A COMPUTER-AIDED SYSTEM FOR STUDYING PLANT CHROMOSOMES

C. Merritt and T. Kasvand
National Research Council Canada

ABSTRACT

The paper describes an interactive system developed to aid biologists in the study of plant chromosomes. The system produces karyotypes of chromosome spreads as well as measurements of individual chromosomes in the spreads.

A 35-mm negative of a chromosome spread is scanned and displayed on a television screen. Each chromosome is straightened and measured interactively; the straightened chromosomes are then displayed in a format chosen by the operator.

The system has been designed for an operator with little experience in computing but whose knowledge of plant chromosomes is required in judging the results of the machine processing. Interaction allows the biologist/operator to point out significant features and to guide the processing when needed; at the same time, the machine provides measurements not easily obtained manually. Hard copy of the final display provides the biologist with the familiar karyotype used in studying chromosomes.

SYSTÈME D'ÉTUDE DE CHROMOSOMES VÉGÉTAUX A L'AIDE D'ORDINATEUR

RÉSUMÉ

Cette communication décrit un système interactif qui a été mis au point pour aider les biologistes à étudier les chromosomes végétaux. Le système produit les caryotypes des étalements de chromosomes ainsi que les mesures de chromosomes individuels dans les étalements.

Un négatif de 35 mm d'un étalement de chromosomes est balayé et affiché sur un écran de télévision. Chaque chromosome est aligné et mesuré interactivement; puis, les chromosomes alignés sont affichés dans un format choisi par l'opérateur.

Le système est destiné à l'usage d'un opérateur à qui manque l'expérience des ordinateurs mais à qui la connaissance des chromosomes végétaux est nécessaire pour évaluer les résultats du traitement automatique. Le traitement interactif permet à l'operateur/biologiste d'indiquer les éléments importants, et de diriger le traitement quand nécessaire; en même temps, l'ordinateur pourvoit des mesures qui sont difficiles à obtenir à la main. Une copie sur papier de l'affichage final pourvoit au biologiste le caryotype familier qui est utilisé pour étudier les chromosomes.

A COMPUTER AIDED SYSTEM FOR STUDYING PLANT CHROMOSOMES

INTRODUCTION

An interactive system has been designed as an aid to biologists in the study of plant chromosomes. Digital image processing techniques are used to facilitate the evaluation of chromosomes spreads of various plant species and hybrids. The system provides both karyotypes of chromosome spreads and measurements of individual chromosomes in the spreads.

Plant chromosome spreads typically contain long chromosomes which are frequently bent and overlapping. Karyotypes are currently prepared manually by cutting out the chromosomes from a photographic print of the spread and arranging them in some order on a new sheet of paper. Several prints of the same spread may be required to extract all the chromosomes when overlapping occurs. In order to keep karyotypes as uniform as possible, spreads are searched for in which the chromosomes are relatively straight. Accurate measurements of the individual chromosomes are difficult to obtain manually.

The interactive system is designed to overcome the problems of manual methods, while retaining the benefits of a human operator. Advantages of the interactive approach are:

1) The operator's knowledge and experience may be used in making decisions concerning the chromosomes. For example, the correct separation of overlapping chromosomes can be indicated by the operator, using his knowledge of the chromosomes in the particular sample spread.

2) Each processing task is performed by the operator or by the machine, according to which is better suited for the task. For example, the operator can quickly point out the approximate location of a chromosome, whereas the machine can easily rearrange the form of a chromosome in order to straighten it.

3) The operator may correct or adjust decisions made by the machine at each stage of the processing, thus preventing propagation of errors throughout the processing.

4) The operator can control the display of the final resulting karyotype and make changes in it as he sees fit.

HARDWARE

The system hardware consists of a MODCOMP II minicomputer with 64K words of memory, two 1.25 megabyte disks and two magnetic tape drives. Input of a 35mm negative transparency of a chromosome spread is accomplished via a flying spot scanner; the digitized image can be stored on tape or disk. For processing, the original picture and resulting images are stored in a Norpak frame buffer and displayed on a colour TV screen. The frame buffer consists of three memory banks, each of which is capable of storing a 640x512 8-bit picture.

