Title
AUTOMATED COUNTING OF CELL CLONES: A PROGRESS REPORT

Author
White, Howard S.

Publication Date
1976-06-01

LBL-4674
UC-48
TID-4500-R64

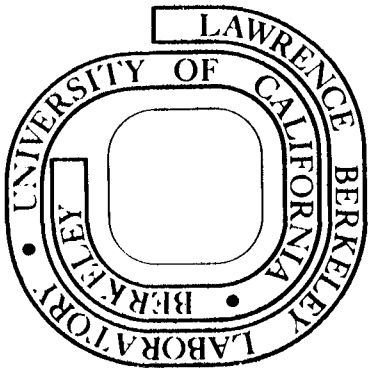
AUTOMATED COUNTING OF CELL CLONES: A PROGRESS REPORT

Howard S. White, Eleanor A. Blakely, and
Tracy C. Yang

DONNER LABORATORY

June 1976

Prepared for the U. S. Energy Research and
Development Administration under Contract W-7405-ENG-48



LBL-4674

DISCLAIMER

LBL 4674

AUTOMATED COUNTING OF CELL CLONES: A PROGRESS REPORT

Howard S. White

Computer Science and Applied Mathematics Department

Lawrence Berkeley Laboratory

and

Eleanor A. Blakely and Tracy C. Yang

Donner Laboratory

University of California

Berkeley, California 94720

June 1976

TABLE OF CONTENTS 22 JUN 76

TABLE OF CONTENTS 22 JUN 76

This report summarizes progress made in adapting the Flying Spot Digitizer, designed for bubble chamber data measurement, for use in enumerating human cell colonies grown in vitro. A description is given of the digitizing hardware and the general pattern recognition programs which are available in the FSD system. The development of the computer program called PETRI is also described. This new software provides the basis for analysis of the digitized information from the FSD.

As with particle track studies from the bubble chamber, experimental data are photographed and the film strip is processed through the FSD for digitizing. Tests of the FSD-PETRI analysis system have been made with film strips of the cell colonies grown on plastic petri dishes. Further work is necessary, but preliminary results presented here demonstrate that the system can be used to enumerate cell colonies with a precision near that of counts made by hand. This precision was made possible by the machine's ability to distinguish overlapping colonies as more than a single large colony. In addition, the FSD-PETRI system provides information which is virtually impossible to obtain by hand; namely, a measurement of the distribution of colony sizes including those of very small colonies. These data potentially will enable the biologist to more clearly understand cellular mechanisms of damage and repair, and will greatly facilitate his ability to obtain measurements of cell survival rapidly and objectively.

Quantative mammalian cell radiobiology traces its origin to the efforts of T. T. Puck and associates (1). This group was responsibile for the introduction of mammalian single-cell techniques *in vitro*. Currently, the method of culturing single cells in a nutrient medium is used as a survival assay for many treatments in laboratories all over the world.

The criterion established is a measure of the reproductive integrity of cells. The cells are usually incubated in petri dishes under controlled conditions of temperature and humidity. If appropriate dilutions are made, the growth and division of the single cells proceeds until a discrete, round clone, whose progeny are derived from a single parent, is visible to the eye. When the clones are of a sufficient size, their growth is terminated and their visibility is enhanced by staining with a dye like methylene blue. Survival determinations are made by handcounting the resultant blue spots left on the bottom of the petri dish.

Even under control conditions (with no known deleterious manipulations), not all of the cells plated onto the petri dishes will produce a clone. The term "plating efficiency" is commonly used to indicate what percentage of control cells actually grow into colonies. It is usualy assumed that the same fraction of cells will be viable from a single population; however, this is not always true as there is some indication that cell survival is dependent on the concentration of cells plated.

Hand counting of control clones is usually not a laborious process if an appropriate concentration of cells was selected. For most established mammalian cell lines, the shape and size of clones from untreated cells are fairly uniform. It becomes a more tedious task however to hand-count clones resulting from cells which have been subjected to a treatment, such as exposure to a radiation dose (see Figure 1). Colony number is reduced with increasing dose of radiation, and the distribution of colony shapes and sizes is widened.

Puck and Marcus (2) originally set 50 cells after 10-20 days as the criterion for colony size to be scored for cell survival determinations. Nias and Fox (3) elegantly demonstrated that the apparent discrepancy between survival data from clone counting and those from growth inhibition cell count studies could be reconciled by scoring a larger minimum clone size. The clone size which represented at least eight cell divisions for HeLa (human carcinoma) cells 10 days post-irradiation corresponded best with the growth extrapolation curve survival data. Survival estimates with a less demanding colony size cut-off were markedly higher.

A few investigators have taken the time to count clones under the microscope, evaluating the cell number present, but this is a very tedious, time-consuming and difficult technique. Most investigators just scan colonies by eye, or with low magnification, and set an arbitrary lower threshold diameter for clone size. Reexaminations of the end-point criterion for survival in cellular

radiobiology emphasize the fact that technical differences exist within the survival literature which must be taken into consideration before meaningful comparisons are made.

The Biomedical Division has begun an extensive program to study the effects of BEVALAC accelerated heavy ions on cultured human cells, in order to explore the feasibility of using these beams for cancer therapy. It is desirable to accurately measure cell survival parameters after doses of accelerated particles in order to know more about the radiation-induced increase in the size distribution of colonies, and to more precisely determine when the colonies stop growing. It is hoped that this information, which is not usually obtainable with conventional techniques, will provide evidence to elucidate cell damage and repair mechanisms after radiation of different qualities.

Experimental data from the BEVALAC runs are generated in several sessions of contiguous days and nights. The backlog of petri dishes to be counted is one of the most severe limitations in present procedures. More data could be accumulated if colony counting were done automatically.

There are several automated systems available to count colonies. One very advanced technique developed in the laboratory of Donald A. Glaser (4) can precisely enumerate and characterize cell colonies. However, its use requires elaborate changes in the entire procedure for handling the cells. Some commercial equipment

is also available to automatically count colonies (5); however these instruments have problems identifying colony clusters, do not resolve small colonies, are subject to artifact and background interference, and are usually relatively slow in operation.

The Flying Spot Digitizer was built at LBL to measure experimental physics data. When recently the demand for physics use diminished, it was possible to explore other applications. One most promising application is to colony counting in petri dishes.

3.1 THE FLYING SPOT DIGITIZER

The Flying Spot Digitizer (FSD) was designed specifically for bubble chamber data measurement. The bubble chamber acts both as a target for particles from the accelerator, and as a detector of the products from the nuclear reactions which take place within the chamber. The bubble chamber is usually filled with liquid hydrogen, so that "events" take place on protons not bound into more complex nucleii. The liquid is heated, but compressed so that boiling does not take place. When pressure is released momentarily, the liquid forms tracks of bubbles along the trajectories of the charged particles. A magnetic field is maintained in the chamber to cause the particles to move in helical orbits. The tracks can then be photographed, showing a mapping of the orbits in the chamber space. Since photographs are taken from three or more camera locations, it is possible to reconsturct the trajectories and determine the momentum vectors of the charged particles participating in the events. This reconstruction depends upon very precise measurement of the track locations with reference to fiducial marks.

A very large number of photographs is produced by bubble chamber experiments. Normally three views of each event are photographed in each Bevatron pulse. A single experiment may continue taking film for several weeks, so that a million three-view photographs are obtained, containing about 300,000 useful events. One such experiment, or else a combination of

experiments yielding about the same number of events to be measured is performed each year. The FSD was designed to meet these measurement needs economically. It has now measured about four million bubble chamber events since beginning operation in 1963.

The design goals for the FSD were precision, speed, and economy of measurement, as well as versatility of application. Since bubble chambers produce film in a variety of formats, the FSD was built with a platen area 150 mm long and 60 mm wide, so as to accomodate film from all chambers. The accuracy requirements were set by film base stability and by the limits imposed by the photographing optics. These requirements suggested that FSD accuracy in the neighborhood of 2 microns would cause measurement errors to be small in comparison to the other factors. Speed of operation and accuracy were made possible by using a Hough-Powell Device (HPD) to scan the film, and by pushing the then state of the art for digital circuitry. Economy was achieved by totally automatic operation under control of the computer which reduced the measurements.

The device as implemented at LBL is able to measure from 100 to 300 film images per hour, depending upon format and scanning density. Images may yield from 50,000 to 1,000,000 position measurements each; those of petri dishes typically yield 100,000. Image features in the size range 15 to 350 microns can be measured with accuracy from 3 to 5 microns, respectively, by the FSD as implemented. This accuracy refers to the reproducibility of

measurements made by single intersections of the flying spot with the image features.

The FSD employes a Hough-Powell Device (6) as its digitizer. A schematic drawing of the FSD optical system is shown in Figure 2.

The device at Berkeley is one of three similar machines that were built concurrently in 1962-1964. Others were built at Brookhaven National Laboratory (BNL) and the European Organization for Nuclear Research (CERN). Although detailed implementation was different for each of these, they shared a common genesis. The basic concepts were worked out by an informal team of persons from the three laboratories, which included Paul Hough from BNL, Tor Lindjarde and Brian Powell from CERN, and Jack Franck, Jerome Russell and Howard White from LBL. Many valuable suggestions were given by physicists Ralph Schutt and Allen Thorndyke at BNL, Lew Kowarski at CERN, and Edwin McMillan, Wilson Powell and Arthur Rosenfeld at LBL, and from others too numerous to name individually.

The HPD contains a fixed slit and a moving slit, whose intersection defines an aperture which moves along the fixed slit. This aperture is illuminated by a mercury-vapor arc lamp, producing a very bright spot that can be imaged upon the film to be scanned. Simultaneously, another part of the light from the spot is imaged upon a reference grating by which the location of the spot can be measured accurately. This reference grating is labeled "picket

fence" in the figure. The spot aperture is about 16 microns in dimension, and is imaged upon the film and reference grating by means of a one-to-one optical system. The lens apertures cause the spot images to be about 20 microns in dimension due to diffraction limits.

The spot moves along the image of the fixed slit. Two optical systems cause this image to be either across or parallel to the long dimension of the film. These are shown in the figure with labels "normal sweep" and "abnormal sweep", respectively. A mechanical shutter allows light to pass through one or the other of these systems.

A mechanical stage carries the film platen, causing it to move with respect to the fixed slit image. By moving parallel to the long dimension of the film, and selecting the normal mode optics, a raster scan of the film is generated. This raster is intended to accurately measure the position of line-like objects oriented along the length of the film. Similarly, by selecting the abnormal mode optics, and moving the stage orthogonally to the long dimension of the film, the abnormal or orthogonal mode scan raster is generated. This raster is intended to accurately measure the position of line-like objects oriented across the film.

The position of the mechanical stage is digitized with a least count accuracy of 2.5 microns. This position is measured with Ferranti linear position resolvers which use moire-fringe patterns

to detect stage motion. A scaler counts fringes bidirectionally, and thus always reflects the stage position.

Each raster scan line is produced by the traversal of the fixed slit by one of the moving slits. Eight moving slits are arranged symmetrically on the disc. These are semi-circular approximations to an involute, so that the intersection moves with uniform velocity along the fixed slit. The disc is driven by a 3600 rpm motor, yielding 480 scan lines per second.

The position of the spot is known to a precision of 1 micron by comparison to the reference grating. A colimated beam of light from the mercury vapor arc lamp is imaged through the slit aperture onto each of the three optical systems. Matched Goetz Artar lenses are used to image the spot upon the film platen and reference grating. These lenses are designed for one-to-one copying, and have extremely low off-axis distortion. Since they are carefully matched, the differential distortion between the lens in the reference grating optical path and either of those in the film measurement path is small enough that it does not need to be included in the calibration equations. The spot image on the reference grating is also approximately 20 microns in dimension. The reference grating is composed of alternate opaque and transparent bands, each band being 16 microns wide. The spot sweeps across these bands, producing a modulated intensity profile resembling a sine curve. The spot position is therefore measured by counting peaks from the beginning of the sweep to give a

fundamental reference to each 32 micron reference position, and by time interpolation within the interval to yield a least count of 1 micron. Because the disc is rotating with considerable inertia, the stability of the time interpolation over the 32 micron interval is extremely good.

The spot moves along the slit at a velocity of 32 microns per microsecond. The basic reference grating cycle has a frequency of 1 MHz. A phase-locked oscillator is operated at the fifth harmonic of this frequency, and provides the standard for spot position, causing a scaler to read directly in least counts whose size is very close to one micron.

The HPD was chosen over a cathode-ray tube (CRT) spot generator for several reasons. Although the choice was made in terms of the much less well developed state of CRT imaging techniques that existed in 1962, many of the reasons are still applicable and would cause a similar decision today. The mechanical spot generator offers long term calibration stability in contrast to the CRT, which is subject to distortions that change as a function of recent history of use. It was not possible in 1962 to generate a spot of 16 micron dimension with a CRT, and to use a larger spot would have downgraded the signal-to-noise ratio for bubble chamber data. Area scanning requires that the entire scanning area be covered with a raster. The speed of measurement is therefore directly dependent upon how fast the spot can move. Because CRT phosphors have a slowly decaying component in their

light emission as a function of time, to move the spot as rapidly with a CRT as is done with the HPD would cause a very long tail of luminescent phosphor to follow the spot, with grave loss of contrast. The speed of the spot motion also requires a very intense spot, since the spot moves through its diameter in about 500 nanoseconds, and therefore the number of photons is so small as to be the principal component in the signal to noise ratio. These considerations of raster stability, spot speed and intensity are even now valid reasons for using the HPD to generate the flying spot.

The output from the HPD is intensity as a function of spot position. The intensity of the light received by a photomultiplier viewing the platen has been modulated by the image on the film. This intensity is expressed as an analog voltage varying in time. The spot position scaler is read out electronically to give the coordinate of the spot.

The center of area of the intensity profile is used to define the location of the feature being digitized. A schematic representation of the intensity trace is defined in Figure 3. Because of the rapid spot motion, the intensity profile has a substantial noise component caused by statistical noise. The center of area is least affected by this random noise of all digitizing techniques.

Digitization is accomplished by recording into memory the

scaler value corresponding to the center of area of features encountered by the spot. The spot coordinate is labeled "w" in the normal mode, "w'" in the orthogonal mode. A digitizing threshold is established as an analog voltage. Whenever the spot voltage exceeds the threshold voltage, the feature is deemed to have begun. It ends when the voltages return to the inverse relationship.

Provision is made in the FSD for digitization either of images having opaque features on a clear background, or else clear features on an opaque background. The image is best digitized when clear areas have transmissivity of 80% or more, and opaque areas have transmissivity less than 20%. Because of the nature of the petri dish images, which have small round features, digitization of these as clear images on opaque background is preferred to eliminate problems of dust particles adhering to the film.

Motion of the spot across the image one time is termed a scan line. The raster scan is composed of a number of consecutive scan lines, each scanning a different area of the picture due to the stage motion between scan lines. The digitized results of a scan line include the X or Y coordinate of the stage, depending on whether the normal or orthogonal mode is selected, followed by the w or w' coordinates of all centers of area for features encountered. Because closely spaced features may cause consecutive w's to occur too rapidly for transmission to the computer, a buffer memory collects digitizings from one scan line, and transmits them to the computer during the measurement of the following scan line.

Two buffers alternate to handle consecutive scan lines. These buffers can contain as many as 120 w-values. This limit was chosen as being several times the maximum number of features encountered in bubble chamber images. It is equally satisfactory for petri dish images. Electronic circuitry exists in the FSD which allows digitizing of feature elements at any spacing, so long as no more than 120 are encountered in one scan line.

Digitizings of consecutive scan lines comprise the data stream of one sweep of the stage. The first scan line transmits the scaler position corresponding to the fixed stage coordinate, followed by that of the moving coordinate and the feature centers. The data stream next has moving stage coordinate and feature centers of the second scan line. Other scan lines follow. Depending on image size and scan line separation, anywhere from a few hundred to several thousand scan lines comprise a sweep, which is the primary unit of data sent to the computer. The attached computer is required to process the data stream in real time, i. e., within the time framework established by the FSD in generating it. This is necessary because the FSD stage has large inertia, and its servo requires several hundred milliseconds to bring it up to speed or to stop it. To interrupt measurement within a sweep would require retracing part of the sweep and restarting, a prohibitively uneconomic process.

