

# A High Sensitivity, High Throughput, Automated Single-Cell Gel Electrophoresis ('Comet') DNA Damage Assay

B. Vojnovic<sup>(1)++</sup>, P.R. Barber<sup>(1)</sup>, P. Johnston<sup>(1)</sup>, H.C. Gregory<sup>(1)</sup>,  
B. Marples<sup>(2)</sup> and M.C. Joiner<sup>(2)</sup>, R.J. Locke<sup>(1)</sup>,

<sup>(1)</sup> Gray Cancer Institute, P.O. Box 100, Mount Vernon Hospital, Northwood, HA6 2JR, Middlesex, United Kingdom

<sup>(2)</sup> Karmanos Cancer Institute, Wayne State University, Hudson Webber Building, 4100 John R, Detroit, MI 48201-2013, U.S.A.

For submission to:        Physics in Medicine and Biology

<sup>++</sup> Author to whom correspondence should be addressed at:

Gray Cancer Institute,  
P.O. Box 100, Mount Vernon Hospital, Northwood,  
Middlesex, HA6 2JR, United Kingdom  
Tel: 01923 828611;  
Fax: 01923 835210;  
email: vojnovic@gci.ac.uk

Pages:                    15

Figures:                7

Running title:        automated comet assay.

Key words:            image processing, comets, fluorescence, Hough transform, automation.

## Abstract:

We present an arrangement for performing, in an automated fashion, measurements of DNA damage and repair at the single cell level, using single-cell gel electrophoresis technique, more commonly known as the Comet Assay. The system is based on a standard motorised microscope and imaging components but sample scanning is performed through the capture of large, but high resolution images. Novel and appropriate image tiling and image processing steps are described in detail. These are used to identify 'comets', distinguish them from debris, segment them to avoid potential overlap and quantify DNA content distributions, using conventional fluorescence methods. The system is capable of automatically quantifying comets at a rate of around one per second. The ability to perform measurements on large numbers of cells, in an unbiased, operator-independent fashion is used to minimise noise and to provide improved experiment statistics. The ability to detect lower levels of damage allows the investigation of cell-cycle dependent damage phenomena. Representative data sets of DNA single-strand and double-strand break damage and repair are presented.

## 1. Introduction

The investigation of DNA damage and repair at the single cell level can be performed, in a very sensitive manner, using single-cell gel electrophoresis (SCGE)<sup>(1, 2, 3, 4, 5, 6)</sup>. Cellular DNA is readily identified using fluorescent markers and the migration of fragments from the cell nucleus through an agarose gel, under the influence of an electric field, results in a characteristic comet-like shape: the assay is thus also commonly known as the ‘Comet Assay’. Lower molecular weight DNA fragments suffer greater migration from the cell nucleus and characteristic information can be obtained from the shape of the comet ‘tail’. This is formed to one side of each cell nucleus, immobilised in the gel, with the cell nucleus commonly referred to as the comet ‘head’. Although SCGE methods have wide applications associated with DNA damage, in our case the primary interest is concerning applications in radiobiology.

Biological variability dictates the acquisition and characterisation of comets from large numbers of cells, typically several hundred. Most work with such assays is performed manually and is clearly both laborious and to some extent inevitably subjective. The user typically ‘finds’ a cell, focuses and acquires a digital image. This is subsequently quantified; a variety of approaches have been used<sup>(2,3,5,6)</sup>, often dictated by the particular biological measure required. A feature of most systems is that relatively high magnifications are used with consequently finer focus ranges. Clearly, sensitivity is ultimately defined by the dynamic range of the imager and more commonly, by the background staining intensity: a small but variable proportion of the fluorescent marker remains present in the gel and results in a non-zero black level. A variety of cell lines exhibit odd-shaped and highly structured nuclei. Although determination of the comet tail moment does not require ‘head’ delineation, it is necessary to determine the head centre as accurately as possible. The ability to delineate the head nevertheless allows other measurements to be made, such as the percentages of DNA in the head and in the tail and the tail length.

We present a fully automated microscopy machine vision image capture and analysis system for the collection of data from slides of ‘comets’. The novel image processing algorithms employed in delineating the ‘comet head’ from the ‘comet tail’ allow us to determine accurately very low levels of damage. In conjunction with calibrated and automated image capture methods, we are able to eliminate operator subjectivity and analyse large numbers of cells (>2500) in a short time (<1 hour). The image processing algorithm presented here is designed to handle particularly difficult nuclei containing a high degree of structure, due to DNA clumping. Previous work on the automation of the comet assay has used thresholding and binary morphology to segment both head and tail<sup>7</sup>. We found such methods to be unsuccessful for routine, unbiased use, particularly on more ‘difficult’ samples that exhibit structure in the ‘head’. Other more advanced systems utilise thresholding with hysteresis to ensure tail pixels are segmented correctly<sup>8</sup> together with a ‘high threshold value’ to segment the comet head. This method was also found to be insufficiently rugged for highly structured heads or when tail intensities are comparable to head intensities; examples of ‘difficult’ comets are shown in figure 1.

Our system allows for the analysis of highly structured comets through the use of an algorithm exploiting a Compact Hough transform and formation of a Radial Map of the comet head (CHARM)<sup>9</sup>. This algorithm has been extensively tested on a previously developed, manual, imaging station where a peak analysis rate of approximately 200 comets per hour was achievable.

We also present techniques used to extend the assay’s dynamic range by removing interfering background fluorescence and to define a region of interest. We primarily use the comet moment, defined by Kent *et al.*<sup>10</sup>, as a measure of DNA damage because of its simplicity to calculate and

reliability over a large dose range. In some instances, the ‘Olive’ moment, as defined by the product of the amount of DNA and the mean distance of migration in the tail <sup>(2)</sup> is more appropriate. If subtle biological variations are to be quantified (e.g. cell cycle dependant damage), then the use of large cell populations is dictated. Under those circumstances, the use of a fully automated system is particularly advantageous providing that the manner in which data is extracted does not introduce any inadvertent bias. In practice, it is essential that the image processing steps are geared towards the correct recognition of an acceptable cell nucleus, i.e. comet ‘head’.

The speed of the assay is however not only determined by the image processing but also by the time required to acquire an image and by the mechanical transit time taken to settle to the next area of interest. The availability of high resolution imagers allows us to minimise these times by reducing the total number of images, and hence stage motions, required to examine a complete sample slide.