Interaction with the operator is via the teletype and a specially built box of pushbuttons and potentiometers. The teletype is used for entering alphanumeric information. Processing is easily controlled by the operator using the pushbuttons; the operator uses the potentiometers to control the Norpak cursor and to indicate points and areas of interest on the display screen.

SOFTWARE

System software consists of two groups of programs, written in FORTRAN: the first group extracts individual chromosomes from the spread, straightening and measuring them; the second group displays the straightened chromosomes as a karyotype. Some of the software has been borrowed from an earlier computer-assisted system for karyotyping human chromosomes[1]; the remainder has been designed particularly for plant chromosome studies and to be used with the available equipment.

The two groups of programs are described, including the actions required by the operator. Before processing commences, the image of the chromosome spread is loaded into one memory bank of the frame buffer from either disk, tape or directly from the scanner (Figure 1).

1. Extraction of Chromosomes

1-1 Threshold the Picture

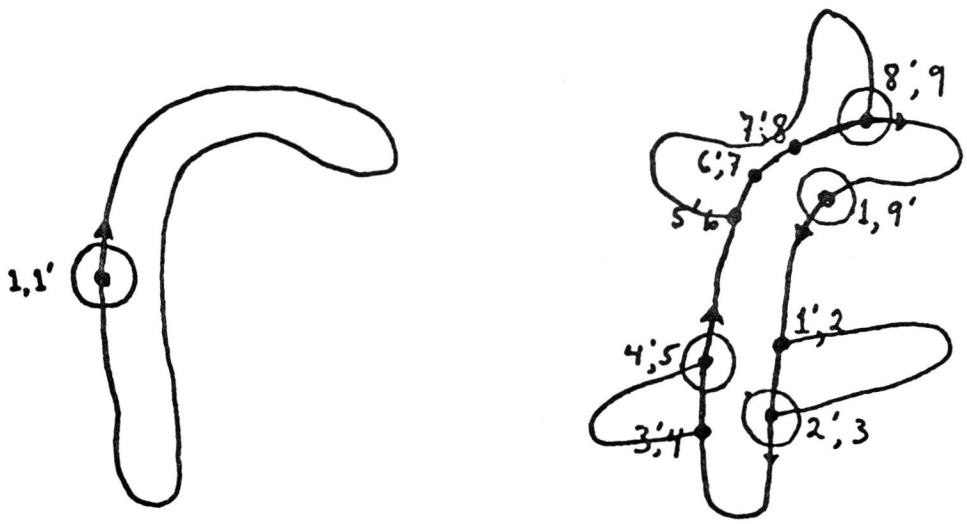
Since the chromosomes are usually very distinct from the background, the picture can be thresholded successfully, simplifying further processing. The operator indicates an area containing a chromosome to be extracted. The binary picture resulting from thresholding this area is loaded into a second memory bank of the frame buffer. A default threshold value is used but it can be adjusted according to the general intensity level of the area. Histograms of the gray levels in typical selected areas have been found to be

bimodally distributed, indicating that suitable thresholds could be selected automatically using well-known techniques.[2]

1-2 Select Points for Contour Follower

The operator selects pairs of points between which segments of the chromosome contour are to be found, and for each pair of points indicates which of two contour following methods is to be used: either the border is to be found by automatic contour following, or the border is to be approximated by a straight line.

For each pair of points between which the contour is to be found automatically, the operator must supply, via the potentiometers, the direction in which the contour follower is to proceed from the starting point and the radius of a search circle (Diagram 1(a)). The approximated border is intended to be used when the contour cannot be found automatically, as when the chromosome overlaps another object (Diagram 1(b)).



(a) (b)

Diagram 1. Points Selected for Contour Follower
(a) The entire border is found automatically from point 1 to point 1'.
(b) The border is approximated by a straight line between point pairs 2,2', 4,4', 6'6', 7,7' and 8,8'.

1-3 Obtain Contour

A list of the coordinates of border points found between each pair of previously selected points represents the chromosome contour. As each border point is added to the list, it is displayed on the TV screen, enabling the operator to check the accuracy of the contour.

When the automatic contour follower is used, the selected starting and ending points are first adjusted to the actual border of the chromosome in the binary image; this is accomplished by searching on either side of each selected point along the circle radius which is perpendicular to the given contour direction until a border point is found. Then, a simple contour follower operating on the binary image[3] is used to locate the border points between the new starting and ending points.