The FSD has a film transport system which positions the desired image on the film platen. Figure 4 shows the Berkeley

FSD with its film transport and film platen visible in front of the operators. Provision is made for rolls of film holding as much as 1000 feet of film, so that time lost in film changing is minimized. The rolls may have 35 or 70 mm film, either perforated or unperforated. Special platens have been used for film of intermediate sizes. When the film is perforated, the positioning is by means of the sproket holes. The FSD and programs assume that the images bear a fixed relationship to the sproket holes, and that they are uniformly spaced. Many bubble chamber formats include a strip of line markings across the film which encode roll and frame number. These allow film to be automatically positioned under control of the computer program. When this data box information is not included in the film format, it is necessary for the operator to position the roll manually before measurement of the first frame. Subsequent frames meeting the uniform spacing criteria can then be positioned automatically. Unperforated film requires special markings to be exposed along the edge of the film to allow automatic centering of the images.

3.2 THE IBM 7094 COMPUTER

The FSD is attached online to an IBM 7094-II computer. This is a large scale general purpose computer built with discrete transistor technology. It was upgraded from a 7090 computer during the early days of the FSD. Its speed and capacity are necessary to maintain the real-time relationship with the FSD.

Several new programs were written to allow use of the existing FSD hardware to explore a variety of pattern recognition applications. The programs which had been used for bubble chamber work were specially designed for the film formats and information content unique to bubble chamber data. A set programs was needed that would control the digitization and handle the measurements of a wide variety of possible film images.

The goal of these programs was to meet the real-time demands of the hardware, to allow the power of a large general purpose computer to address the pattern recognition problem, and to implement procedures which were sufficiently economical as to be directly suitable for use in actual data processing applications.

4.1 THE MEASUREMENT CONTROL PROGRAM

A program (701) was written for the IBM 7094 computer that would control the measurement and transmit the digitizings in most compact form to tape for subsequent analysis. It is in this program that the real-time constraints of the FSD hardware must be met. These limit the amount of image processing that can be done concurrently with the measurement. Hardware sensed error conditions are dealt with, usually by forcing some hardware reinitialization, and repeating the measurement of the image in which the error was detected. The data stream from the hardware is analysed for validity while the digitizings of interest are being abstracted for output to tape.

The measurement control program is supplied with a deck of control cards which specify the format and descriptors of the film images to be digitized. Identifying labels as roll, frame, and image serial number, together with the description of the image locations on the film are contained in the control deck. This deck also contains parameters describing the number and mode of the digitizing sweeps, the digitizing threshold, and the area of the image to be transmitted to tape for each sweep.

Under control of this program on the IBM 7094, the FSD is operated to measure the film. Once the proper roll of film has been placed upon the FSD and has been moved to the correct initial position, further operation is under control of the computer

program. Sweeps of the first image are begun, and as each sweep is made, the digitizings of the appropriate area are transmitted to tape. The data stream is checked for scan line integrity, and error flags set by the hardware are monitored. Should any error of significance be detected, the digitizings are discarded without being written to tape, and the sweep is repeated. A summary of operations and of detected errors is gathered for quality control purposes at the end of the run.

The output tape contains digitizings of the specified image area. Each sweep yields a block of physical records containing digitizings from consecutive scan lines. The 18-bit words characterizing XY and W coordinates are stored in order of their occurance within the data stream, except that they are packed two per 36-bit machine word. One XY and all W values within the selected region are written for each scan line. A final record for each sweep contains a summary of all resets and other error corrections taken by the program. Each image yields a contiguous group of physical records on the output tape. About 6,000,000 coordinates can be written on one reel of tape. The number of images written per reel depends upon the number of digitizings contained within each image.

4.2 THE CALIBRATION AND MAINTENANCE PROGRAMS

Three calibration and maintenance programs also operate on the IBM 7094 computer. These programs were developed for use in bubble chamber measurement control, and relate more to the FSD hardware than to the particular kind of data being digitized.

One program (434) uses the two measurement modes to calibrate the optics of the HPD. It does this by finding and measuring a field of fiducials distributed throughout the platen area to be digitized. Locations for each fiducial are calculatd from both orthogonal and normal modes. The coefficients relating the micron coordinate system, common to both modes, to the two mode-dependent raw coordinate systems are then adjusted by a least-squares procedure. The program repeats this cycle of measurement followed by calculation until all fiducial measurements yield the same values in both modes within specified tolerances. The new set of coefficients is then punched out on cards for inclusion in the control deck of program 701. This procedure is done every week when the FSD is running continuously. It only rarely detects a significant change in the calibration coefficients, but is used to certify the accuracy of the FSD digitizations.

One of the maintenance programs (431) is run just before beginning a measurement run. It is used to check the FSD data stream for integrity. Even very low frequency intermittant errors of digital transmission can be found by the program. The second

maintenance program (411) is used to check the analog track center circuits and the interpolation oscillators, which only infrequently require adjustment. The program is used periodically as a check of the adjustment, and as a very necessary analysis tool while the adjustments are being performed.

4.3 THE DATA SEQUENCING PROGRAM

The program CFIN gathers into contiguous words all digitizings from each local image area. This program operates on the CDC 7600, and is written in FORTRAN language. It reads the tape produced by program 701, in which data are ordered according to the natural progression of the FSD in each sweep. For pattern recognition purposes, it is desirable that digitizings for each local area being considered be gathered together, regardless of which sweep was the source. Program CFIN achieves this result by storing the data in packed form in the large core memory (LCM) of the CDC 7600. Provision is made to overflow onto disc if the number of digitizings in one image exceeds the capacity of the LCM. This limit is about 1,000,000 digitizings, and is not reached in PETRI data.

The program divides the image area into sectors and zones, and collects digitizings of all sweeps into these groupings. Because the typical image may yield 100,000 digitizings, it is impossible for the pattern recognition program to store all of the picture image at one time. Since the FSD produces digitizings of one picture region in at least two sweeps (normal and orthogonal), the data stream must be reordered to gather all digitizings of one region together.

The larger division unit of the image is the sector. A retangular array of sectors is placed upon the picture, with sector

size being chosen so that a few thousand digitizings will lie within the typical sector. This choice is predetermined for a given class of input data. Because the sector divisions are laid out without reference to the specific information content of an image, it is expected that image elements will sometimes be divided into two or more sectors. The intent is to have the sector size be several times the picture element size. For PETRI use, the image is divided into 96 sectors, each about 3 mm square on the image. These form an array 12 by 8 in the X and Y coordinates, respectively.

The smaller unit is the zone, which normally is a square with 512 micron side. Digitizings are represented by 20-bit entries containing two 9-bit fields of X and Y coordinates and a 2-bit flag field which identifies the sweep mode of the digitizing. Because it is necessary to use 18 bits to describe a complete X coordinate in the image, the digitizings within the zone are expressed in a coordinate system whose origin is at the corner of the zone. It is therefore possible to pack three digitizings into one 60-bit word of the CDC 7600. Two other words per zone give the zone identity and the origin of the zone with respect to the image origin.

The CFIN output tape contains the digitizings organized into sectors and zones. Each sector is a physical record whose length is determined by the density of digitizings in that picture region. The median value for PETRI data is about 500 digitizings per sector. The sector record contains one logical record of sector

information, and one logical record for each non-empty zone within the sector. The entire sector record is omitted if the sector is empty of digitizings. The sector header record contains the identity of the sector, its reference location in the image coordinates, and an index to the variable length zone records. The zone records contain the digitizings in the packed form previously described. The physical records on tape contain record serial number, word count and a final checkword to verify the accurate transmission of data.

4.4 THE INTERACTIVE DISPLAY PROGRAM

The program PRISM operates on the CDC 6600 computer to interactively display the information of a CFIN tape. A Tektronix 4012 Graphical Display terminal is used to interact with the computer. The standard LBL interactive system SESAME (7) provides the host environment for PRISM. A standard graphics support package GRAFPAC (8) is used to operate the display.

PRISM conducts a dialog with the terminal operator that allows him to move from image to image, to set the display parameters, and to display desired images. The display may be made at any scale and center, and can show points from selected digitizing modes. Sector and zone boundaries can be superimposed under control of the operator. Since the 4012 is a storage screen device, no computer cost is expended in refreshing the display as would be the case for terminals which do not have this property.

The PRISM program serves two purposes: it allows the digitized data to be studied in detail as a graphic presentation, and it serves as the host executive for the development of pattern recognition programs for specific applications. Because most operations are more economically performed on the CDC 7600, the application program, once developed, is then transferred to the larger computer for production runs. But PRISM still remains as a very useful tool for studying problem images and testing new improvements.

In July, 1975, two petri dishes were selected as being typical of those for which automatic colony counting techniques were to be developed. These were 60 mm dishes, having about 100 countable colonies well distributed throughout the dish. One dish was selected from a control set, and the other from a set that had been exposed to a 500 rad. dose of X-rays. This dose was selected because it lies below the 1% survival level for cells irradiated with X-rays on glass, and results in the broad distribution of colony sizes we see after high doses. The dishes were photographed using a fine-grained, high contrast emulsion. Illumination was from behind the dishes.

Figures 5 and 6 show photographs of the control and irradiated dishes, respectively. It may be seen that in both control and irradiated pictures, the colonies are generally round, solid, and are relatively smooth-edged. Their opacity is good, so that the contrast to the surrounding background is high. The background field is relatively clear of unwanted marks, or optical noise, even in the irradiated dish. The photographs were made with 35 mm film, and the images are therefore almost exactly a two-times demagnification of the original dishes. The diameter of the colonies is approximately 250 microns on the control dish image. There are very few closely-spaced colonies in either dish.

One considerable problem in these photographs is the heavy rind-like ring, caused by the reflections on the vertical walls of the petri dishes. Since the colonies extend into the rind, they

cannot be distinguished from it on the basis of their location in the dish. The rind is very broad in the photograph, and its edges are uneven. There are multiple lines associated with it. Because the major band of the rind is very much larger than the FSD's maximum digitizing size of 400 microns, this part is drastically suppressed in the digitizings, and the bulk of the points come from the uneven edges of the rind. This causes the location of the dish edge to be quite uncertain, and makes counting in the vicinity of the edge to be inaccurate. Several attempts were made for the program to locate the edges using the uneven swath of digitizings, but confusion between colonies and edges always gave uncertain discrimination near the line of division. It was decided to solve this problem by changing the dish shape, and no further effort was taken in the program to remedy this problem.

The colony counting program PETRI was developed on the basis of the two sample dishes just described. It was recognized that the sample data base was small, but it was felt that the development of a relatively simple program to count this limited case would be useful in pointing the way toward improvements in the technique not only in the software, but in the preparation of the dishes and in the photography as well. And indeed such has been the case.

PETRI accepts data in the format of the CFIN output tape, and performs two principal operations of gathering the digitizings into clusters, and identifying the clusters. It was developed within the CDC 6600 host program PRISM, and then transferred to the CDC 7600 computer for operation on larger quantities of data than a few frames. The flow of data through the entire system is shown in Figure 7.

Due to the center-of-area digitization by the FSD, colonies are percieved in skeletonized form. This is illustrated in Figure 8 for an isolated colony and for two colonies lying close to each other. The size of the colonies on the film image, about 250 microns, allows about eight sweeps of each mode to intersect the colony when the scan line spacing is 30 microns. These give rise to an equal number of digitizings, to form a characteristic "cross" in the graphical display. Usually a few additional digitizings are produced from intersections of the scan line with irregular edges or other imperfections in the colonies. Therefore 20 digitizings

are typically received for colonies of average size.

Digitizings of individual image feature elements are gathered into separate lists by a process called "clustering". The clustering algorithm used is simple in order to save computer time and memory space. Its object is to gather into a list all digitizings lying close to each other, but separated by some gap dimension from all other digitizings in the image. The algorithm begins in the first zone of the first sector. The first point initializes (opens) the first cluster. A cluster region is defined that extends a distance GAPX in both X-coordinate directions, and a distance GAPY in both Y-coordinate directions from the point. The next point is compared to existing open cluster lists, and either added if it lies within the region defined for an open cluster, or else it is used to open a new cluster defining a new region. When a point is added to an existing cluster, the region defined for that cluster is extended, so that a border of size determined by GAPX and GAPY is always maintained. When the addition of a new point to an existing cluster causes two cluster regions to overlap, the clusters are merged, and their lists are combined. A single retangular region is defined with a gap of the same size as before surrounding it. The search is continued for all points contained in the zone. At the completion of each zone, all open clusters are examined. Those which do not extend into as yet unsearched zones are closed to further additions, and are written to an intermediate file. Only active clusters reside in

core memory during the clustering process. The process is then continued to the next zone, and to the next sector. When all sectors of the image have been processed, the remaining open clusters are closed and written to the cluster file, which is then rewound to prepare for the next phase of the PETRI procedure.

Each cluster in turn is analyzed to determine its identity. Only three identities are presently recognized: single colonies, double colonies, and "other". It was felt that it is better to develop and operate this simplistic model than to delay its operation while developing more sophisticated models while not having yet sampled enough data to gain real insight into the problems which other data might present.

Recognition of the number of colonies within a cluster is based upon the nature of the skeletons formed by the digitizings. A single colony is cross-shaped, with the normal mode digitizings comprising the vertical bar, and the orthogonal mode digitizings comprising the horizontal bar. In the graphical display, the X-coordinate is vertical, directed downward; the Y-coordinate is horizontal, directed to the right. The dispersion about the mean for the Y-coordinate of the normal mode points is small for skeletons having a single vertical bar. Similarly, the dispersion about the mean for the X-coordinate of orthogonal mode points is small for skeletons having a single horizontal bar. Clusters containing only one colony are therefore characterized by small dispersion in both dimensions. Clusters with two colonies will

have large dispersion in at least one or the other dimension. The classification "other" is applied to very large clusters having larger dispersion than expected for double colonies. These primarily result from portions of the rind.

A size is associated with each cluster which is the mean of the lengths of the skeleton bars in the two dimensions. This is simplistic for clusters having two colonies, but the frequency of such features in the sample dishes was sufficiently small that this was unimportant for these dishes. It is one of several details which need attention before the program is ready for experimental data.

Those experienced in counting of colonies by hand assert that size is a sufficient criterion for determining which colonies are to be counted. The program is based upon this assumption. All clusters are tested against a size threshold, and those which exceed this are counted once if single, and twice if double. For convenience, ten thresholds are used by the program, so that ten different counts are obtained as a function of size.

The program PETRI yields as output a listing of all clusters in the image whose dimension is greater than some small threshold, and gives for these their identity number, coordinate location, class, and size. For each image the total count previously described is listed, along with identifying numbers which relate the counts to the film images, and thereby to the actual dish

identities.

During the time in which the PETRI program was being developed, several modifications in technique were made with the intent of improving the contrast and measurability of the film images. Since substantial numbers of photographs are expected to be measured, attention was given to making the task of photographing the dishes one which would be as routine as possible. A Flight Research camera, capable of automatic operation with long rolls of 35 mm film, was adapted to this use. A jig was designed which will allow rapid positioning of the dishes so that the images will always bear the same relationship to the frame index perforations. A variety of emulsions and exposures was tested on the FSD to find optimal settings.

The problem of the rind was obviated, at least in the next intended groups of dishes, by developing a technique of center plating. In this, the tissue cells are innoculated into only the central region of a 100 mm petri dish. Instead of swirling the dish to distribute the sample, the innoculating pipette is used to meter the sample throughout the central 60 mm portion of the culture medium. The procedure generally works well, and can produce relatively well distributed colonies in the dishes. It does have the drawbacks of using larger and thus more expensive dishes, with more nutrient medium and more incubator space. The plating technique on the larger dishes is also more tedious.

A better solution to the rind problem is probably found in the design of a new shape for the petri dish. By sloping the dish

bottom downward from the outside wall for several millimeters, so that only the central portion of the dish is flat, it is probably possible to prevent adherance of the colonies close to the wall so that colonies and rind will occupy mutually exclusive regions. Such a dish has been designed by Jack Gunn and George Gates of the Mechanical Engineering Department at LBL, but has not yet been tested.

An improvement was made to the staining technique. To get maximum contrast between colony and background, the dishes were rinsed in tap water after staining. This procedure removed the somewhat mottled appearance found in earlier dishes, and did improve contrast. It has the unwanted property of washing away the central regions of some colonies, leaving empty circles or "donuts" in the film images. An adjustment was made to PETRI to reduce its propensity for identifying these features as double clusters.

