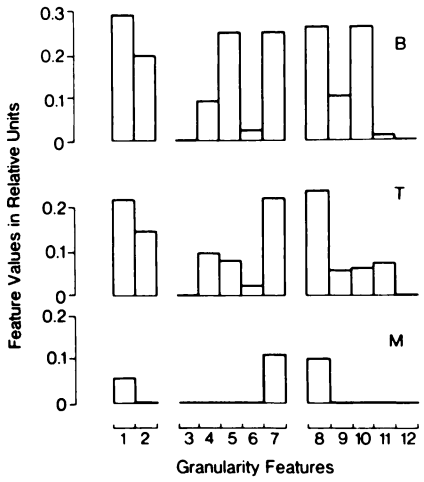

Fig. 8. Profiles of chromatin granularity measures for B lymphocytes, T lymphocytes, and monocytes (M). As explained in the text, a threshold was set so as to render optimum discrimination between monocytes and T cells. Profile formed by twelve descriptors: these fall into one group of two granularity measures and then two groups combining five granularity measures each. Each of the descriptors is expressed in relative units, with a range of zero to unity.

DISCUSSION

Various methods have been used to distinguish human T and B lymphocytes and monocytes,[11,20,21] each with its relative merits and disadvantages. Beginning from studies of mouse thoracic duct lymphocytes,[7,10] we attempted in this study to identify a number of image features with good potential for discrimination between T and B lymphocytes, and monocytes in Feulgen-stained cytocentrifuge preparations of human peripheral blood. Detailed analysis of the chromatin distribution patterns in cell samples collected from normal human adults produced, for every feature extraction algorithm, cell image features with statistically highly significant differences between B cells, T cells, and monocytes.

The significance of the cell-type-specific differences and the reassuring finding that differences between samples from seven normal adults were statistically not significant both suggest that a computer classification of these cells based on discriminators derived from the chromatin distribution patterns is promising.

Statistical differences between the distributions of any feature's value for two cell populations are a necessary but not sufficient condition for good discrimination. In automated discrimination the classification of an individual cell is the objective, and for this the dispersion of feature values in the classification space is decisive. Such classification will undoubtedly be based on several features, and the classification success will depend on the variance-covariance matrices for the cell populations to be discriminated as well as on the differences in mean feature vectors.

[Part II of this study, to be published separately, will report the success attained in establishing a classification rule using the features described here. It will report the error rates and the relative contribution of each feature to the discrimination between B and T cells.]

Careful inspection of the collected cell image data showed that even the purified T and B lymphocyte populations were not entirely homogenous. In addition to the contamination of the T cell preparations with small percentages of B cells and the contamination of the B cell preparations with small numbers of T cells, and especially monocytes, there was evidence that subgroups of human B and T cells existed.

In prior studies of mouse thoracic duct lymphocytes, two and possibly four subpopulations of T cells were identified.[10] Thus there was the potential for accurate subclassification of peripheral lymphocytes in health and disease using scanning microphotometry.[7,9] There is already considerable information available linking specific lymphocyte and monocyte abnormalities with known disease states. For example, Sézary syndrome was recently identified as probably a disease of the "helper" T cell.[22] Very subtle changes in the quantity or quality of small subpopulations could be identified by the methods presented here.

The specific advantages of computer-assisted scanning microphotometry are the potential for high reproducibility (as compared to rosetting, fluorescent, and culture techniques,[11]) the ability to make very fine discriminations (e.g., identifying subpopulations), the requirement for only small cell samples, the speed of analysis (immediate on completion of scanning), and the ability to

work with mixed cell populations. Discrimination within a mixed cell population is a particular advantage in that any cell separation method has the inherent risk of depletion of cell subgroups. Additionally, rosette techniques depend upon cell surface antigenic features, which are not always expressed in disease states, leaving some cells classified as "null cells."[23]

An apparent major initial disadvantage of scanning microphotometry is the cost of the equipment and the need for expertise in setting up the system. However, once basic computer algorithms have been established, there is no reason why a center utilizing the method cannot serve as a facility shared by scientists in a variety of disciplines.

Nonetheless, questions arise as to the general biologic significance of the observed differences. Pragmatically, provided that the discriminant capacity is reproducible, such questions do not have to be answered immediately. But prior to any eventual general use of the system the effects of known biologic parameters, e.g., cell cycle phase, immune stimulation, age, and various disease states, must be carefully evaluated.