## **2. Materials and methods**

### **2.1 Comet preparation**

Slides were prepared by coating semi-frosted glass slides (BDH) with 100 µl 1% PFGE agarose (Biorad) in ddH<sub>2</sub>O. After allowing to gel at room temperature, slides were dried at 60°C overnight. Precoating with agarose improves attachment of subsequent agarose layers. Cells were prepared by trypsinising 75% confluent cultures and diluting the single cell suspension to 5 x 10<sup>4</sup> cells ml<sup>-1</sup> using cold medium. A 1% (w/v) solution of low gelling temperature, type VII agarose (Sigma) containing 1 x Hank’s balanced salt solution (HBSS) was diluted 3:1 with the cell suspension and maintained at 37°C. 500 µl aliquots of the mixture were quickly spread onto glass slides, After gelling on a cold metal plate for 1 minute, the agarose embedded cells were transferred to 100mm petri dishes containing 25 ml cold medium and maintained on ice prior to irradiation.

Cells were irradiated on ice using a 240 kVp x-ray source at a dose rate of 0.23 Gy minute<sup>-1</sup>. For dose response curves, cells were irradiated with single doses of x-rays (0-10 Gy) according to the appropriate dosimetry schedule and then immediately submersed in lysis solution at room temperature.

#### **Neutral assay.**

For the measurement of DNA double strand breaks cells were lysed in a buffer containing 30mM EDTA, 0.5% SDS pH 8.3. The slides were placed in an oven at 50°C, and left to lyse for 3.5 hours in darkness. After this time, they were gradually cooled by transferring them to the refrigerator for 30 minutes, and maintained in darkness. The slides were washed for 1 hour in three changes of rinse solution (1 x TBE buffer) and electrophoresis was performed in fresh rinse buffer at 0.6 V cm<sup>-1</sup> for 30 minutes.

#### **Alkaline assay.**

For the measurement of DNA single strand breaks and alkali labile base damage, lysis was performed for 4 hours at 4°C in a buffer containing 1.2 M NaCl, 0.1% N-lauryl sarcosine, 0.26 M NaOH, 100mM Na<sub>2</sub>EDTA, pH 12.6. The slides were gently rinsed for 1 hour in three changes of an alkaline rinse solution (0.03 M NaOH, 2 mM Na<sub>2</sub>EDTA) and electrophoresis was performed in fresh rinse buffer at 0.6 V cm<sup>-1</sup> for 30 minutes.

Immediately after electrophoresis, slides were dehydrated by sequential treatment with 70, 95 and 100% ethanol solutions for 10 minutes each and air dried overnight. Slides were stained for 30 minutes in a 1:10000 dilution of SYBR-Gold (Molecular Probes Inc.) in TBS, rinsed in distilled

water and mounted using Miowiol-DABCO antifade and sealed using clear nail-varnish. Individual comets were analysed within 24 hours.

## 2.2 Image Acquisition

Slides were viewed by conventional wide-field fluorescence microscopy (excitation 450-490 nm, emission >590 nm, 505 nm dichromatic mirror) using a modified Nikon TE300 inverted microscope (Nikon UK Ltd, Kingston, Surrey, United Kingdom). Image capture is performed by a 2/3", 12-bit dynamic range, Peltier-cooled megapixel (1344 x 1024) CCD camera (type ORCA II-ER, Hamamatsu KK, Shizuoka, Japan), used with a IEEE 1394 (Firewire OHCI) interface. A feature of our arrangement is to maximise the captured field of view (approximately 1150 x 875  $\mu\text{m}$ ) through the use of a low power (10x) objective. The particular objective used (Nikon CF160 Super-Fluor) has a numerical aperture of 0.5, resulting in around 2.9-fold increase in collected emission light compared to the more usual, *ca* 0.3 numerical aperture, 10x objectives. This approach makes optimal use of the camera's resolution while capturing a reasonable number of individual comets in a single image. The camera is usually operated in 12 bit mode, though in some cases, where faster image transfer is required, only an 8 bit dynamic range is used, along with on-chip 4 x 4 binning. This operation mode is primarily used during autofocusing operations.

In order to minimise photobleaching, a high speed excitation shutter (type Uniblitz V24, Vincent Associates, Rochester, NY, USA) is fitted following the conventional mercury lamp burner (75 W). An in-house developed electronic control unit, interfaced to the host PC through a USB interface, is used to control the operation and timing of this shutter, as well as a variety of other microscope functions (objective turret, trans-illumination intensity etc.) used in other types of automated microscopy. The camera is operated in an externally triggered mode, whereby a feedback signal at the time that the shutter is fully open is used to initiate image integration on the CCD chip. Following a programmable integration time (typically 100 ms) the shutter is closed.

The comet slide is scanned in raster fashion in x and y directions using a motorised stage, building up a mosaic-like tiled image of the selected area. The stepper motor driven stage (type Scan IM, Märzhäuser Wetzlar GmbH & CO KG, Wetzlar-Steindorf, Germany) has a 200 nm resolution and its position is controlled by a three axis driver (type L-step, Lang GmbH & Co. KG, Huetttenberg, Germany), through a serial RS232 interface. The third axis is used to control a stepper motor (type MA-42, Märzhäuser GmbH, Wetzlar, Germany), mechanically coupled to the microscope's fine focus spindle. This arrangement is used for automatic focus operations.

The acquisition control software was developed in-house and assembled from software development tools provided by the commercial sources (Lang, Hamamatsu) and from in-house developed USB low-level routines. The complete software package (acquisition and analysis) is developed in 'C' code, running under the LabWindows/CVI™ development environment (National Instruments, Austin, Texas, USA), providing a standard graphical user interface. This allows the integration of image capture, image processing and the presentation of data from within one application. The software is executed on a conventional 800 MHz / 256 Mbyte RAM PC running Microsoft Windows 2000™. The complete system occupies approximately 1.5 m<sup>2</sup> of bench space, as shown in figure 2.

The analysis of a single image frame would restrict the number of comets that can be analysed, since they are randomly distributed, and a proportion would be present on the image frame edges. Conventional approaches would allow for a small overlap region somewhat greater than the

dimensions of the object of interest. Hence objects of interest found at the right hand edge, for example, would be ignored in the left edge frame of the next frame in the right direction. In the case of comets, which are relatively wide objects, this is an inefficient process, resulting in a high proportion of comets being exposed to the excitation light twice, potentially enhancing photobleaching. Instead, we adopt the following procedure, shown in figure 3: At the start of the scan, a single 'line' of adjacent image frames is captured but not analysed. After two frames of the next 'line', we construct a composite image of 2 x 2 raw image frames and identify all the comets. As the scan progresses, a moving 2 x 2 array is constructed and analysed, ignoring all comets already analysed. Although inevitable geometric image distortions are present, these have been found to have negligible effects on the results. It is of course essential to ensure (i) that camera rotation is such that it's imager is parallel to the stage motion and (ii) that the stage motion per frame is an accurate reflection of the imager's field of view, so that adjacent images are 'joined' accurately. Using this method, the larger, 2 x 2, images have an overlap of 50% yet each comet experiences just one exposure to the excitation light.