When the contour is to be approximated by a straight line, the coordinates of adjacent points lying on that line between the selected pair of points are calculated.

1-4 Load Contour Into Mask Array

The contour is scaled to fit into a 50x50 mask array. Points in the mask array are marked as being on the contour, outside the contour, or inside the contour. By displaying the mask array, the operator can evalute the performance of the contour follower. A closed contour is necessary for the straightening of the chromosome; if this has not been found or if the contour is inaccurate, the operator may request additional contour segments to be found to close the contour, or obtain a new contour by selecting new pairs of points.

1-5 Locate Backbone

The backbone, whose location is necessary for straightening the chromosome, is found by thinning the chromosome in the mask array to a single line using Hilditch's algorithm.[4] A number of points on this line are selected as the backbone points (Diagram 2). Alternatively, the backbone points can be pointed out by the operator if thinning the mask does not result in a single unbranching line.

1-6 Straighten Chromosome

A cubic spline curve is fitted to the backbone points and, if necessary, extrapolated to the ends of the chromosome by referring to the mask. The length of the chromosome is obtained from this curve.

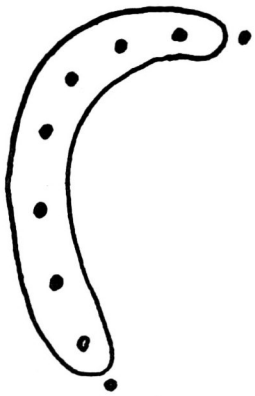


Diagram 2.
Located Backbone Points.

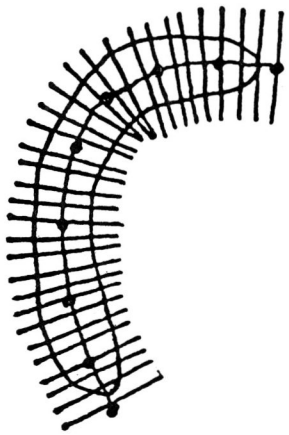


Diagram 3.
Vectors Scanned to
Straighten Chromosome.

The chromosome is straightened by scanning the picture on vectors centered at consecutive points along the backbone curve and perpendicular to the curve (Diagram 3). Two passes are made: on the first pass the scan vectors are displayed in position, allowing the operator to check the results of the straightening process; during the first pass, the area of the straightened chromosome is calculated. Before the second pass is made, the operator is asked to supply an identification number and description for the chromosome; this information plus the calculated length and area measurements are written on magnetic tape as a file label for the straightened chromosome. On the second pass, a tape file for the straightened chromosome is created by writing each scan vector on the tape. Referring to the mask array, only those points within the chromosome on each scan line are actually read; the remaining vector points are set to a minimum value.

2. Display of Karyotype

2-1 Partition the Display

The display area on the TV screen is divided into a number of equal rectangular partitions in which the straightened chromosomes are to be displayed.

Using the potentiometers, the operator chooses the length and width of the partitions and the total size of the display area. By varying these parameters, the partitioning format and number of available partitions can be adjusted to suit the karyotype being displayed. The partition boundaries are displayed and the partitions may be labelled by number for reference by the operator if desired.

2-2 Assign Chromosomes to Partitions

Each chromosome to be displayed is assigned to one of the partitions. Assignment can be done either individually for each chromosome or automatically for the entire set of chromosomes in the spread. In the former case, the operator must give the identification number of each chromosome and the number of the partition it is to be displayed in. In the automatic case, all the chromosomes are assigned to partitions automatically in an order chosen by the operator. The default order is that in which the chromosomes were extracted; ordering may also be done on any of the measurements contained in the file labels, such as chromosome length or area.

2-3 Display Chromosomes in Partitions

Each straightened chromosome is displayed in the partition to which it has been assigned; the partition may optionally be labelled with the identification number, length and area of the chromosome. If the operator is not satisfied with the location or order of the displayed chromosomes, changes are easily made by reassigning individual chromosomes to different partitions or by using a different ordering method for automatic assignment to partitions. The partitioning format may be changed without affecting the assignment of chromosomes to partitions. Figure 2 shows sample karyotypes for the chromosome spread of Figure 1.