Film of dishes especially prepared for FSD measurement became available in February, 1976. Two tissue strains were used: 49 dishes contained T1 strain human kidney tissue, and 41 dishes contained V79 strain Chinese hamster lung tissue. The dishes were approximately equally divided between control and irradiated samples, with 25 control dishes for each strain, and 24 T1 strain and 16 V79 strain irradiated dishes. The radiation dose was 500 rad. in all cases. The dishes were prepared using the center-plating and tap water rinsing techniques previously described. The first dish from the T1 control series is shown in Figure 9, and the first dish from the T1 irradiated series is shown in Figure 10. We identify these dishes as T1C1 and T1X1, respectively.

The 49 dishes of T1 culture were digitized in two separate runs having substantially different analog thresholds. The runs were made at different thresholds to demonstrate that counting techniques do not critically depend upon threshold nor photographic exposure. It is convenient to identify the runs by the name of their data tapes. CFIN25 digitized the T1 images at high threshold, resulting in smaller apparent colony diameters and the supression of faint or small image features. CFIN27 digitized the same images independently of the previous run, at a low threshold, so that small and faint features were included, and colony diameters appeared larger. Similar high and low threshold runs were made on the V79 images. These were labeled CFIN26 and

CFIN28, respectively. A final digitizing run was made at the same time, with the low threshold used for CFIN27 and CFIN28, processing only the first image of each of the four series of dishes. This run is labeled CFIN29. It allows comparison of two independent measurements having the same instrumental settings.

Displays of these digitizings are contained in several figures which follow. Figure 11 shows CFIN29 digitizings of the first T1 control dish (T1C1). This display was made by the program PRISM, using the Tektronix 4012 Graphical Display terminal. A part of the same image, but digitized in the CFIN25 run at high threshold, is shown in Figure 12. For reasons that are not entirely clear, the colony images on this film are not uniformly opaque, but show considerable amount of structure. The cross-shaped digitizings which had characterized previous petri dish images are not as dominant as expected. The enlarged display gives indication of the amount of structure detail digitized in the pictures. The program still gives reasonably satisfactory performance, and was not changed since future images are expected to return to the more disc-like shapes which produce the idealized skeletons.

Digital displays of the first irradiated dish of the T1 series (T1X1) are shown in Figures 13 and 14. Again, the full dish display is from the low threshold run CFIN29, and the enlarged portion from CFIN25.

A comparison of the diameters measured at the same threshold

setting for single colonies in the upper half of T1X1 was made, and is plotted in Figure 15. Colonies were matched by their coordinate locations in the two measurements, and sizes from the CFIN27 and CFIN29 tapes were compared. The differences in diameter between colonies of CFIN27 and the mean of the two runs were calculated. The standard deviation of these differences is 10.2 microns. When one recalls that the scan line spacing is 30 microns, and that the scan rasters have an arbitrary shift in zero causing the lines to fall randomly in the two digitizations, one would expect this value to be near 15 microns.

Comparison of diameters measured at different thresholds includes not only instrumental scatter but also intrinsic differences in the opacity and structure of the colonies. This is most pronounced in the larger colonies of the control dishes. Such a comparison is shown in Figure 16, with data from CFIN25 and CFIN29 for dish T1C1. The most obvious feature is the shift of approximately 100 microns in colony diameter, which results from the difference in digitizing threshold. There is more pronounced scatter in the distribution, caused by different opacity gradients in different colonies. Some colonies in the photograph are grey while others are black. The standard deviation of CFIN25 diameters from the adjusted mean for individual colonies is 46 microns.

A study of measured colony diameters in ten dish samples was made. Figure 17 shows results for the first ten control and the first ten irradiated dishes of the CFIN25 digitization. The

distribution of each sample into the various size groupings is shown as percent of the total sample. As is expected, the control dishes show a sharp peaking of size in the neighborhood of 300 microns, while the irradiated dishes contain colonies which become much more frequent with decreasing size. Figure 18 gives the same distribution for the CFIN27 digitization. Here the control dishes have colony sizes clustering around 400 microns. This 100 micron difference in apparent diameter is due to the different digitizing threshold in the two runs. The same comparisons were made for dishes of the V79 cultures, and are shown in Figures 19 and 20. Even though these colonies have a different edge structure, the difference between diameters of controls is not far from 100 microns.

The dishes of the T1 culture were counted twice by hand, as well as by the automated technique in the runs CFIN25 and CFIN27. The hand counts were performed independently by two persons who have had much experience counting dishes. Agreement was made before counting as to the minimum size for inclusion in the counts.

The machine counts were functions of the minimum size threshold. A constant 100 microns was taken as representing the difference in size between the two runs, based upon the difference in size observed in the colonies of T1X1 (Figure 15), and upon the difference between peaks in the control colony size distributions (Figures 17 and 18). The machine counts were therefore determined by the choce of only one threshold parameter. This was chosen so

that the number of colonies counted in the control dishes of the T1 sample, averaged between the CFIN25 and CFIN27 runs, equaled the average of the hand counts for the same dishes. The final value was then rounded to the nearest 10 microns. For CFIN25, the minimum size counted was 140 microns, whereas for CFIN27 it was 240 microns.

A comparison of the counts is shown in Table I.

Table I

COMPARISON OF T1 COUNTS

Control (24 dishes)		
Hand count (EB)	152.5	clones/dish
Hand count (TY)	153.8	"
CFIN25 (High thresh, 140 mu)	151.8	"
CFIN27 (Low thresh, 240 mu)	156.5	"
X-ray (24 dishes)		
Hand count (EB)	103.0	clones/dish
Hand count (TY)	92.9	"
CFIN25 (High thresh, 140 mu)	96.2	"
CFIN27 (Low thresh, 240 mu)	93.6	"

These entries reflect the average of the counted dishes of each series. It may be noted that the count of irradiated clusters made

by machine lies quite close to that obtained by hand.

So that the variation in counts from dish to dish can be studied, the counts are plotted in Figure 21 as a function of dish number. Four counts are plotted for each dish, reflecting the two hand and two machine values. It may be seen that the hand counts agree quite well with each other, having a calculated average deviation for the control dishes of 3, and for the irradiated dishes of 8. The machine counts also give good internal agreement, suggesting that the effects of digitizing threshold upon colony edges are well compensated by the 100 micron threshold difference. However, the machine and hand counts agree less well than do either series alone. The average deviations in count calculated for the two series "EB" and CFIN27 are 9 and 25 for control and irradiated dishes, respectively.

The differences between hand and machine counts almost certainly reflect more subtle discriminations than size alone on the part of those making the hand counts. The most significant differences between counts are in dishes T1X8 and T1X9, where the CFIN27 count exceeds EB by about 15%, and in dishes T1X14 and T1X15, where the machine count is about 50% smaller than EB. A detailed study of these comparisons has not yet been made. To do so will require mapping the colonies meeting the hand counting criteria, and studying the machine measurements on a colony by colony basis. No obvious features distinguish these dishes from other dishes of the experiment, as can be seen by studying Figures

22 - 25.

A comparison of the machine counts for the V79 culture was made also, and is shown in Figure 26. No comparison was made with hand counts for these. Because the edges of the V79 colonies have different appearance from those of the T1 strain, it is not suprising to find that the same difference in size used to correct for threshold differences would not be applicable. The figure shows that using a substantially larger minimum size of 444 microns for counting the control dishes brings excellent agreement. The seemingly large jump in minimum required to bring agreement is a consequence of the deep valley near size 250 microns in Figure 20. The value of 444 microns was chosen to force the counts of CFIN26 and CFIN28 into agreement for the control dishes. Again it appears that the machine counts agree very well for the two runs, even at widely different thresholds, once the constant difference in size due to threshold is taken into account.

The Flying Spot Digitizer has been used successfully to enumerate normal and irradiated human cell colonies. In addition, clone size distributions have been measured over a wider range than is currently available with comercially available systems. Some information on the clone shape and structure have also been obtained. Together these measurements provide new data which go beyond that to which radiation biologists usually have access. The FSD, which was developed in the Physics Division of LBL with the financial support of the U. S. Atomic Energy Commission, is available for the relatively moderate cost of operation and maintenance.

There are some drawbacks with regard to the use of this system for colony counting. The IBM 7094 computer is bulky and old. Maintenance costs are higher than would be the case for a newer computer; however, the high speed of operation, efficient use of memory, and capability for precise measurements of the total system warrants its serious consideration for counting of cell clones. It should also be mentioned that because the FSD-PETRI system uses film, it can analyse data obtained remotely at other laboratories.

Analysis of biological data has been made feasible by supplementing the FSD and its general pattern recognition programs with the new computer program PETRI. A procedure has been developed whereby human cell colonies grown <u>in vitro</u> and stained for contrast, are photographed on long strips of 35 mm film. The FSD-PETRI system digitizes and analyses the colony data. There are

three threshold settings which can be manipulated in this system: (1) the light intensity setting used for digitizing the colony patterns, (2) the distance threshold used to resolve the integrity of the digitized clusters, and (3) the threshold setting used to determine the minimum colony diameter to be enumerated.

In a single ninety dish experiment, preliminary tests were made at various threshold settings in order to find agreement with counts of the colonies made by hand. One group of threshold settings resulted in colony counts which showed a high degree of correspondence between hand and machine counts of control and irradiated human cell colonies.

Work toward continuing improvement of the clone identification algorithm needs to be done, but from this early progress, we conclude that the FSD-PETRI system holds promise for its successful use as a means of routinely enumerating cell colonies grown on petri dishes, more rapidly, with greater objectivity and providing considerably more information than is possible with hand counts.

The authors would like to acknowledge the excellent programming assistance of Shirley Buckman, Joan Franz, and Wen-Sue Gee in developing the procedures applicable to petri dish measurements. We are also grateful to Karen Smith for her valuable technical assistance in the cell culture work, to Laurie Craise for drafting the figures of this report, and to Robert Bigelow of the Technical Photography section at LBL for his assistance in testing film parameters.

Special thanks are extended to Cornelius Tobias of the Biomedical Division for his helpful suggestions; to James Baker, Head, and Robert Harvey, Associate Head of the former Mathematics and Computing Group at LBL; and to Carl Quong, Head of the Computer Science and Applied Mathematics Department for their support and encouragement.

The original FSD development was under the auspices of the U. S. Atomic Energy Commission, predecessor to the U. S. Energy Research and Development Administration which funded the present development. Work was performed under contract number W-7405-eng-48.

1. Puck, T. T., Marcus, P. I. and Ciecura, S. J. (1956) Clonal growth of mammalian cells in vitro: Growth characteristics of HeLa cells with and without a feeder layer. J. Exp. Med. 103, 273-284.

2. Puck, T. T. and Marcus, P. I. (1956) Action of X-rays on mammalian cells. J. Exp. Med. 103, 653-669.

3. Nias, A. H. W. and Fox, M. (1968) Minimum clone size for estimating normal reproductive capacity of cultured cells. Br. J. Radiol. 41, 468-474.

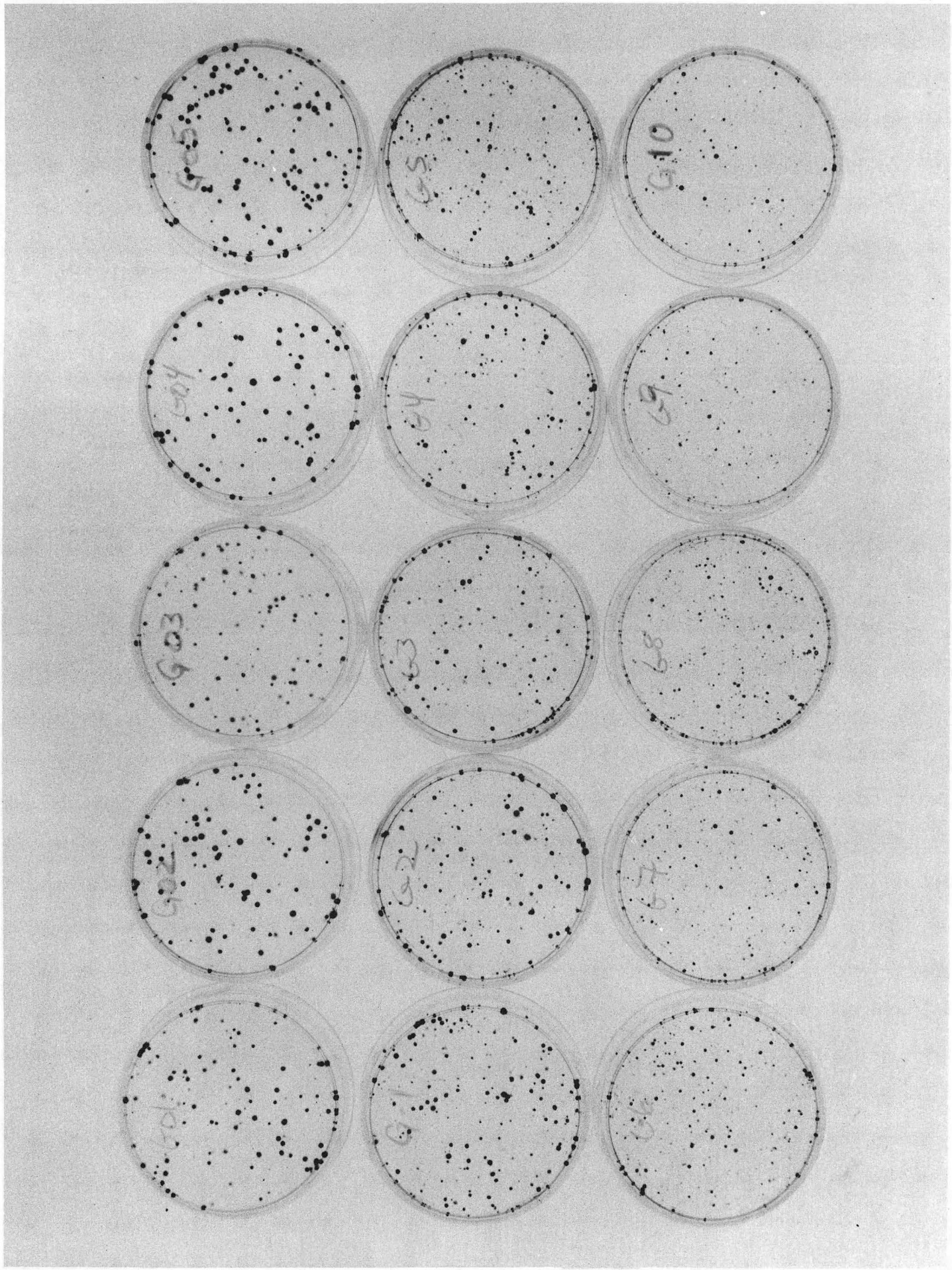
4. Sevastopoulos, C. (1976) Automated isolation and identification of thermal sensitive bacterial mutants defective in DNA replication. Ph.D. thesis. University of California, Berkeley.

5. Fisher Automatic Bacterial and Cell Colony Counter. Fisher Scientific Company, Pittsburgh, Pa.

6. P. V. C. Hough and B. W. Powell, "A Method for Faster Analysis of Bubble Chamber Photographs", Nuovo Cimento X, 18, 1184 (1960).