When scanning large areas, it is likely that the optimal focus position varies across the large scanned area of the sample due to misalignment of the sample within the microscope<sup>(23)</sup>. We overcome this problem by establishing a 'focus surface' at the start of the scan. We have found it adequate to define a plane surface from three points in the scanned area on which the comets are optimally focused. The objective position is thus adjusted at each image frame location during the scan. The initial focusing at the three locations may be performed manually, or more usually using an autofocus algorithm. We derive a focus indication signal based on the content of high spatial frequencies<sup>(17, 18)</sup>. During autofocus operations, fast image transfer is ensured by operating the camera in 8 bit intensity resolution mode and with on-chip 4 x 4 binning. The highest spatial frequency that can be derived is thus approximately  $0.3 \mu\text{m}^{-1}$ . Since it is likely that significant proportions of the image will contain no useful information, the most intense objects are identified. The criterion used is the selection of threshold level so that the interclass variance between dark and bright regions is maximized (inter-variance threshold)<sup>21, 22</sup>. A 64 x 64 pixel image is thus extracted with the object at the centre. A two-dimensional complex Fast Fourier Transform (FFT) is performed on this sub-image and the real part (square of magnitude) is used to generate an FFT image which is further segmented into components containing frequencies up to  $0.15 \mu\text{m}^{-1}$  and up to  $0.04 \mu\text{m}^{-1}$ . The ratio of the amplitudes of these band-limited components provides us with a normalised focus indication signal. These particular spatial frequencies are optimum for comets where the 'heads' have typically  $20 \mu\text{m}$  diameters. This focus indication signal is used as the error signal in a feedback loop which includes the z-axis drive motor. The loop iterations are performed using a Fibonacci optimisation algorithm<sup>(18, 19, 20)</sup> with a search range of +/-  $100 \mu\text{m}$  in 13 steps. The autofocus operations take approximately 3 seconds.

## 2.3 Image Processing

Reliable image processing starts with image normalisation, i.e. imager black/white level corrections and corrections to compensate for inevitable deviations from a 'flat' illumination field, due to the combined effects of the finite dimensions of the arc lamp source as well as slight, but noticeable vignetting in the microscope objective and optics. We have found it useful to use a 'standard' thin fluorescent sample of  $10 \mu\text{m}$  PMMA spun-coated on a standard coverslip, (supplied by Agar Scientific Ltd, Stanstead, Essex). The fluorescence image generated by this sample ( $I_{\text{white}}$ ) is used to normalise all subsequent comet images, following the usual subtraction of the camera's black level image ( $I_{\text{black}}$ ). However, inevitable small debris is likely to be present on the raw  $I_{\text{white}}$  image, and prior to normalisation, the image is passed through a 21 x 21 pixel wide median filter. Since it is

unlikely that there would be abrupt changes in illumination intensity, the use of such a filter is justified and is effective in preserving the broad features of the illumination field. This new filtered  $I_{\text{white}}$  image is then used to correct the comet image ( $I_{\text{comet}}$ ) such that:

$$\text{Normalised image} = [(I_{\text{comet}}) - (I_{\text{black}}) / (I_{\text{white}}) - (I_{\text{black}})] \times p$$

where  $p$  = peak intensity value of ( $I_{\text{white}}$ )

### 2.3.1 Region of Interest Definition

This next stage involves defining a region of interest around each comet. This is achieved by first thresholding the normalised image to identify all ‘bright’ objects, the threshold value being set to the intensity value corresponding to 20% of the peak frequency of the histogram; this assumes that the histogram peak corresponds to the background. This process segments whole comets, including both heads and tails. These are separated into individual comets by binary region growing. A rectangular region of interest (ROI) is then determined for each comet. This is typically around 10% larger in  $x$  and  $y$  than the maximum width and height of the binarised object. These ROI’s are used to perform a preliminary selection of candidate objects, according to user-set criteria, using prior knowledge about expected comet head sizes and approximate tail directions. It is of course possible that ROI’s of comets that are too close will overlap (Figure 4). If this overlap is unacceptable, both comets are rejected. The user can modify these criteria, but commonly the ROI is rejected if (i) it is too large, i.e. typically if  $x > 25$  times the expected head diameter (ii) it is not horizontal (i.e. if  $y > x$ ), (iii) if the overlap in *both*  $x$  and  $y$  of the ROI’s is greater than the expected head diameter, (iv) if  $y < 2$  times the expected head diameter and (v) if  $x < 4$  times the expected head diameter.

### 2.3.2 Background Removal

Background fluorescence may be present in the image due to auto-fluorescence, scattered excitation light, or from unwashed stain in the gel. This background fluorescence signal not only interferes with comet identification but also leads to erroneous measurements of comet moment and tail length. We have found it adequate in all cases to derive the average intensity of the pixels at the edge of the ROI and to subtract that value from the ROI ‘thumbnail’ image. This assumes that there is minimal change of excitation intensity within the ROI (i.e. that the image has been normalised correctly) and that the distribution of any unbound fluorophore across the ROI is essentially homogeneous.

If higher magnifications are warranted and the comet occupies a larger fraction of the image frame, or if normalisation cannot be performed, the background signal cannot be assumed to be ‘flat’ across the comet. In those cases, the background signal is estimated for each image independently from the image data surrounding the comets. A sparse data set, of typically 48 points, is sampled from the image on a regular grid. Each point comprises a 5 x 5 pixel box average. A general 2-dimensional quadratic function is fitted to this data set with the use of a general least-squares algorithm with singular value decomposition<sup>11</sup>. Data from the sparse set may, by chance, lay on a comet. To reduce the effect of this, the fit is recalculated with a revised data set in which these outlying points are removed. They are identified by the proportionately large error generated at these points when the data set is compared to the first fit. The final 2-dimensional quadratic fit to the data set is used to correct the original image by subtracting the corresponding fit value from each pixel.

### 2.3.3 Comet Head Delineation

The comet head delineation was achieved with the CHARM algorithm (Compact Hough And Radial Map). The compact Hough transform was recently used successfully in an automated colony counter<sup>(9)</sup> and has been reapplied here to find the approximate centre position of bright circular objects. The CHARM algorithm starts with the comet ROI and applies two perpendicular Sobel<sup>12</sup> edge-detection filters. A compact Hough transform is applied to the resulting edge maps: this gives a large response to bright circular objects. The pixel with the greatest response is taken as a point within the comet head. From this point, radial searching is performed to find likely object boundary points<sup>13, 14</sup> based on the response of the Sobel filters. The edges, identified by a significant Sobel response of the correct orientation, form a radial map that describes the head shape. Some filtering of this map is required to eliminate distraction from structure due to DNA clumping within the head. The 64-point maps are median filtered with a 7-point kernel and forced to be convex.