OPERATOR CONVENIENCES

The system has been designed for use by an operator whose experience with computers is minimal. Several features have been included in the system to make it easy to operate:

1) Pushbutton control. The processing can be controlled entirely with the pushbutton box. When a processing task has been completed, the operator is presented with a list of possible actions which he may request: i) continue with the next task, ii) make a correction and repeat a certain task, iii) start processing a new chromosome,

iv) control the colour display, v) print a table of statistics gathered for the already straightened chromosomes, vi) display the karyotype. If an error in processing has occurred, only possible courses of action to correct the situation are presented, along with the reason for the error. The pushbutton control requires the operator to have only a general understanding of the processing sequence; where action by the operator is required, the options are presented in a language easily understood.

2) Automatic mode. Although processing consists of a number of individual tasks, automatic mode allows the tasks to be executed one after the other until a checkpoint is reached where the operator's approval to continue processing is required. Checkpoints occur at only a few places: i) after the mask array is displayed, allowing the operator to correct the contour if necessary; ii) before the straightened chromosome is written on tape, allowing the operator to select backbone points manually, if straightening was unsuccessful; iii) when initiating the display of the karyotype; iv) after the karyotype has been displayed, allowing the operator to make changes in it.

3) Negative display. The scanned and stored image is of a negative transparency of the chromosome spread. This image is normally displayed as this negative but features provided with the colour display allow the corresponding positive image to be displayed. (Figure 3). The operator, who is accustomed to looking at positive prints and through the microscope, is able to work with the more familiar image.

CONCLUSIONS

The interactive system presented provides the biologist with a useful tool for studying plant chromosomes. The use of interaction allows the machine and the operator to perform the tasks for which each is best suited. Measurements of individual chromosomes, which are difficult to obtain manually, are provided. The ability to straighten the chromosomes allows some uniformity between karyotypes of different sample spreads. Emphasis has been placed on ease of operation of the system and ease of making changes, both in correcting the processing results when necessary, and in changing the display to obtain suitable karyotypes.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. G. Fedak of the Ottawa Research Station, Agriculture Canada for providing the data.

REFERENCES

1. Kasvand, T., Hamill, P., Bora, K.C., Douglas, G., "Experimental Online Karyotyping at the National Research Council of Canada", Automation of Cytogenetics, Asilomar Workshop, Pacific Grove, California. November 30-December 2, 1975.

2. Rosenfeld, A. and A.C. Kak, Digital Picture Processing, New York: Academic Press, 1976.

3. Duda, R.O. and Hart, P.E., Pattern Classification and Scene Analysis, New York: John Wiley, 1973.

4. Tamura, H., "Further Considerations on Line Thinning Schemes", Paper of Technical Group on Pattern Recognition and Learning, Institute of Electronics and Communication Engineers of Japan (IECEJ), PRL75-66, Dec. 1975 (in Japanese).

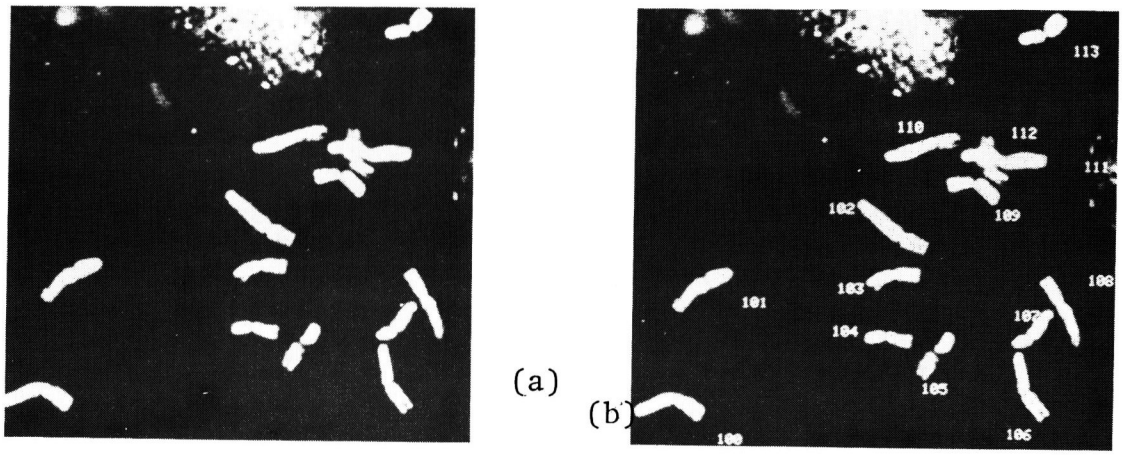


Figure 1.
(a) Display of original digitized negative film.
Chromosome spread of barley x rye
(b) Chromosomes labelled with identification numbers.