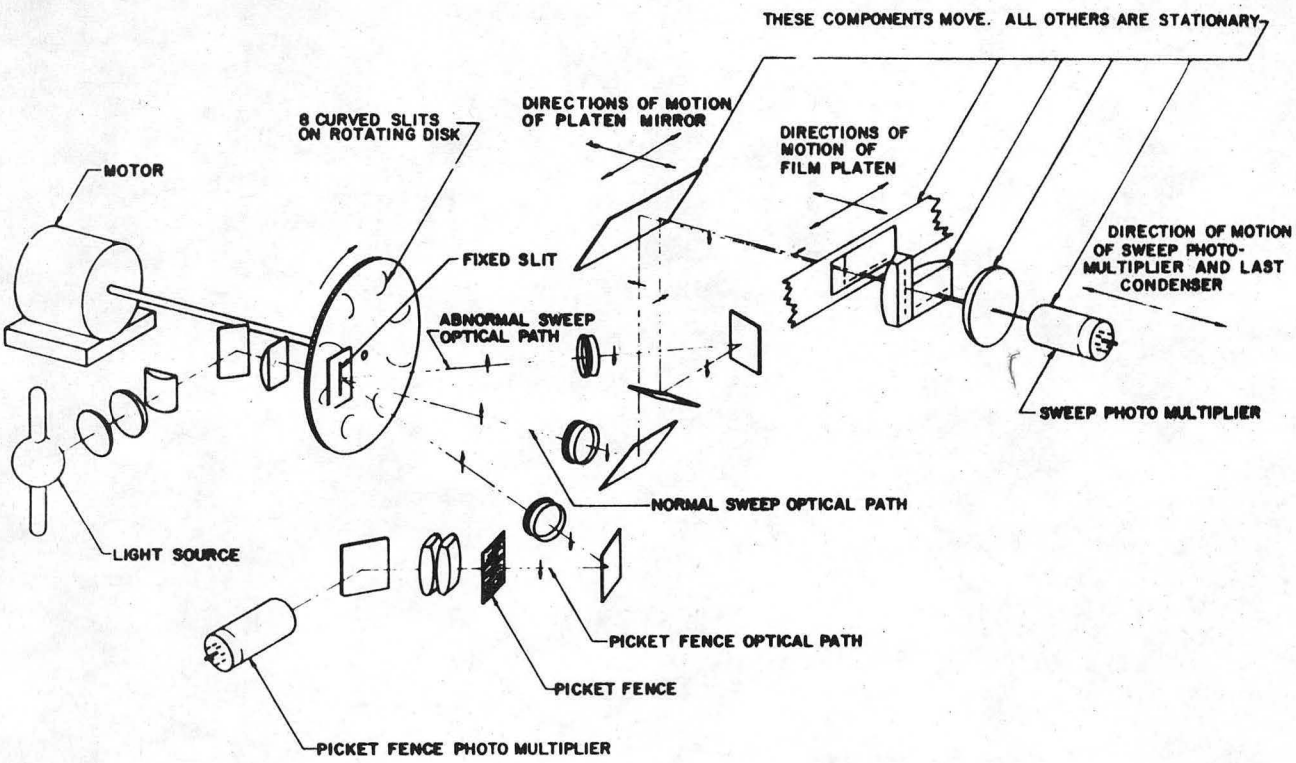
7. H. Holmes, "GRAFPAC", LBL Computer Center Writeups, UCID 3600.

8. "SESAME", LBL Computer Center Writeups, UCID 3600.