### 2.3.4 Comet rejection criteria

While most ‘unwanted’ objects are rejected according to previously described criteria, it is possible that comets are generated from multi-nucleated cells. Such cells are readily identified by the CHARM algorithm. Examples of the results of segmentation by the algorithm, and of the overall performance, are shown in figure 5.

## 2.4 Comet Assay Measurements

Several statistics are calculated for each comet, as defined below.

**Comet moment:** The comet profile is used to determine the ‘Kent’ comet moment<sup>(15)</sup> by binning pixel intensities across the comet. This is performed in two orthogonal directions, horizontal (x) and vertical (y), the moment results vector summed, to make the measurement independent of comet direction. The comet angle compared to the image is also determined.

$$\text{comet moment} = \frac{\int \{(\text{amount of DNA at distance } r) \times (\text{distance } r)\} dr}{\text{total DNA}}$$

The distance  $r$  is measured relative to the centre of mass of the head such that cells with undamaged DNA will have a comet moment close to zero. The centre of mass of the head is taken as the previously determined centroid of the mask<sup>(16)</sup> of the delineated comet head from the CHARM algorithm (Section 2.3.3). While most of the data presented use the Kent moment, the system also performs a similar vector sum to determine the ‘Olive’ moment, i.e. the product of the tail DNA content and the mean distance of migration in the tail.

**% DNA in tail and head:** The total fluorescence intensity within the delineated comet head is compared to the total comet intensity within the region of interest to determine the percentage of DNA in the head.

**Tail length:** The tail length has been defined as the distance between the centre of the comet head and the last non-zero pixel of the comet profile<sup>(2)</sup>. In our case the profile in the x-direction is used and corrected for comet angle.

The full analysis of a single comet requires approximately 1.7 seconds, including data collection in an Excel™ (Microsoft Corp., USA) table. The software allows for storage of comet thumbnail images and their display during the analysis procedure, as well as display of profiles. Examples of typical images are shown in figure 5. However, when performing an automated scan, the processing time is primarily a function of the number of frames, rather than the number of objects in a frame. The rejection of comets clearly influences the throughput, which is maximised when the seeding density is optimised. An average time per acquired 'valid' comet is thus typically <1.5 seconds, i.e. >2500 comets per hour is achievable. In practice, this figure is exceeded (even up to 3500 comets per hour when dealing with 'short' comets where the seeding density may be increased with reduced risk of overlap of comet ROI's).

### 3. Results

The validation of the automated system was performed with reference to an existing, manually operated system, based on a CCIR video rate camera. The resolution of this camera was lower and the more usual, higher magnifications were employed (typically 60x). Summary data comparing systems are shown in figure 6.

Both the automated and manual systems show similar, linear responses for the tail moment. The two comet tail parameters of length (figure 6b) and percentage of DNA in the tail (figure 6c) showed evident dose-dependent responses. However, the use of the Kent moment (figure 6a) is confirmed to be a more superior indicator than tail length or percentage DNA in the tail. The measurement of total comet DNA (head + tail), (figure 6d) is expected to be independent of radiation dose at doses above 0.2 Gy<sup>(2,6)</sup>. This is observed by the more sensitive automated system, while the manual system is clearly more variable.

The ability to collect information from very large numbers of cells allows us to discriminate between subpopulations. Examples of this include differentially damaged subpopulations (e.g. drug resistant, radiation resistant or apoptotic fractions of tumours). In figure 7, we present the application of the automated system to cell-cycle discriminations. This discrimination may be based both on DNA content and on the behaviour of replicating DNA during electrophoresis. It has been recognised that S-phase replicating DNA is retarded under neutral conditions due to the occurrence of replication forks<sup>(2,4)</sup> while the single strand breaks at these forks are observed as increased background damage. This is clearly shown by the clustering of G1, S, G2 phases in figure 7b. The discrimination of cell cycle phase based on both DNA content and tail moment allows us to produce separate dose response curves (figure 7c) for the different phases. Both figure 7a and 7c demonstrate the independence of both single- and double-strand break induction for cell-cycle phase.

In the case of alkaline lysis (used as a measure of DNA single-strand breaks), the dose response is essentially linear above 0.1 Gy. Under neutral conditions (used as a measure of DNA double-strand breaks), the initial slope (<5 Gy) is in fact likely to be caused by the relaxation of supercoiled DNA by single-strand breaks<sup>(3)</sup>. At doses >5Gy, the dose response is essentially linear. The ability to resolve this in detail and with a high degree of confidence is a direct consequence of processing numbers of cells greater than those commonly used with manual assays.

## **4. Conclusion**

We have demonstrated that a fully automated system, where image analysis is seamlessly combined with image capture and sample positioning is capable of very fast throughput. The ability to analyse many thousands of comets in a short time is particularly useful in minimising biological noise, thereby extending the applicability of this assay to a range of biological investigations where cell-cycle effects or sub-population discrimination are important. Furthermore, the system removes inevitable subjectivity and operator fatigue.

A secondary advantage of the automated approach, not exploited here, is that the spatial coordinates of each comet are known. It is therefore straightforward to re-examine an individual comet should it be identified by an unusual profile or other measure from e.g. population analyses. The ability to revisit and re-image the comets is afforded by the use of a minimum level of excitation, through the use of an excitation shutter and of a sensitive, wide dynamic range imager. While photo-bleaching is conventionally minimised by the use of a high magnification objective, we have taken the approach of using a low power objective, with a numerical aperture high enough to resolve comet structure. It is unfortunate that even lower power lenses (e.g. 4x) of appropriately high numerical aperture are not readily available; the use of such lenses would further enhance throughput by reducing the number of stage movements (i.e. image frames) needed to quantify a given area. The data presented are obtained using a commonly available intercalating dye (SYBR-Gold), but it is expected that the use of intercalating dyes more specific to single-strand breaks and double-strand breaks would enhance the sensitivity and specificity even further. In conclusion, we have shown that the use of optimised software, in combination with a readily available imaging platform and appropriate optics results in a high throughput capable of quantifying DNA damage in a statistically unbiased manner. Such an approach increases the applicability of this type of assay to a range of biological and radiobiological questions, particularly those associated with quantifying sub-populations.

## **5. Acknowledgements**

We acknowledge the financial support of CRUK, Programme Grant C133/A1812 – SP 2195-01/02 and the US Department of Energy Low Dose Radiation Research Program grant DE-FG07-99ER62878. It is a pleasure to acknowledge the technical support of Robert Newman and John Prentice for electronic and mechanical workshop assistance, of Tim Idzenga who patiently optimised autofocus algorithms and Julian Gilbey for assistance with the development of a fast 2-D median filter algorithm.