ACKNOWLEDGMENT

We thank Barbara Soehnlen for skilled technical assistance in preparation of the purified lymphocyte fractions.

REFERENCES

1. Wied GL, Bartels PH, Bahr GF, Oldfield DG: Taxonomic intracellular analytic system (TICAS) for cell identification. Acta Cytol 12:177-210, 1968

2. Bartels PH, Bibbo M, Bahr GF, Taylor J and Wied GL: Cervical cytology: Descriptive statistics for nuclei of normal and atypical cell types. Acta Cytol 17:449-453, 1973

3. Kiefer G, Kiefer R, Moore GW, Salm R, Sandritter W: Nuclear images of cells in different functional states. J Histochem Cytochem 22:569-576, 1974

4. Sandritter W, Kiefer G, Schlucter G, Moore W: Eine cytophotometrische methode zur objektivierung der morphologie von zellkernen. Histochemie 10:341-352, 1967

5. Brenner JF, Gelsema ES, Necheles TF, Neurath PW, Selles WO, Vastola E: Automatic classification of normal and abnormal leukocytes. J Histochem Cytochem 22:697, 1974

6. Bibbo M, Bartels PH, Bahr GF, Taylor J, Wied GL: Computer recognition of cell nuclei from the uterine cervix. Acta Cytol 17:340-350, 1973

7. Olson GB, Anderson RE, Bartels PH: Differentiation of murine thoracic duct lymphocytes into T and B subpopulations by computer-cell scanning techniques. Cell Immunol 13:347-355, 1974

8. Anderson RE, Olson GB, Shonk C, Howarth JL, Wied GL, Bartels PH: Computer analysis of defined populations of lymphocytes irradiated in vitro. I. Evaluation of murine thoracic duct lymphocytes. Acta Cytol 19:126-133, 1975

9. Bartels PH, Jeter WS, Olson GB, Taylor J, Wied GL: Evaluation of correlational information in digitized cell images. J Histochem Cytochem 22:69-79, 1974

10. Bartels PH, Olson GB, Layton JM, Anderson RE, Wied GL: Computer discrimination of T and B lymphocytes. Acta Cytol 19:53-57, 1975

11. Report of a WHO/LARC-sponsored workshop on human B and T cells, London, 15-17 July 1974. Identification, enumeration and isolation of B and T lymphocytes from human peripheral blood. Scand J Immunol 3:521-532, 1974

12. Bøyum A: Separation of leucocytes from blood and bone marrow. Scand J Clin Lab Invest 21 [Suppl 97], 1968

13. Yam LT, Li CY, Crosby WH: Cytochemical identification of monocytes and granulocytes. Am J Clin Pathol 55:283-290, 1971

14. Deitch AD: Cytophotometry of nucleic acids, in Wied GL (ed): Introduction to Quantitative Cytochemistry. New York, Academic, 1966, pp 327-354

15. Bartels PH, Bahr GF, Wied GL. Cell recognition from line scan transition probability profiles. Acta Cytol 13:210–217, 1969

16. Pressman NJ: Markovian analysis of cervical cell images. J Histochem Cytochem 24:138–144, 1976

17. Vidal B de Campos, Schlueter G, Moore GW: Cell nucleus pattern recognition: Influence of staining. Acta Cytol 17:510–521, 1973

18. Ostle B, Mansing R: Statistics in Research (ed 3). Ames, Iowa, Iowa State Univ Pr, 1975

19. Bradley JV: Distribution-free Statistical Tests. Englewood Cliffs, N.J., Prentice Hall, 1968

20. Holm G, Pettersson D, Mellstedt H, Hedfors E, Bloth B: Lymphocyte subpopulations in peripheral blood of healthy persons. Clin Exp Immunol 20:443–457, 1975

21. Greaves MF: Clinical applications of cell surface markers. Prog Hematol 9:255, 1975

22. Broder S, Edelson RL, Lutzner MA, Nelson DL, MacDermott RP, Durm ME, Goldman CK, Meade BD, Waldmann TA: A malignant proliferation of helper T cells. J Clin Invest 58:1297–1306, 1976

23. Greaves MF, Brown G, Rapson NT, Lister TA: Antisera to acute lymphoblastic leukaemia cells. Clin Immunol Immunopathol 4:67, 1975

24. Sherwood EM, Bartels PH, Wied GL: Feature selection in cell image analysis: Use of the ROC curve. Acta Cytol 20:255–261, 1976