XBB 766-5414

Figure 1

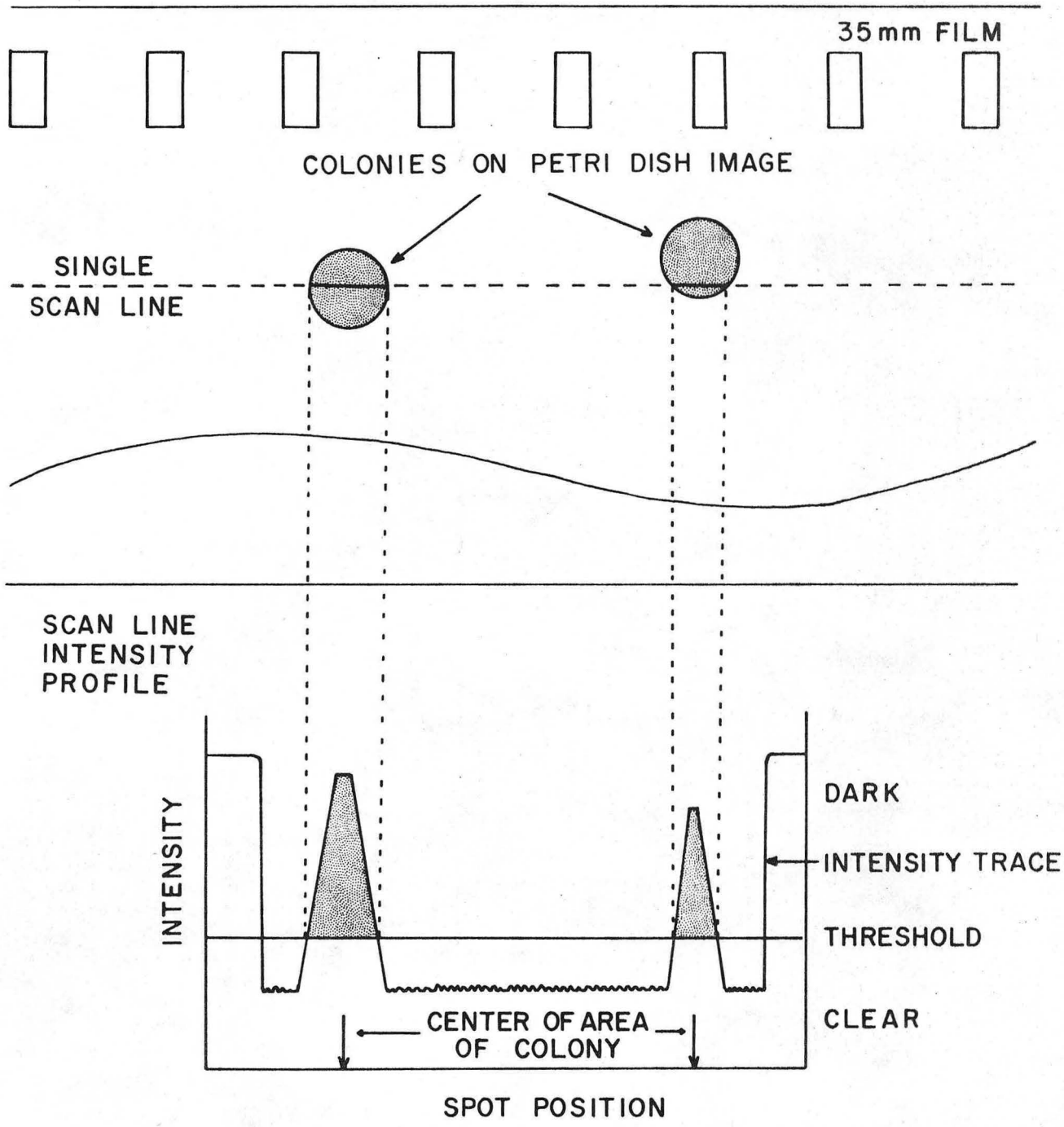


FLYING SPOT DIGITIZER
OPTICAL SCHEMATIC
UCLRL BERKELEY

XBL 764-1219

Figure 2

IMAGE DIGITIZATION PROCESS

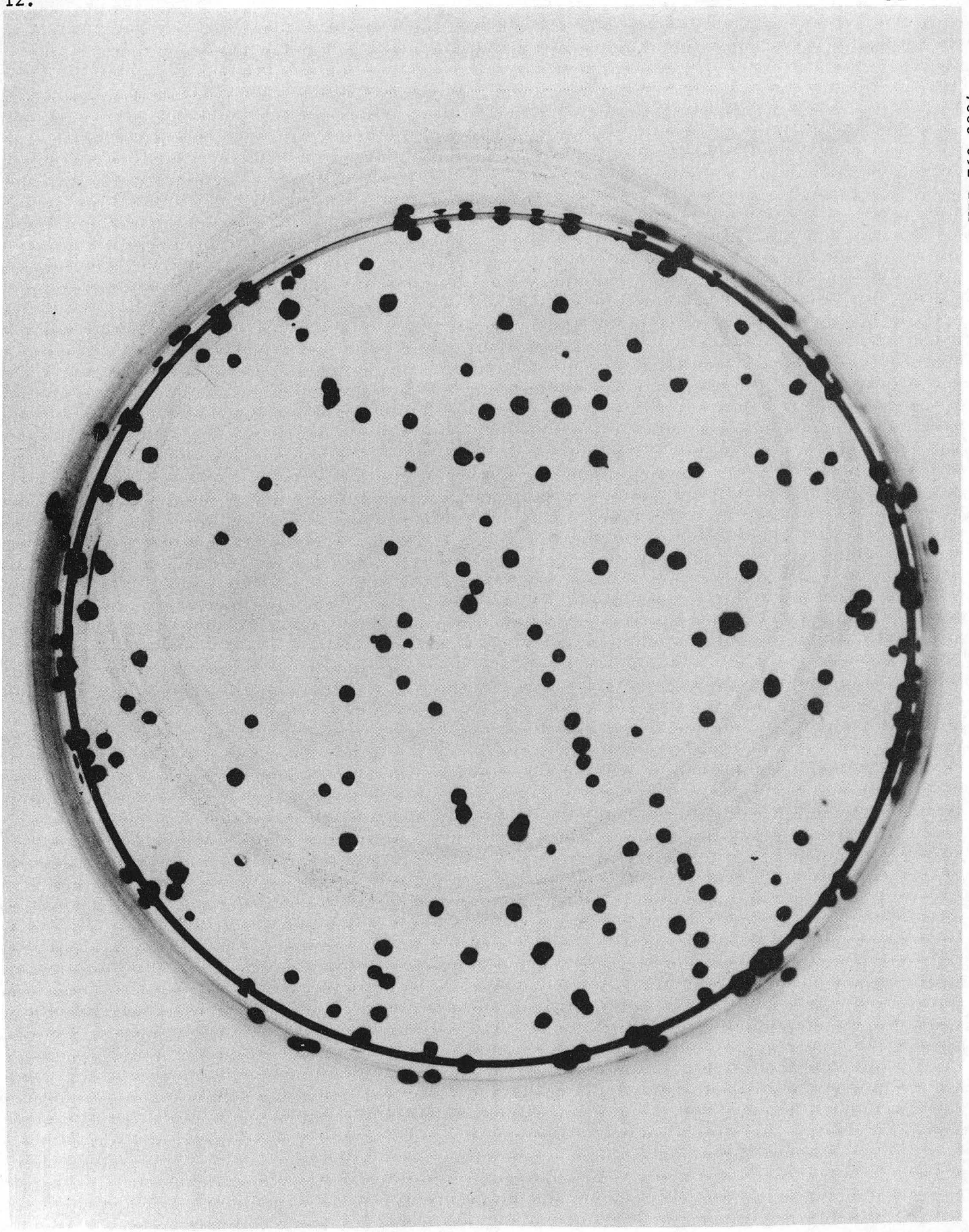


XBL 766-8489

Figure 3

XBB 7511-8632

Figure 4



XBB 763-3004

Figure 5

Dish No. 93
500R
Dark-long exposure

XBL 766-8490

Figure 6

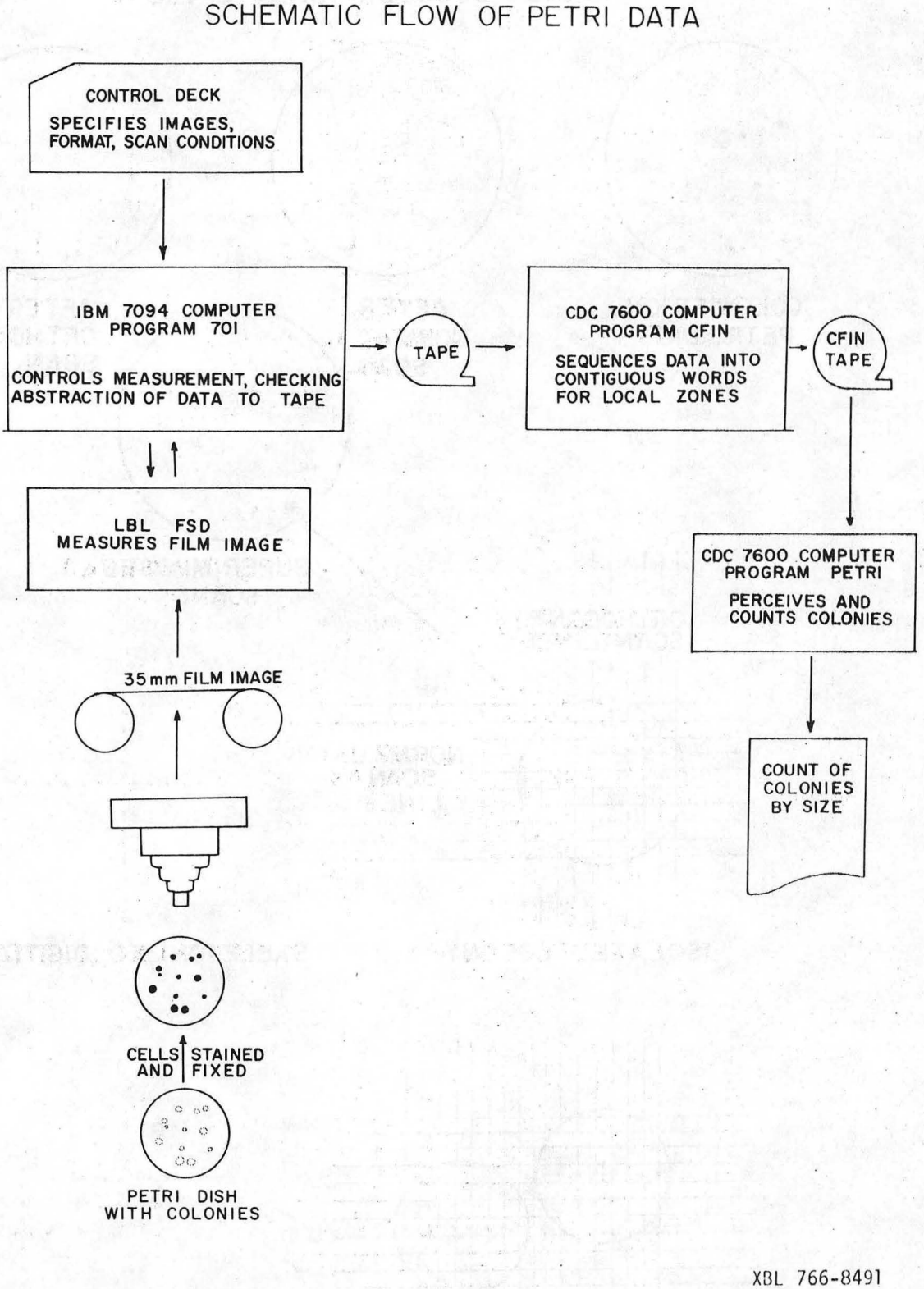


Figure 7

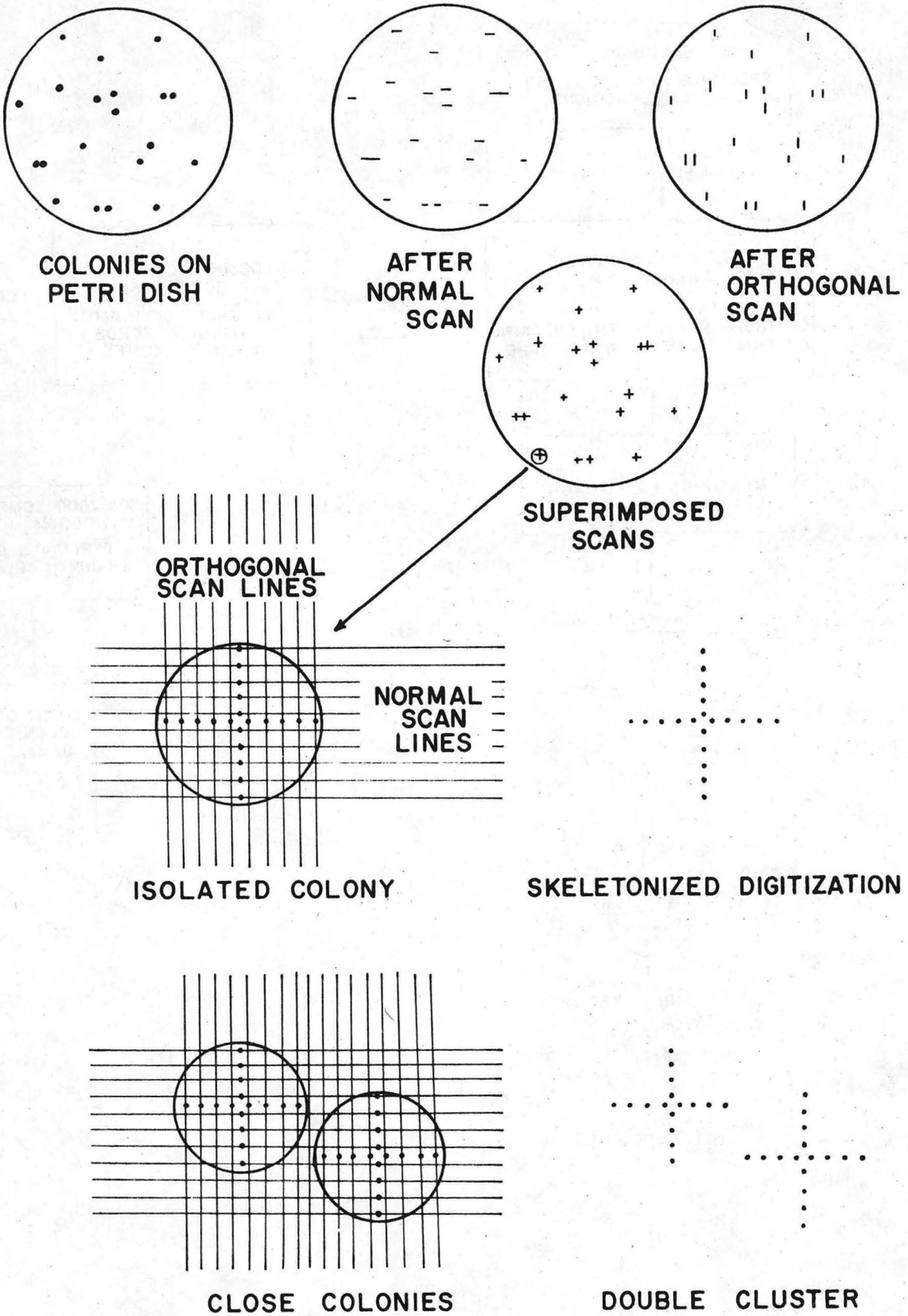
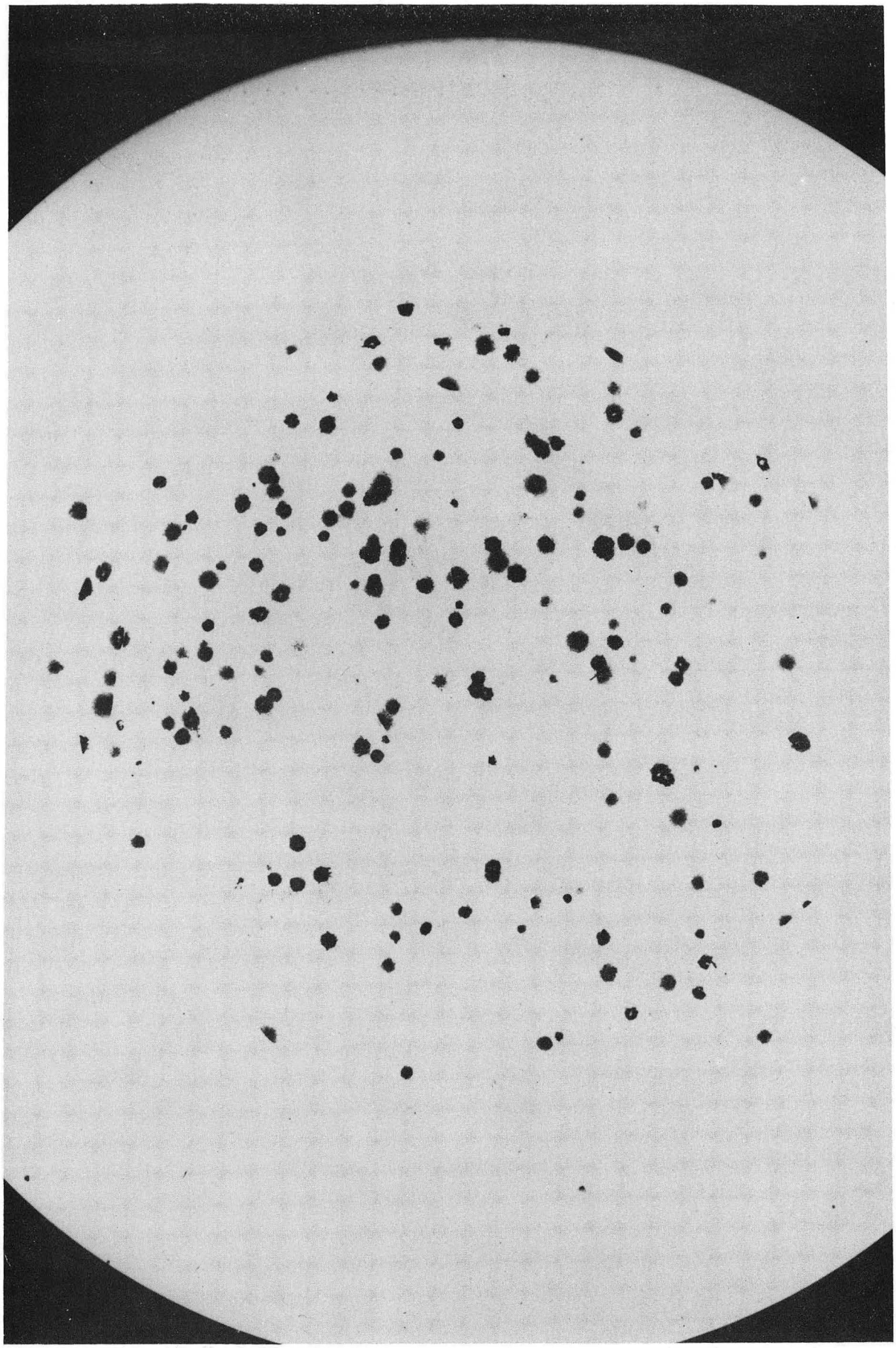
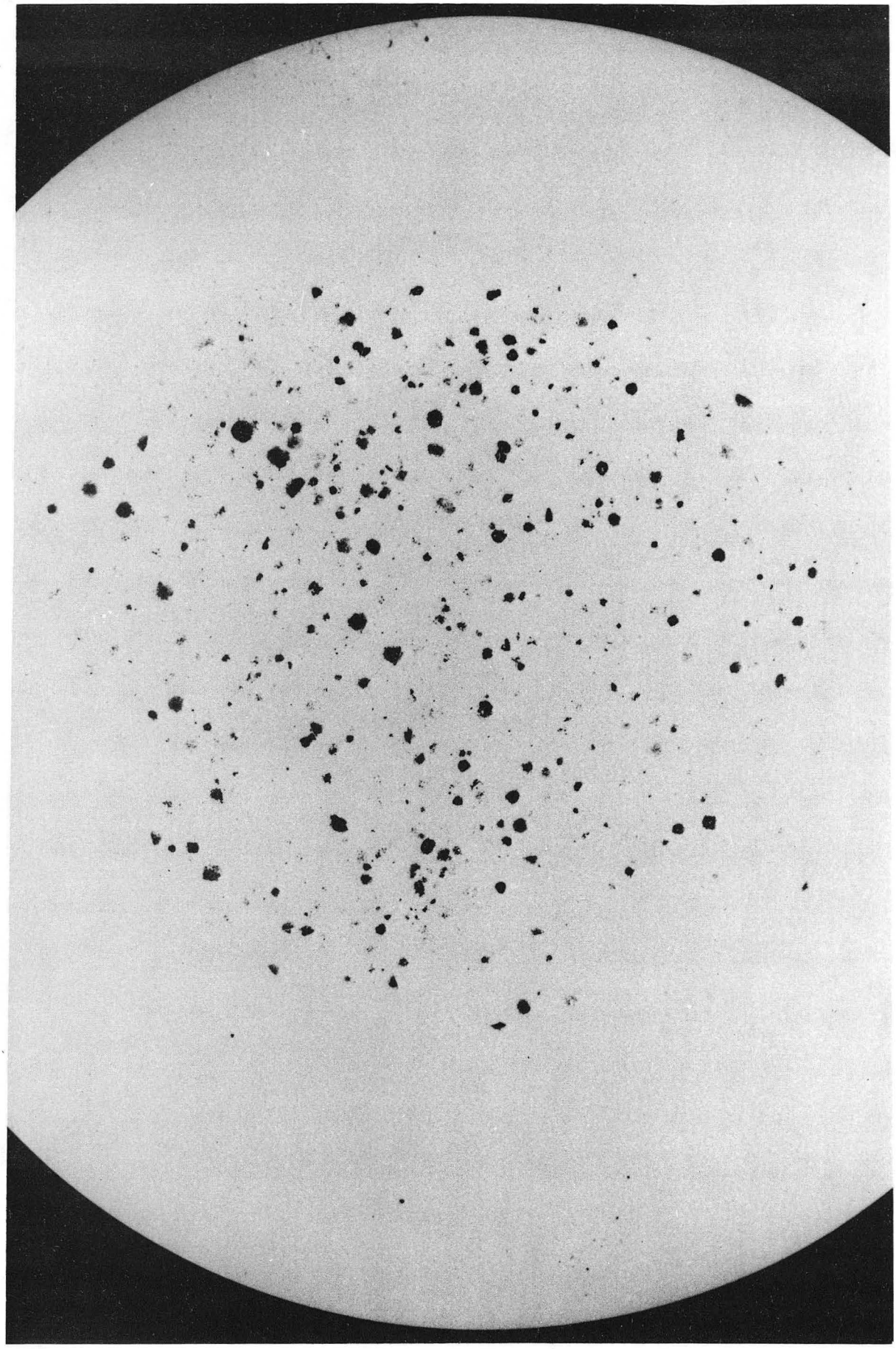

IDEAL CLUSTER DIGITIZINGS
COLONIES ON PETRI DISH
AFTER NORMAL SCAN
AFTER ORTHOGONAL SCAN
SUPERIMPOSED SCANS
ORTHOGONAL SCAN LINES
NORMAL SCAN LINES
ISOLATED COLONY
SKELETONIZED DIGITIZATION
CLOSE COLONIES
DOUBLE CLUSTER
XBL 766-8480