## References

1. Olive, P. L. Overnight lysis improves the efficiency of DNA damage detection in the alkaline comet assay. *Radiat Res* **155**, 564-571, 2001.
2. Olive, P. L., Banath, J. P., and Durand, R. E. Heterogeneity in radiation-induced DNA damage and repair in tumor and normal cells measured using the "comet" assay. *Radiat Res*, **122**: 86-94, 1990.
3. Olive, P. L., Durand, R. E., Banath, J. P., and Johnston, P. J. Analysis of DNA damage in individual cells. *Methods Cell Biol*, **64**: 235-249., 2001.
4. Olive, P. L. DNA damage and repair in individual cells: applications of the comet assay in radiobiology. *Int J Radiat Biol*, **75**: 395-405., 1999.
5. Olive, P. L., Johnston, P. J., Banath, J. P., and Durand, R. E. The comet assay: a new method to examine heterogeneity associated with solid tumors. *Nat Med*, **4**: 103-105, 1998.
6. Fairbairn, D. W., Olive, P. L., and O'Neill, K. L. The comet assay: a comprehensive review. *Mutat Res*, **339**: 37-59, 1995.
7. Friauff, W., A. Hartmann, et al. (2001). "Automatic analysis of slides processed in the comet assay." *Mutagenesis* **16**(2): 133-137.
8. Bocker, W., W. Rolf, et al. (1999). "Automated comet assay analysis." *Cytometry* **35**: 134-144.
9. Barber, P. R., B. Vojnovic, et al. (2001). "Automated Counting of Mammalian Cell Colonies." *Physics in Medicine and Biology* **46**(1): 63-76.
10. Kent, C. R. H., J. J. Eady, et al. (1995). "The comet moment as a measure of DNA damage in the comet assay." *International Journal of Radiation Biology* **67**(6): 655-660.
11. Press, W.H., "Numerical Recipes in Fortran" 2<sup>nd</sup> Ed., Chapter 15, ISBN 052143064X.
12. Sobel, I. (1990) "An isotropic 3 x 3 image gradient operator." In H. Freeman, editor, *Machine Vision for Three-Dimensional Scenes*. Academic Press.
13. Mouroutis, T., S. J. Roberts, et al. (1998). "Robust cell nuclei segmentation using statistical modelling." *Bioimaging* **6**: 79-91
14. Barber, P. R., B. Vojnovic, et al. (2000). An automated colony counter utilising a compact Hough transform. *Proceedings of Medical Image Understanding and Analysis, MIUA2000*, London, UK, UCL, London.
15. Kent, C. R. H., J. J. Eady, et al. (1995). "The comet moment as a measure of DNA damage in the comet assay." *International Journal of Radiation Biology* **67**(6): 655-660.
16. Russ J.C., "The Image Processing Handbook", 3<sup>rd</sup> Ed. ISBN 0849325323, p. 517.
17. M. Subbarao, and J. L. Tyan, "The Optimal Focus Measures for Passive Autofocusing and Depth-From-Focus", *Proceedings of SPIE Conference on Videometrics IV*, Philadelphia, Oct. 1995
18. M. Subbarao, T. Choi, A. Nikzad, "Focusing Techniques", *Optical Engineering*, Nov. 1993, Vol. 32, No. 11, pp. 2824 - 2836
19. F. R. Boddeke, L. J. van Vliet, H. Netten and I. T. Young, "Autofocusing in Microscopy Based on the OTF and Sampling", *BioImaging*, Vol. 2, No. 4, pp. 193 - 203
20. J. M. Geusebroek, F. Cornelissen, A. W. M. Smeulders and H. Geerts, "Robust Autofocusing in Microscopy", *Cytometry*, 2000, Vol. 36, No. 1, pp. 1 - 9
21. P. K. Sahoo, S. Soltani, and A. K. C. Wong, A survey of thresholding techniques, *Comput. Vision Graphics Image Process.* **41**, 1988, 233-260.
22. Papamarkos, N. and Atsalakis, A. (2000) Gray-Level Reduction Using Local Spatial Features. *Computer Vision and Image Understanding* **78**, 336350.
23. Bravo-Zanoguera, M, 1998. *Rev. Sci. Instrum.* **69**(11), 3966-3977.

Figure 1: Three examples of 'difficult' comet images. The white outline shows the result of head segmentation by the CHARM algorithm.