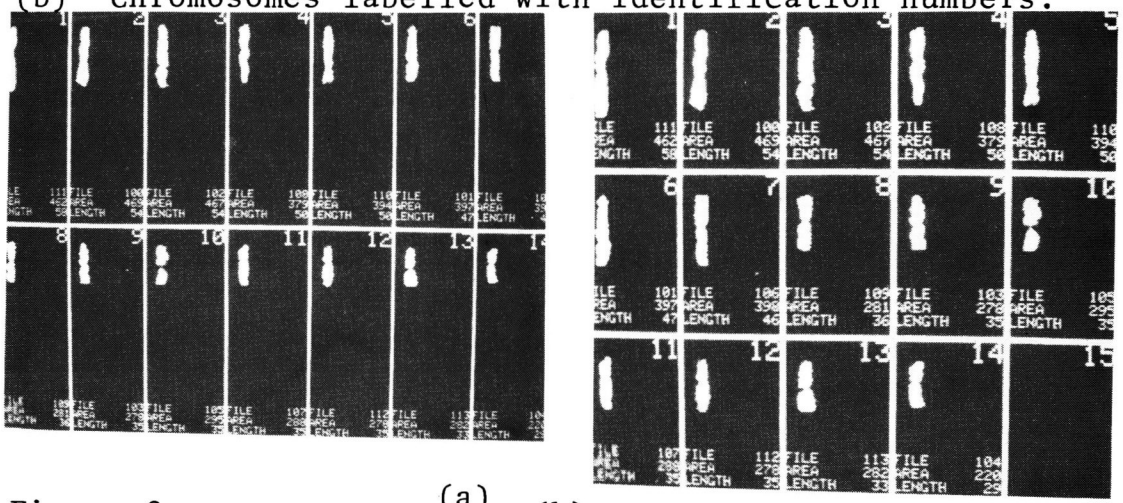


Figure 2.
(a) Karyotype of chromosomes in Figure 1 ordered on length.
(b) Alternate partitioning format for karyotype of 2(a).

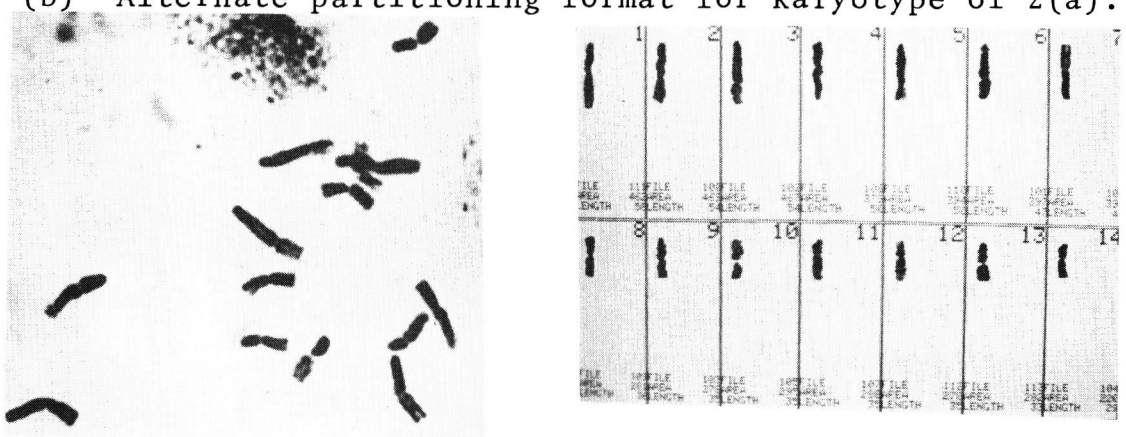


Figure 3 (a) Display of positive for Figure 1(a).
(b) Display of positive for Figure 2(a).