Figure 8



XBB 763-3004

Figure 9



XBB 763-3012

Figure 10

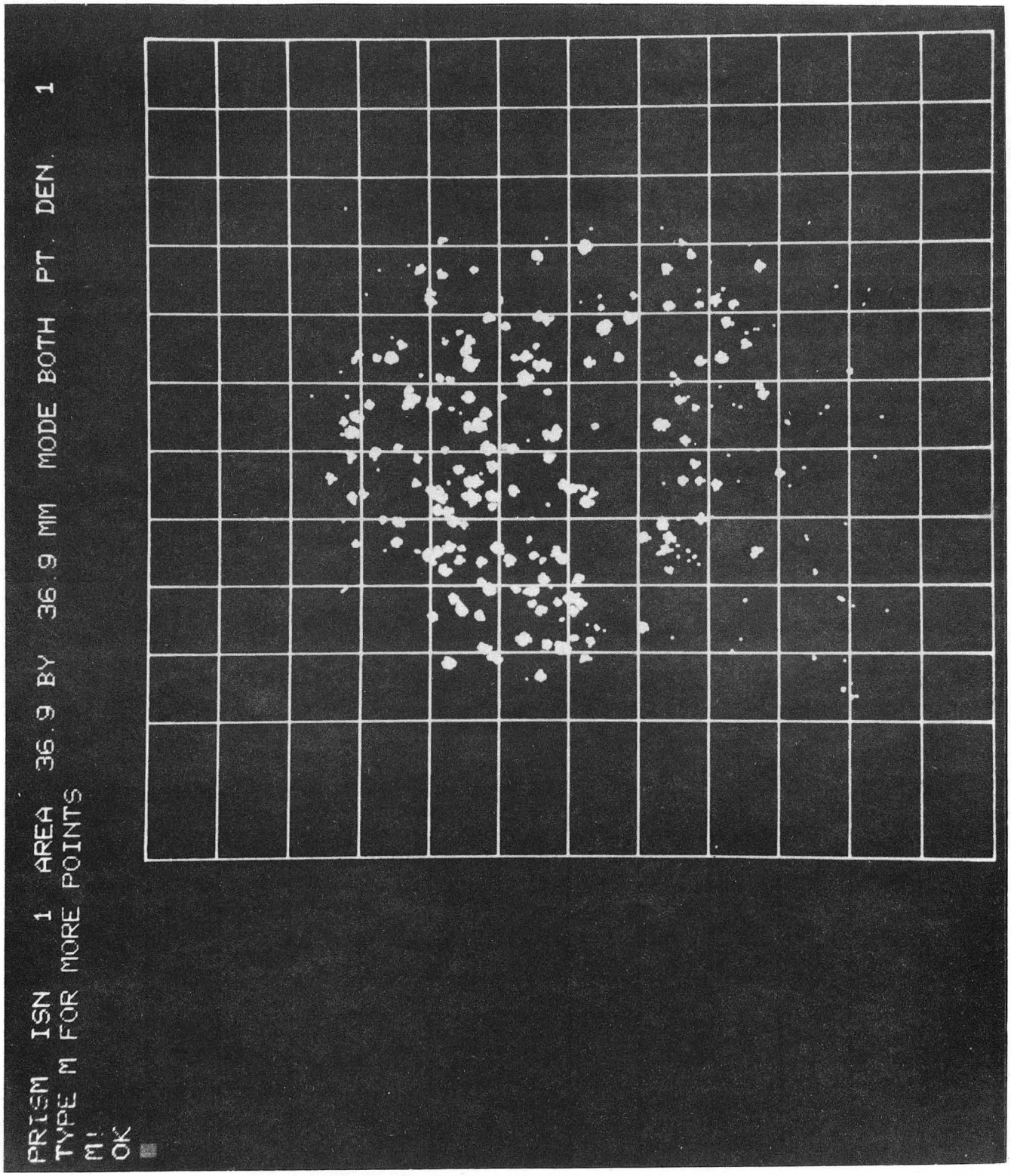


Figure 11

XBB 763-3001

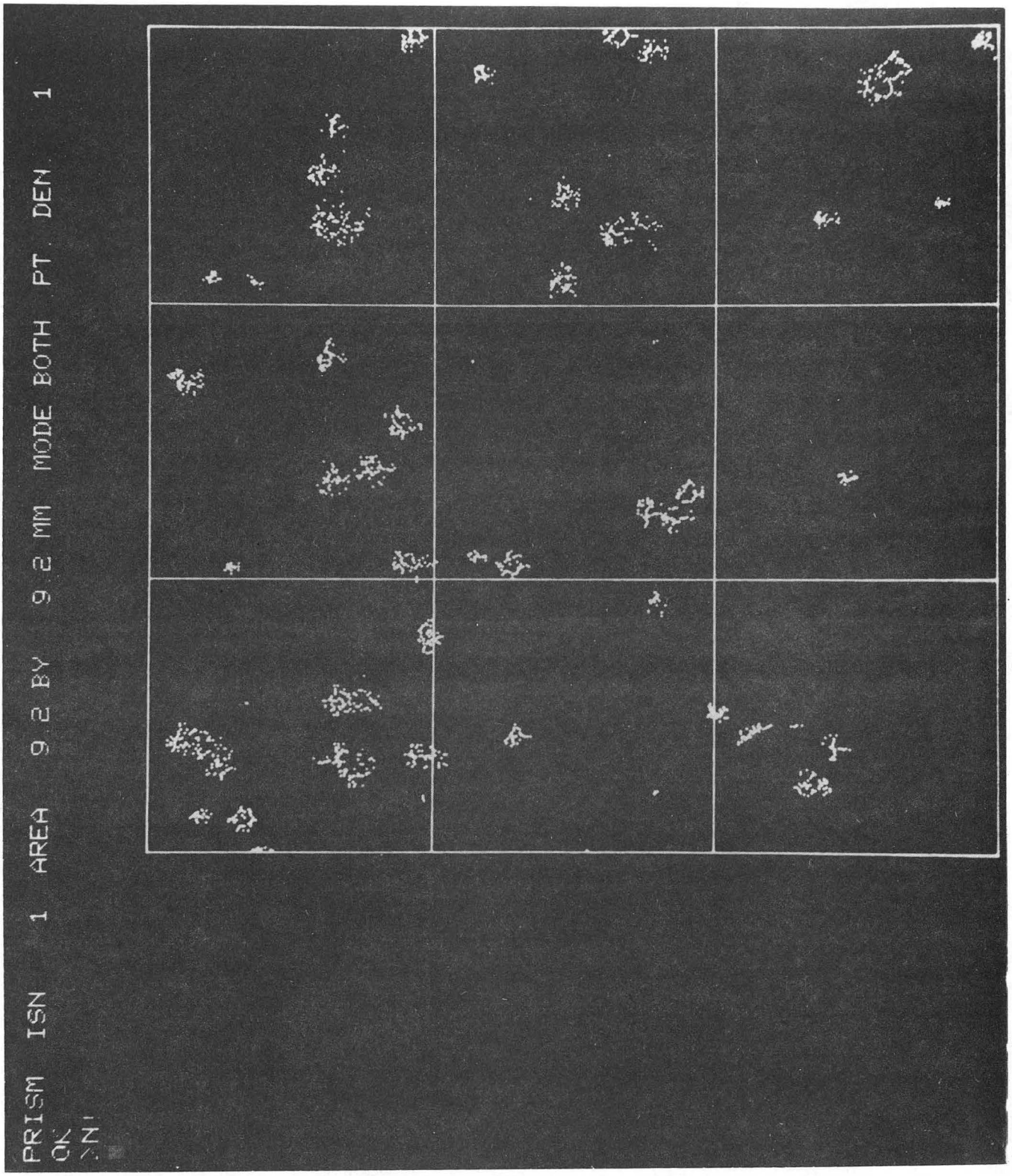


Figure 12

XBB 763-2999

Figure 13

XBB 763-3002

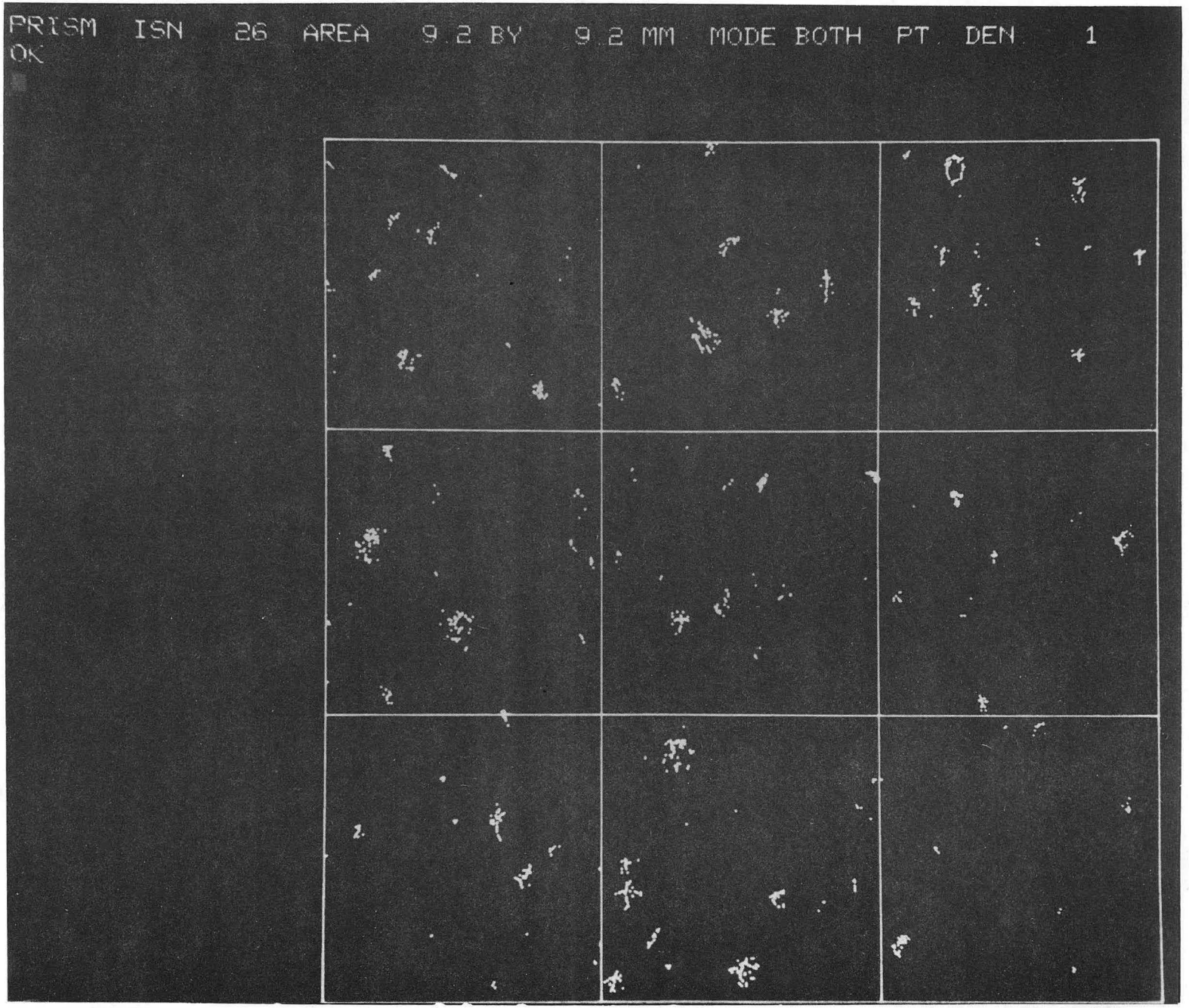
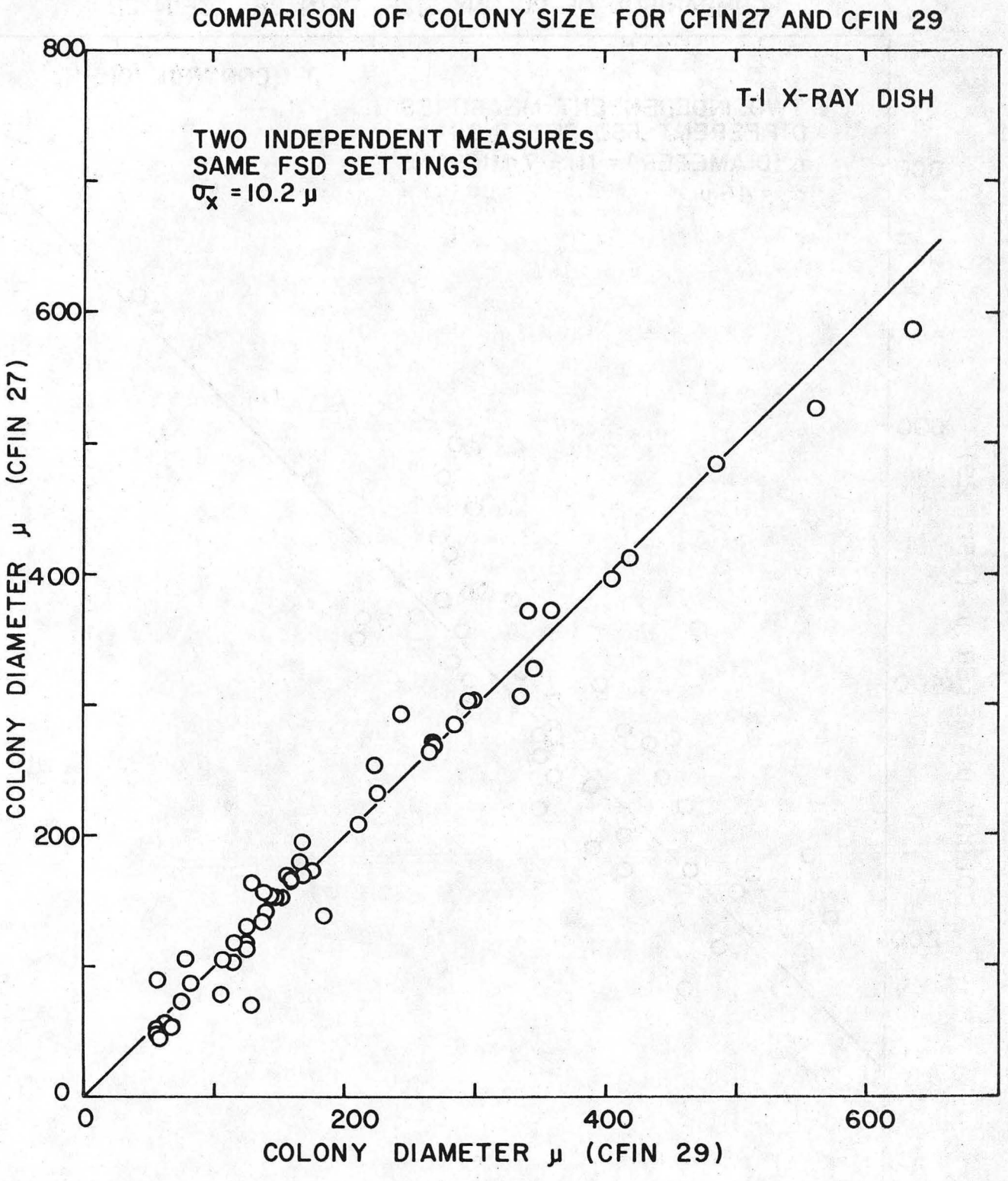


Figure 14 XBB 763-2998



XBL 766-8485

Figure 15

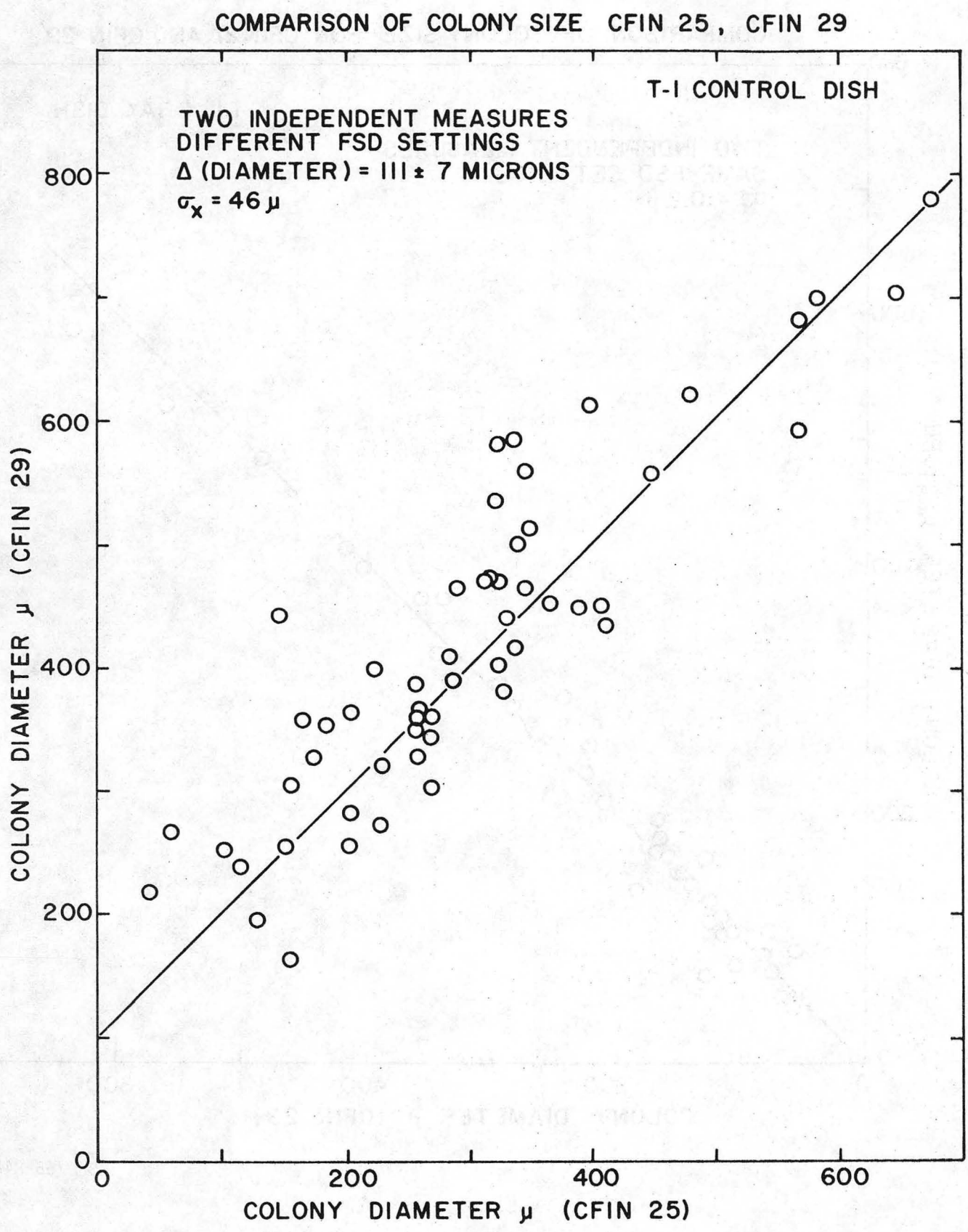
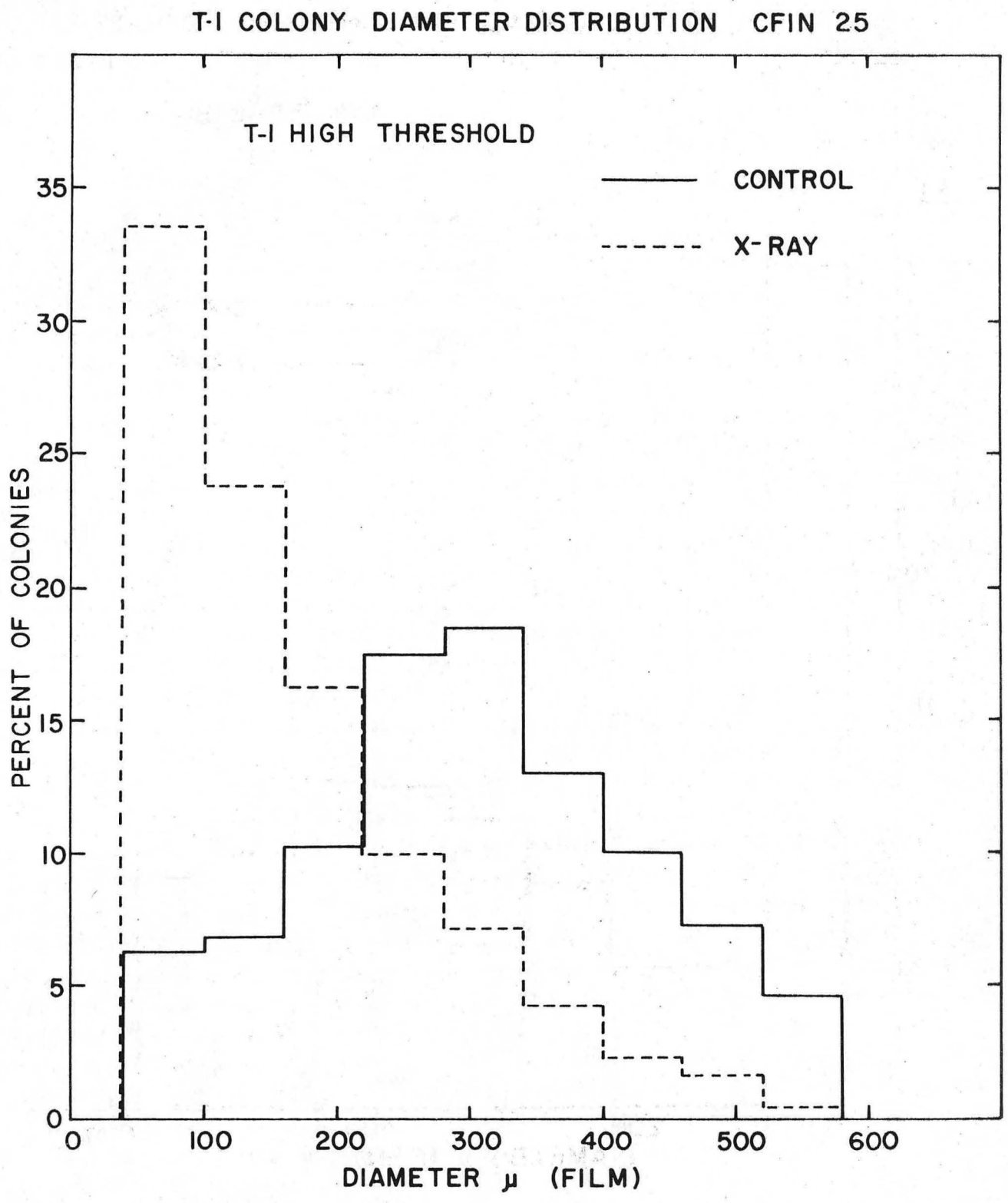