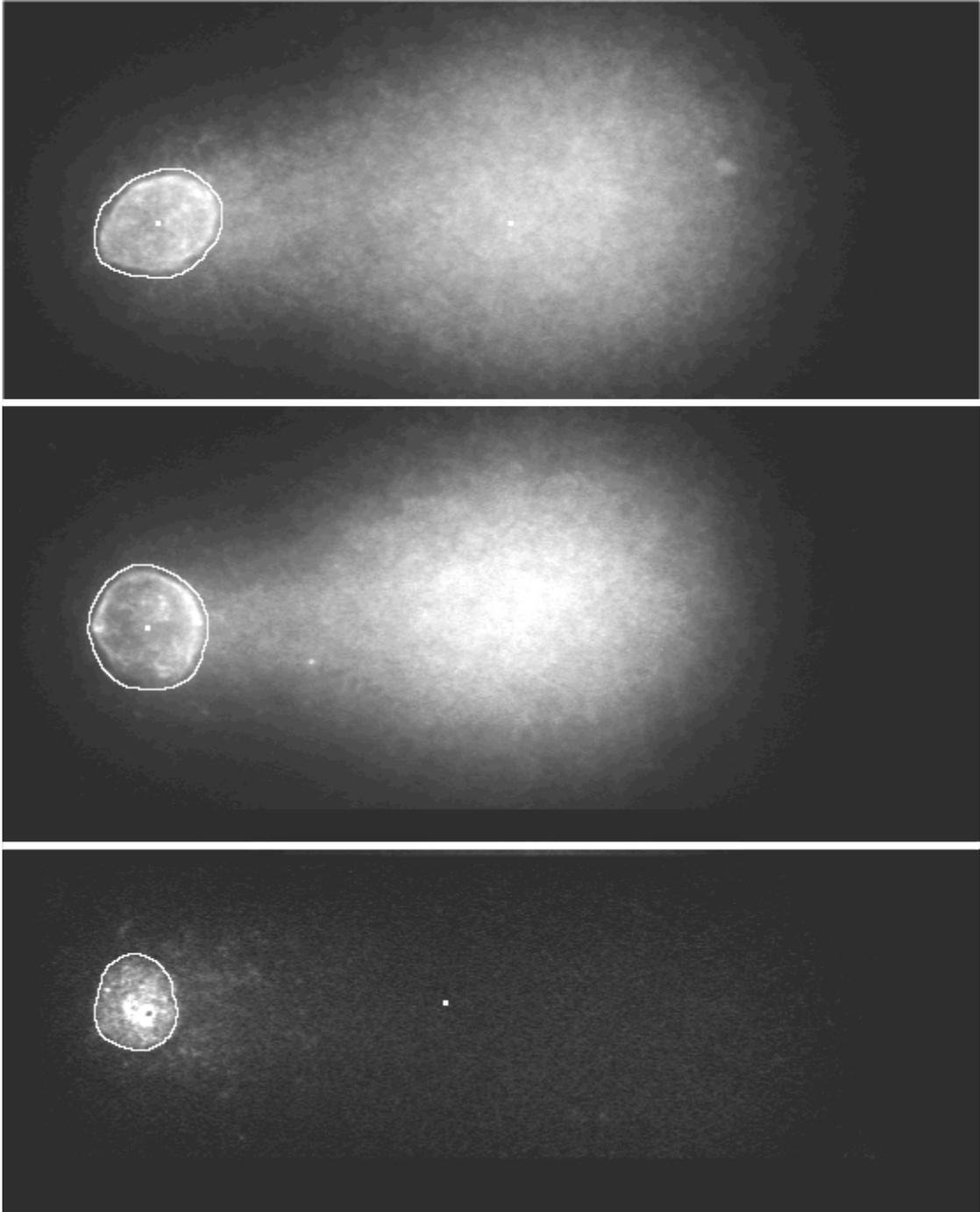


Figure 2: The complete automated comet analysis system. The inset shows the user interface.

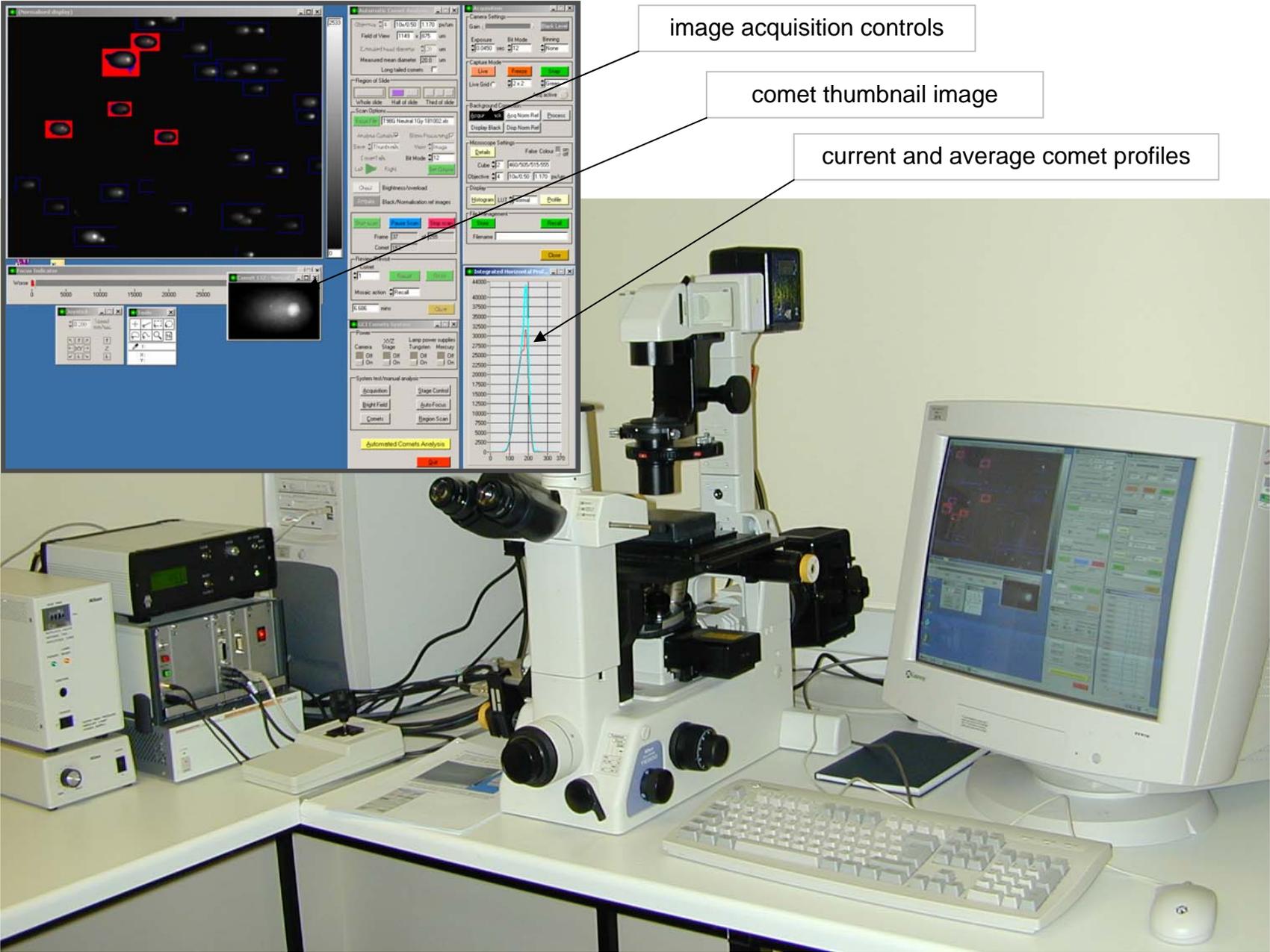


Figure 3: Stage scanning sequence: the slide is scanned in a raster-like pattern (panel (1)), with analysis starting only after the second line of frames (panel (2)). All comets (identified as A) completely enclosed by the first 2 x 2 image block are found. Comets identified as (B), straddling the adjacent 2 x 2 block are found next, while comets (C) are ignored. This process is repeated until the full area of interest is scanned.

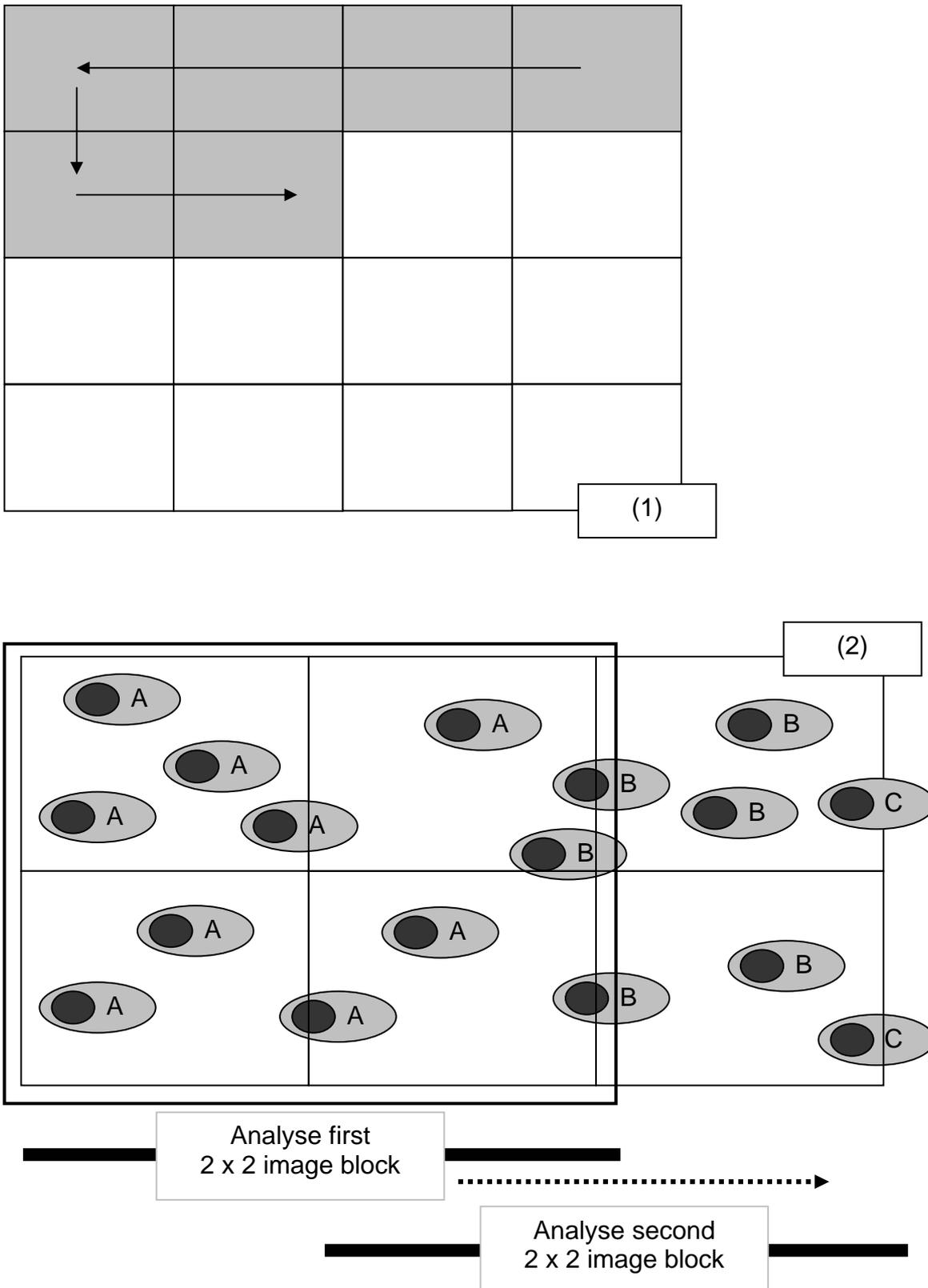


Figure 4: Criteria for rejecting comets which are too close from the automatically placed regions of interest (ROI). Rejection of both comets is invoked when the ROI overlap exceeds criteria described in the text.

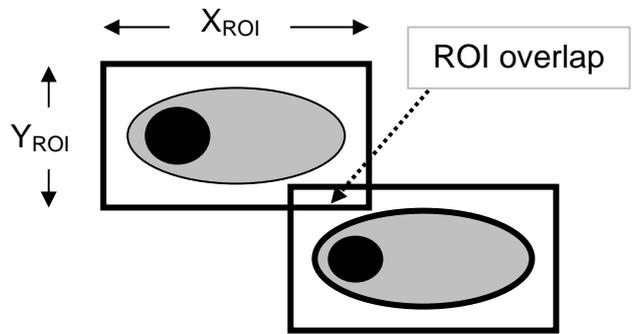


Figure 5: Typical results obtained with automated system. Panel (a) shows the composite image of the central third of the sample slide, consisting of 17 x 10 raw image frames. Panel (b) shows one of those frames with a linear intensity scale. Panel (c) is the same as (b), but through a logarithmic display look-up table, to enhance the visibility of the weak comet tails. Panels (d) and (e) show two correctly identified examples of bi-nucleated cells. Panel (f) shows a delineated comet head and its centre and the centre of gravity of the tail, i.e. with the head 'removed'. The profiles shown below panels (e) and (f) are used to calculate DNA content. The profile below (e) indicates the difficulty in resolving bi-nucleated cells solely on the basis of profile.