COMPARISON OF COLONY SIZE CFIN 25, CFIN 29
T-I CONTROL DISH
TWO INDEPENDENT MEASURES
DIFFERENT FSD SETTINGS
Δ (DIAMETER) = 111 ± 7 MICRONS
σ_x = 46 μ
COLONY DIAMETER μ (CFIN 29)
COLONY DIAMETER μ (CFIN 25)
0
200
400
600
800
XBL 766-8482

Figure 16



XBL 766- 8483

Figure 17

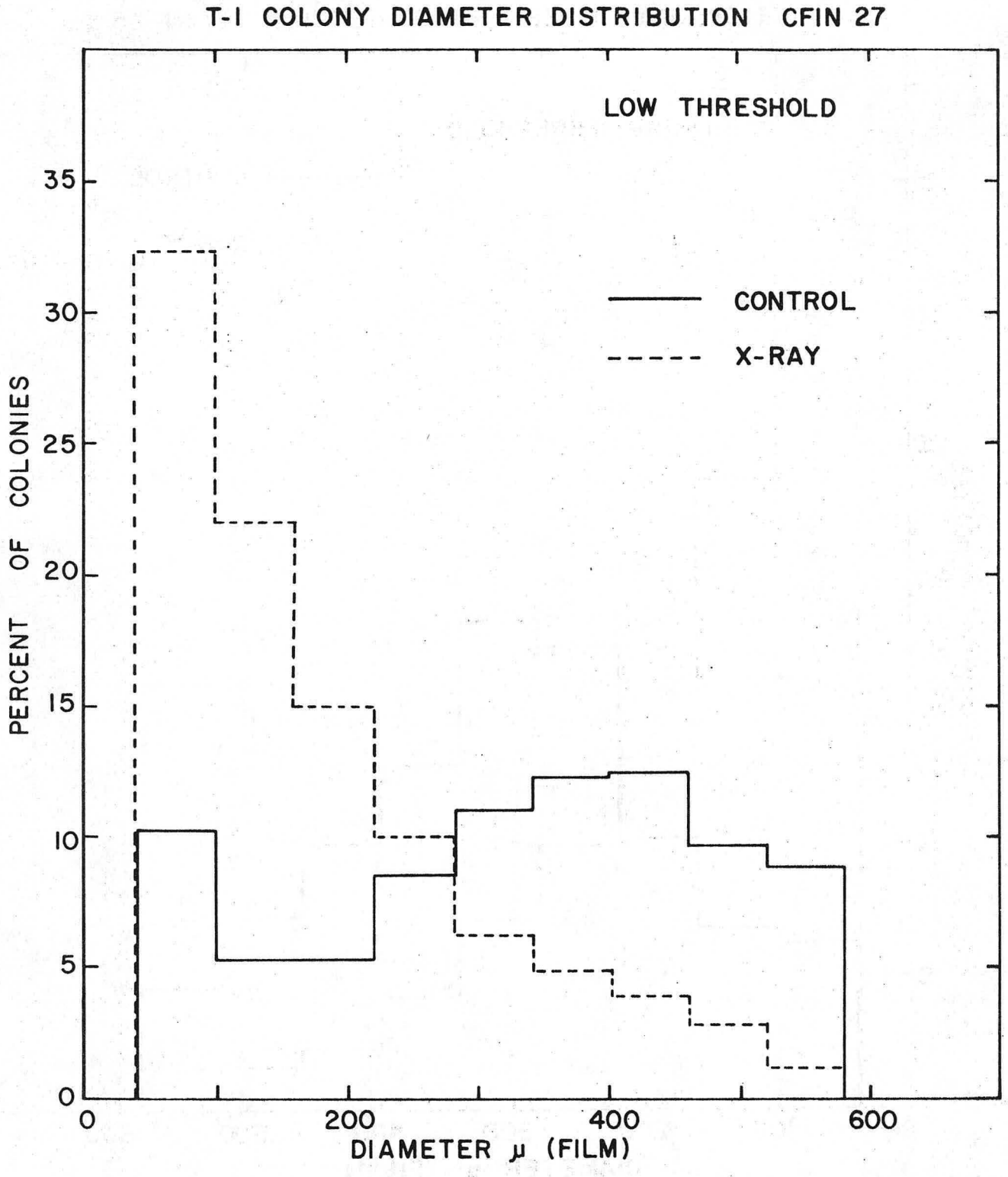

T-I COLONY DIAMETER DISTRIBUTION CFIN 27
LOW THRESHOLD
CONTROL
X-RAY
PERCENT OF COLONIES
0
5
10
15
20
25
30
35
0
200
400
600
DIAMETER μ (FILM)
XBL 766-8486

Figure 18

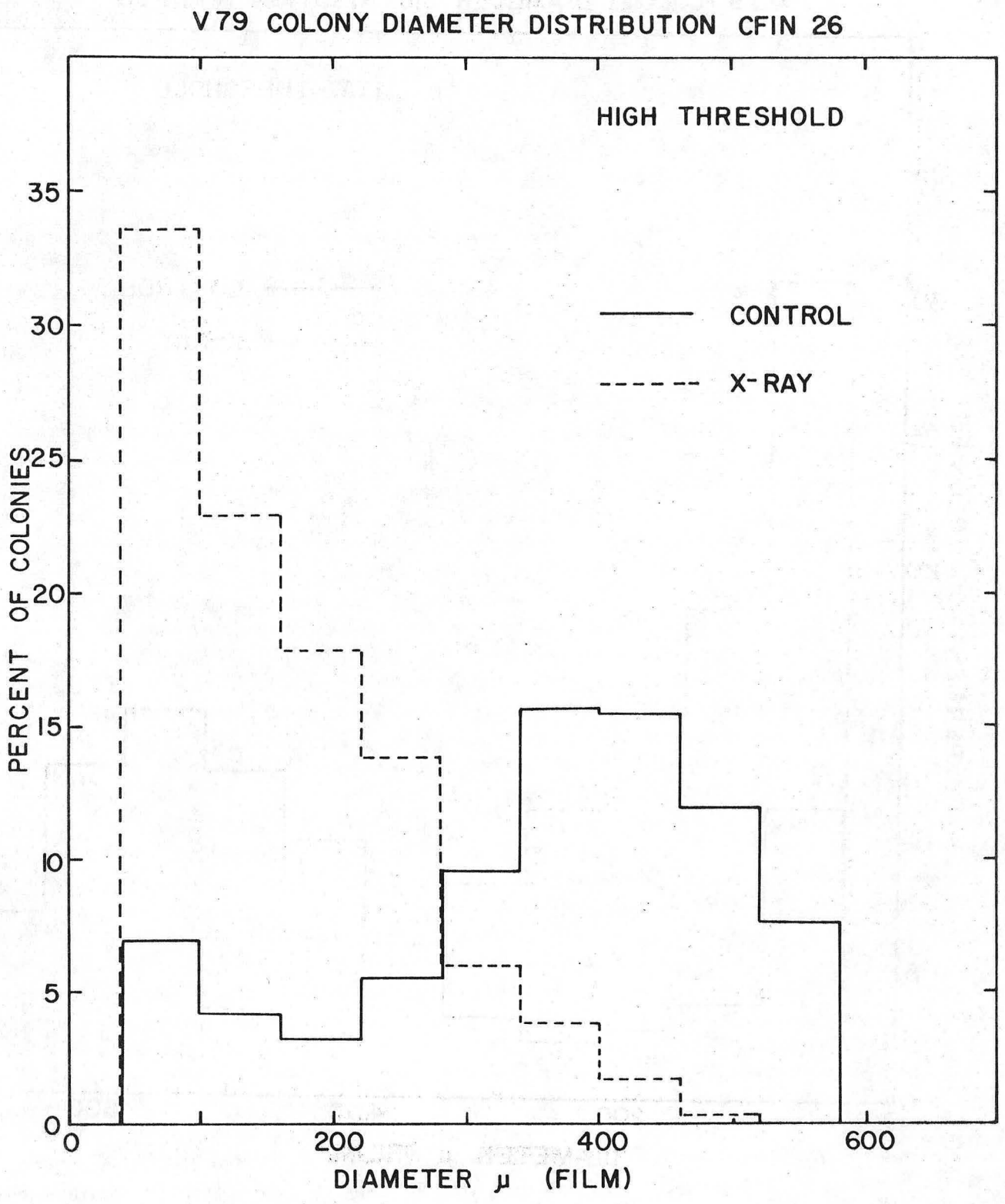

V 79 COLONY DIAMETER DISTRIBUTION CFIN 26
HIGH THRESHOLD
CONTROL
X-RAY
PERCENT OF COLONIES
0
5
10
15
20
25
30
35
0
200
400
600
DIAMETER μ (FILM)
XBL 766-8487

Figure 19

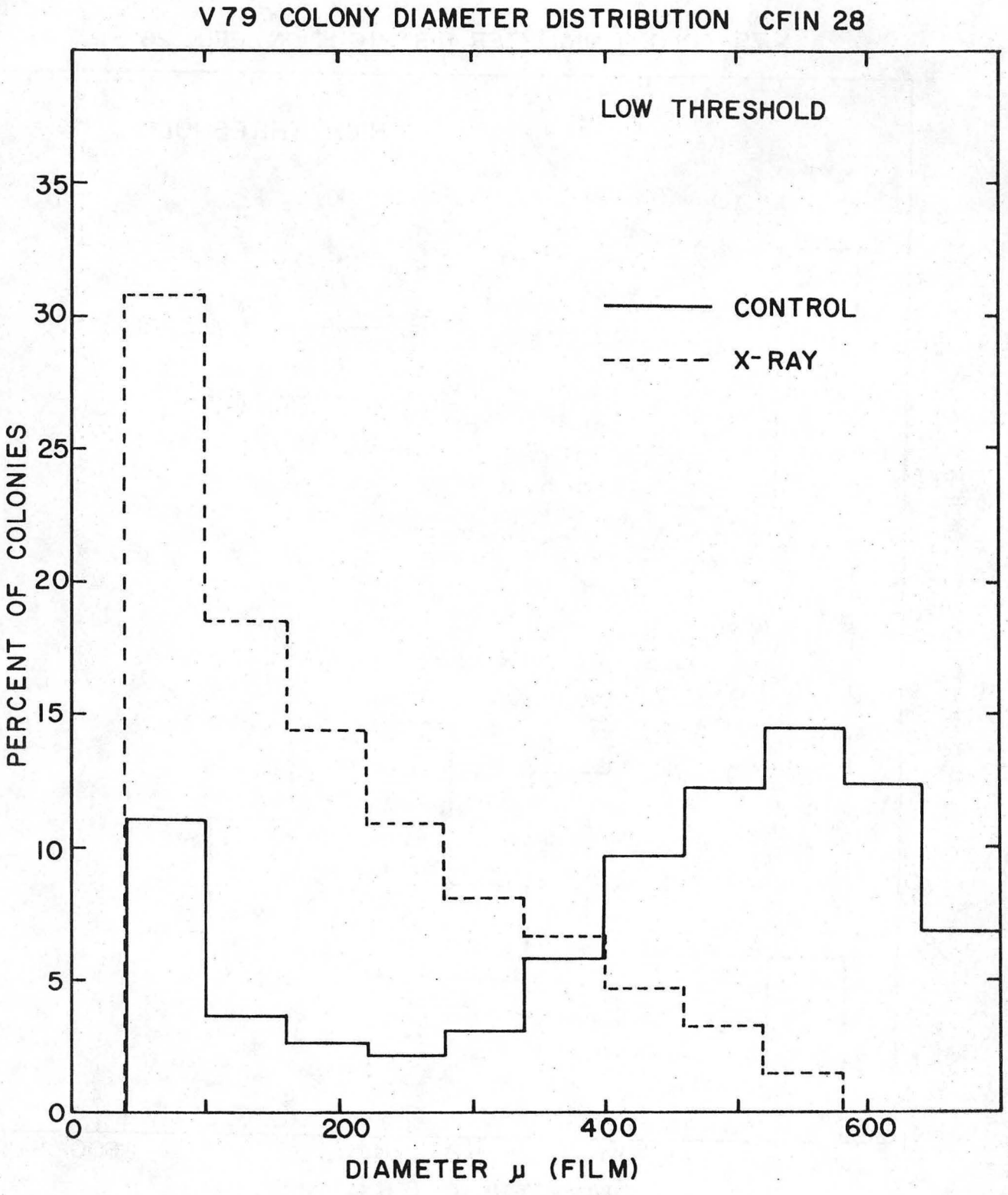

V79 COLONY DIAMETER DISTRIBUTION CFIN 28
LOW THRESHOLD
CONTROL
X-RAY
PERCENT OF COLONIES
0
5
10
15
20
25
30
35
0
200
400
600
DIAMETER μ (FILM)
XBL 766-8488

Figure 20

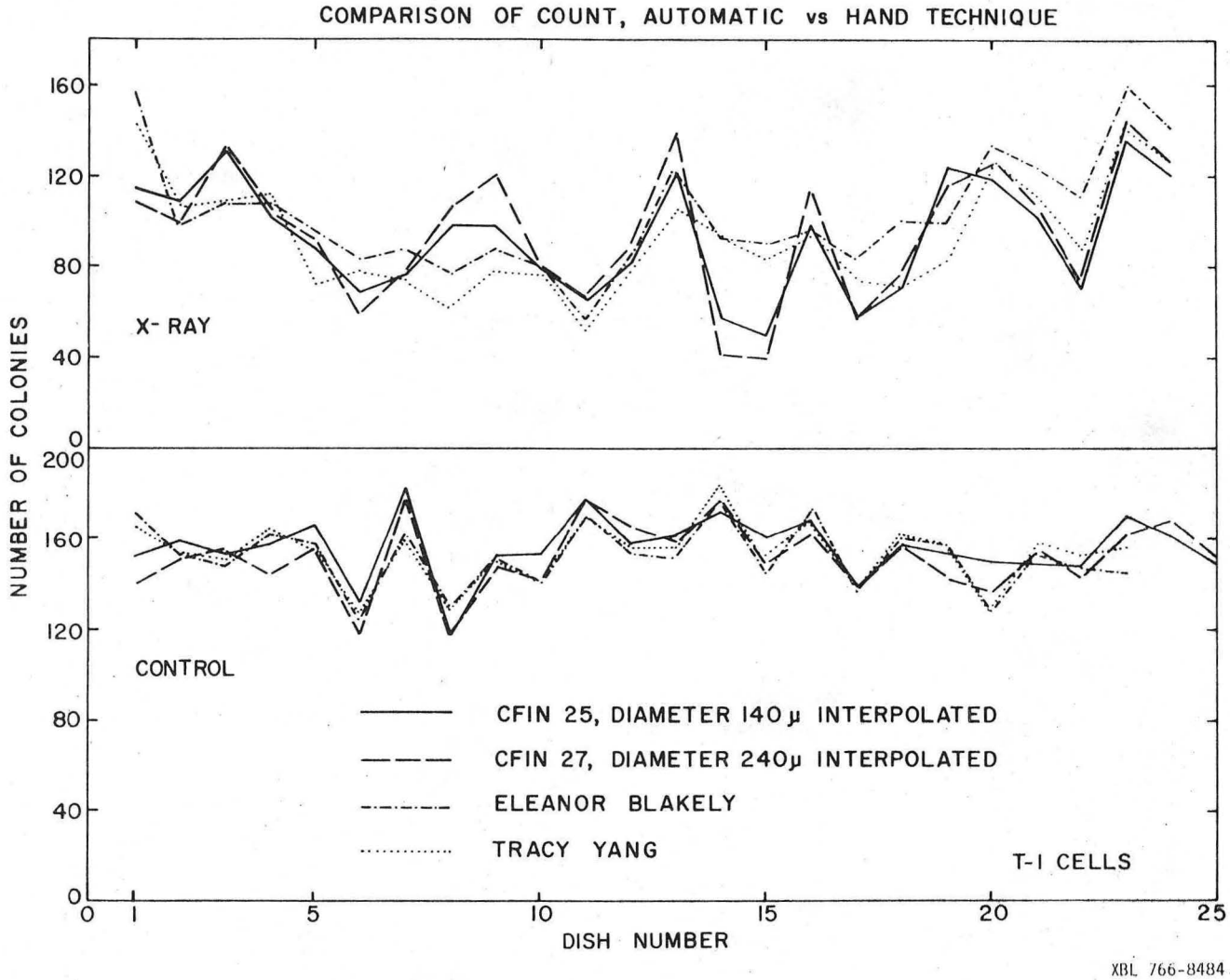
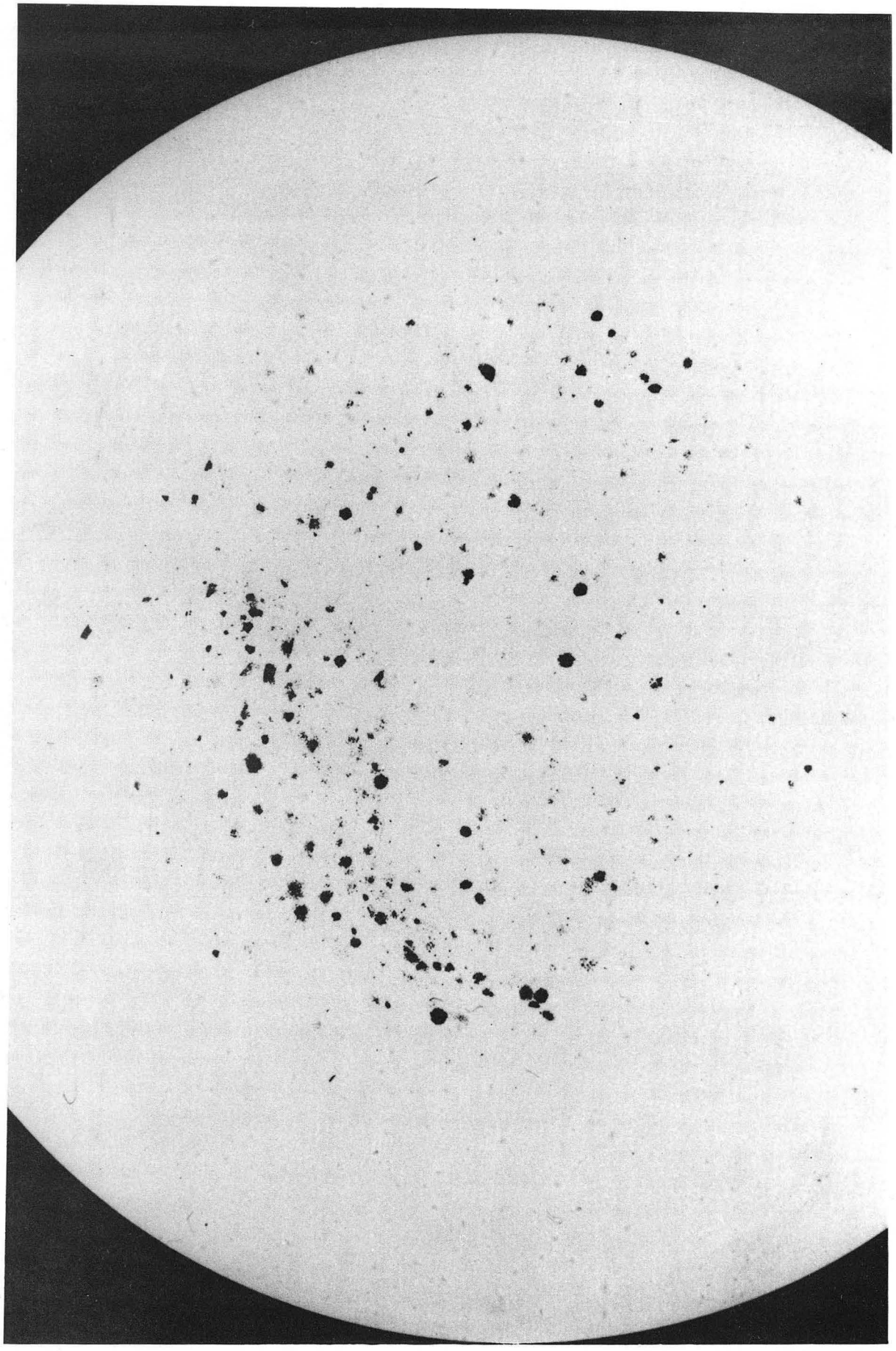
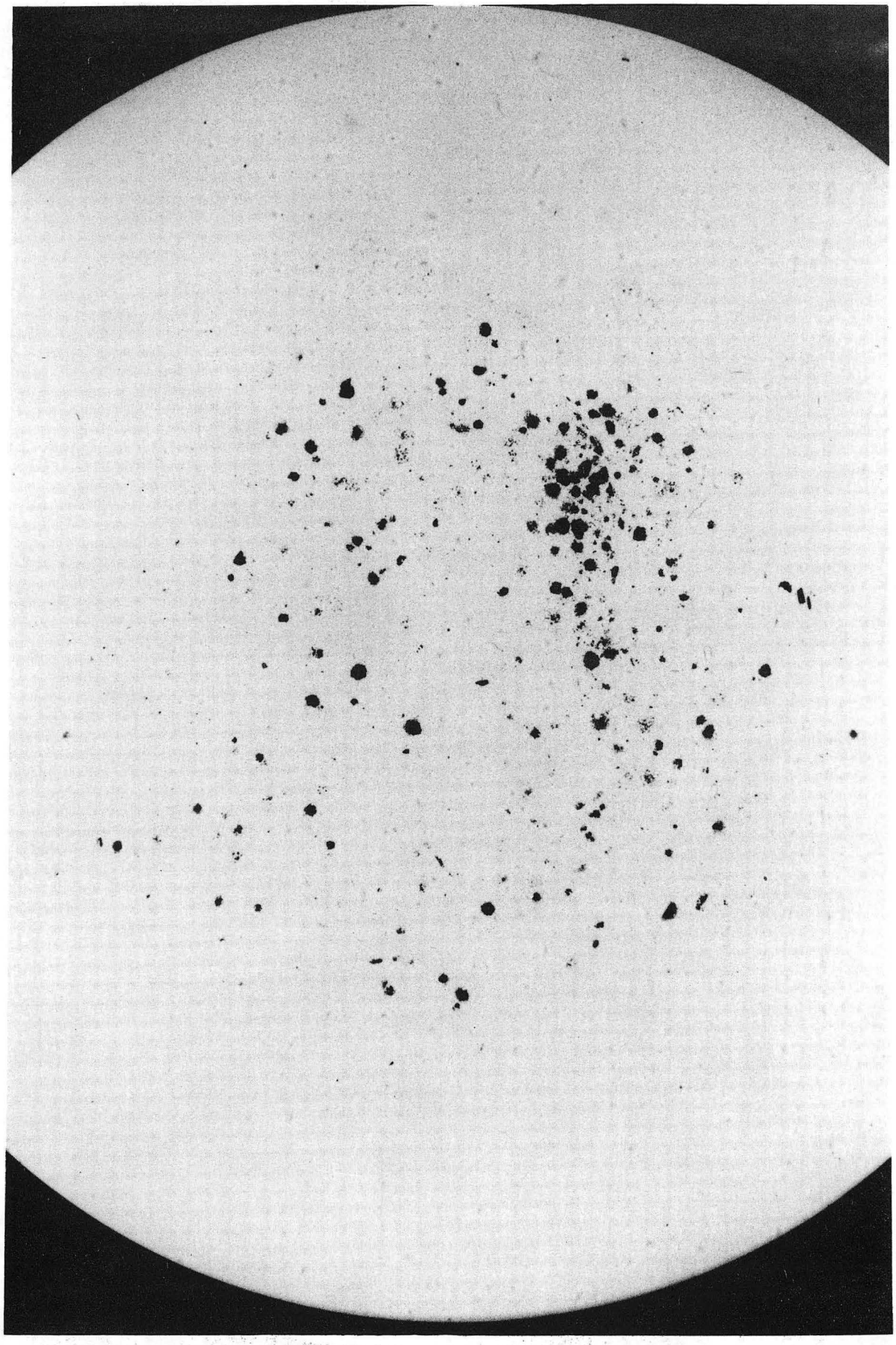


Figure 21



XBB 763-3011

Figure 22

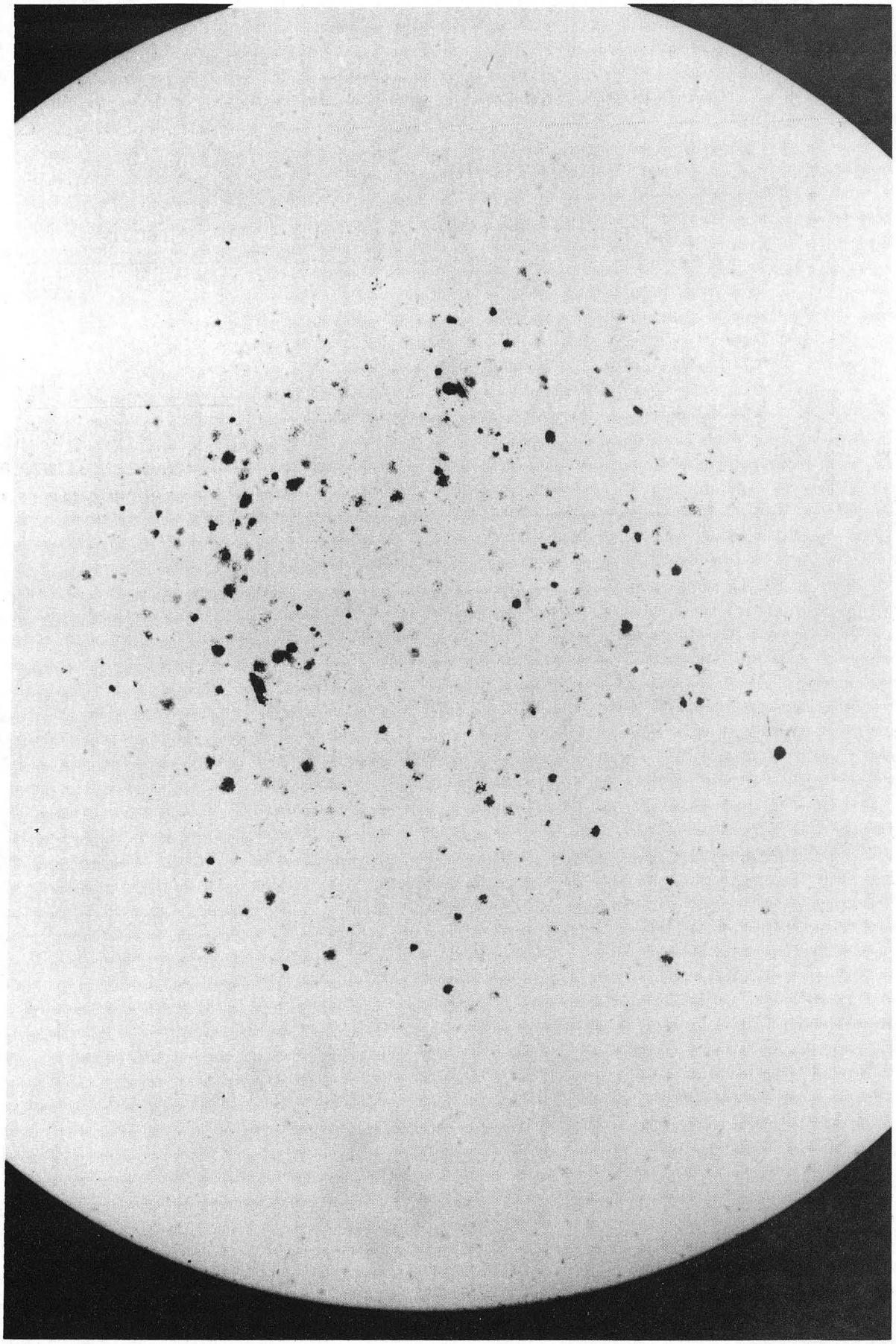


XBB 763-3009

Figure 23

XBB 763-3007

Figure 24



XBB 763-3008

Figure 25

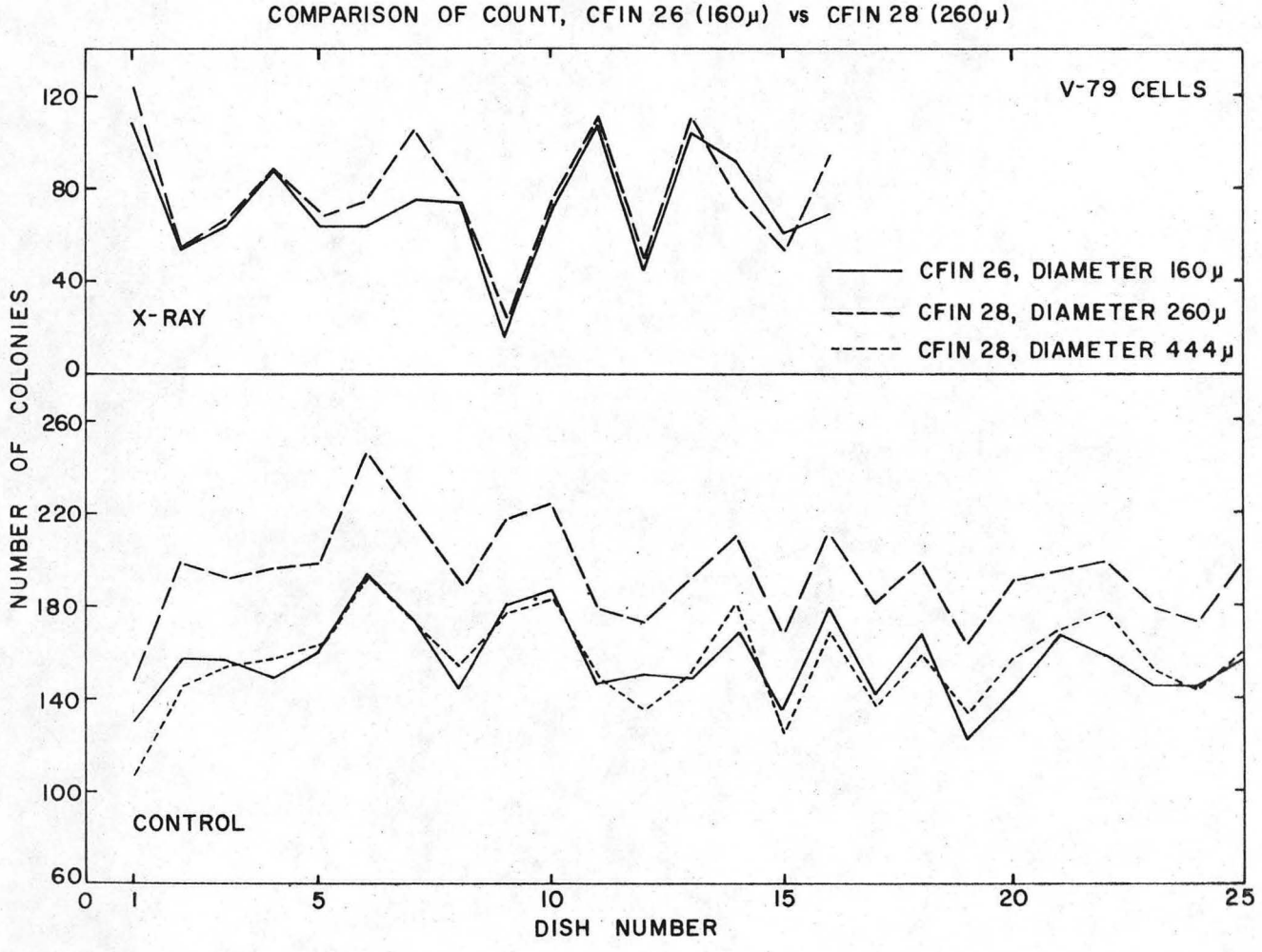

COMPARISON OF COUNT, CFIN 26 (160μ) vs CFIN 28 (260μ)
V-79 CELLS
X-RAY
CONTROL
CFIN 26, DIAMETER 160μ
CFIN 28, DIAMETER 260μ
CFIN 28, DIAMETER 444μ
NUMBER OF COLONIES
DISH NUMBER
120
80
40
0
260
220
180
140
100
60
0
1
5
10
15
20
25
XBL 766-8481

Figure 26

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720