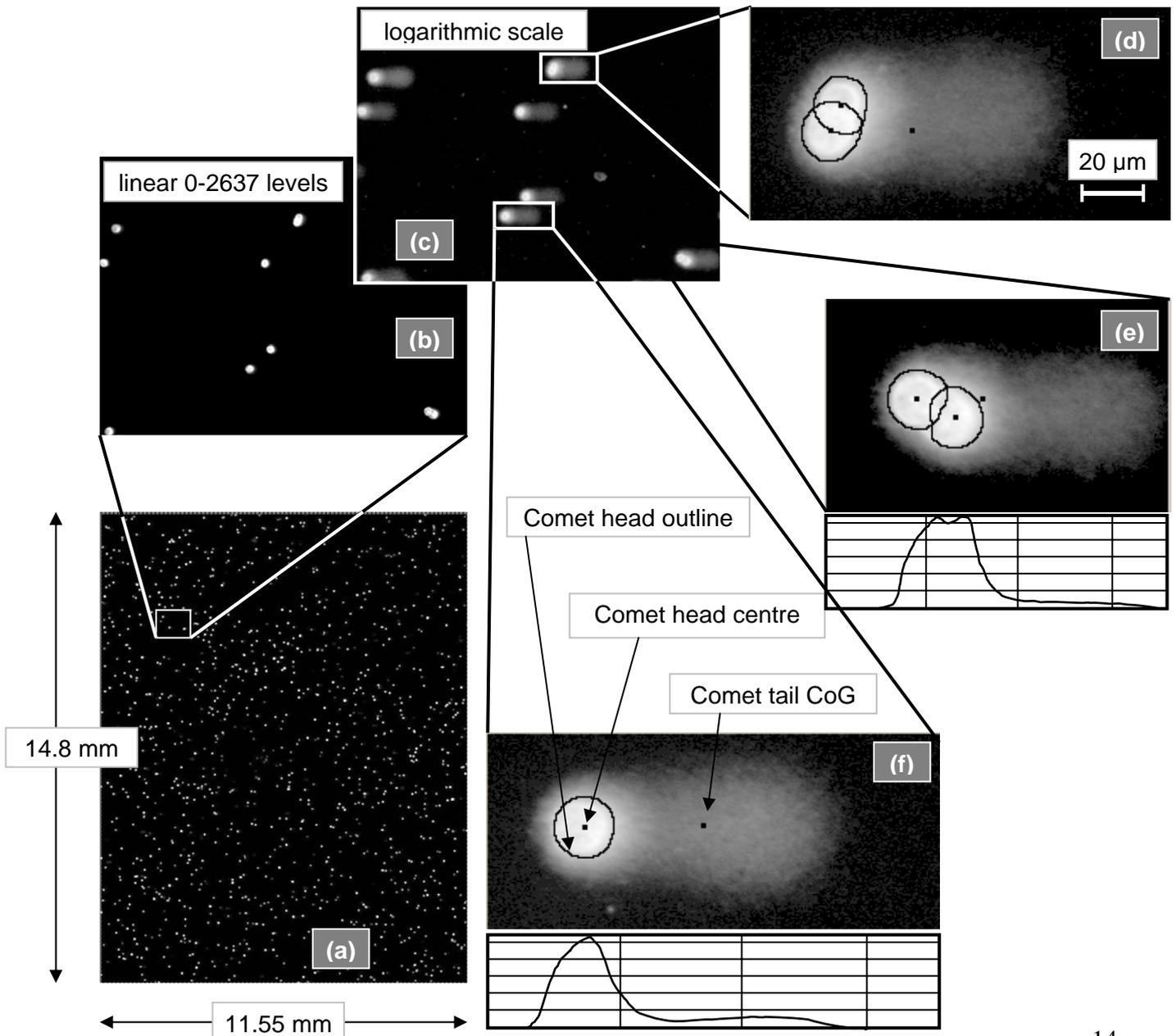


Figure 6: Pooled data from 3 experiments of T98G glioblastoma cells lysed and electrophoresed under alkaline conditions. Manual analysis was performed on 3 x 100 cells, while 3 x 600 cells were processed automatically. Inter sample variations observed in panels (b), (c) and (d) lead to systematic errors which are compensated by the use of the Kent tail moment, shown in the dose response shown in (a). The improved consistency of the automated system (panel (d)) is demonstrated by the flat graph of fluorescence intensity versus dose.

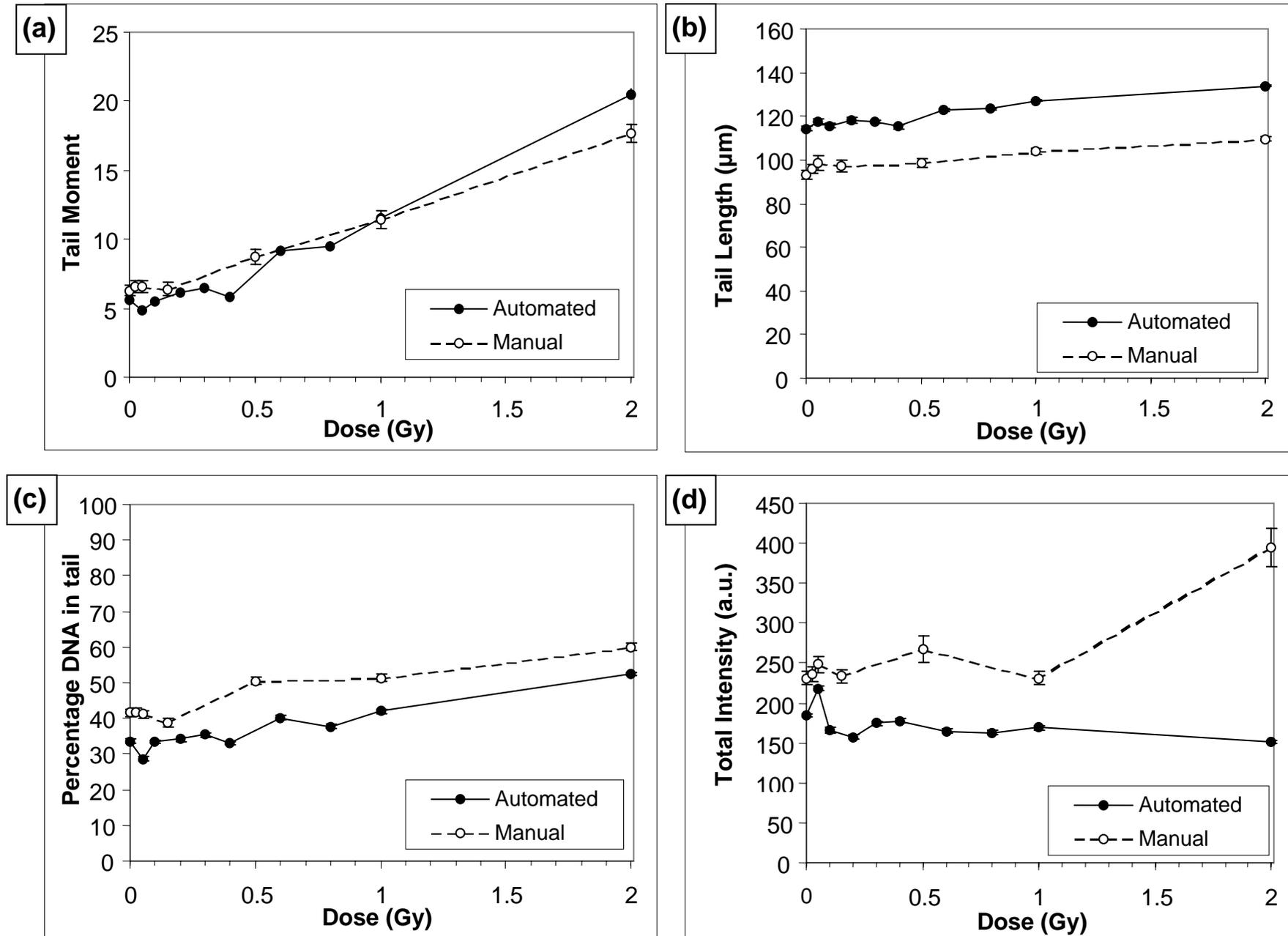


Figure 7. T98G glioblastoma cells were embedded in agarose, irradiated and lysed under neutral or alkaline conditions as described above. 600-1000 comets were scored for each dose point. Panel a) shows the density distribution of tail moment measurements versus the DNA content (proportional to fluorescence intensity) after differing doses. Panel b) shows the distribution of Tail moment and DNA content after 1 Gy (alkaline) or 2 Gy (neutral) X-rays with circles representing the position of cells in the G<sub>1</sub>, S and G<sub>2</sub> phases of the cell cycle. Note that the migration of DNA in S-phase cells is retarded under neutral and enhanced under alkaline lysis conditions. Panel c) mean tail moment of cells in G<sub>1</sub>, S, G<sub>2</sub> and the total population of cells. Discrimination of cell cycle phase was by K-means iterative cluster analysis (JMP statistical discovery software, SAS Institute NC, USA). Error bars represent the 99% confidence intervals. Tukey-Kramer HSD comparison of the means for the total population indicates significant differences at the 99% confidence level between all doses after neutral lysis and at doses > 0.2 Gy and the control under alkaline conditions.

