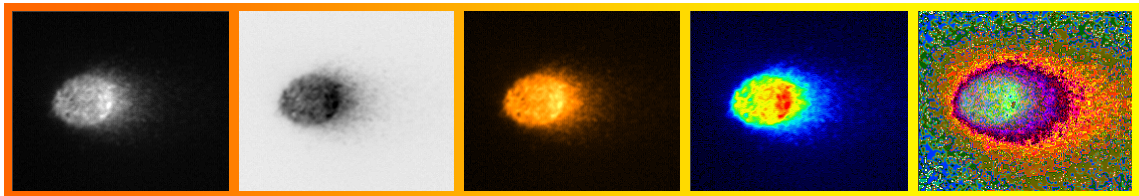


# Help!

## There is a comet in my computer!



**A dummy's guide to image analysis  
used in the comet assay**

**Barbara Vilhar**

**University of Ljubljana  
Biotechnical Faculty  
Department of Biology**

# Contents

<b>Do you have a problem?</b>	<b>1</b>
<b>1 Fluorescence microscopy</b>	<b>3</b>
<b>2 Detection of emitted light</b>	<b>7</b>
<b>3 Measurement of the comet parameters</b>	<b>9</b>
3.1 Grey values	9
3.2 Segmentation	10
3.3 Geometric calibration	11
3.4 Measurement of geometric parameters	12
3.4.1 And I measured my comet at 400× magnification	13
3.5 Measurement of light intensity parameters	14
3.5.1 Integrated light intensity	14
3.5.2 Percentage of DNA in the tail	17
3.5.3 Tail extent moment	17
3.5.4 Comet profile	18
3.5.5 Olive tail moment	19
<b>4 Recorded image and displayed image</b>	<b>21</b>
4.1 Look-up table (LUT)	21
4.2 Live image and grabbed image	23
<b>5 Image acquisition</b>	<b>26</b>
5.1 Histogram of image grey values	27
5.2 Image contrast and brightness	27
5.3 Image saturation	30
5.4 How can I evaluate the quality of my image?	32
5.5 Shading	35
5.6 How can I detect shading in my image?	36
5.7 Why a 10-bit camera beats an 8-bit camera	38
5.8 Binning	40
5.9 Which image does my software measure?	41
5.10 How do I get a good quality image out of my comet equipment?	42
<b>6 Concluding remarks</b>	<b>45</b>
<b>Notes</b>	<b>46</b>

# Figures

<b>Figure 1.</b> The principle of fluorescence and absorption and emission spectra of DNA-bound ethidium bromide .....	4
<b>Figure 2.</b> The path of light through a microscope during observation of fluorescence of ethidium bromide .....	6
<b>Figure 3.</b> Light emitted from a comet on a slide is detected as an image .....	8
<b>Figure 4.</b> Information on an image is coded as grey values .....	10
<b>Figure 5.</b> Segmentation of the comet image.....	11
<b>Figure 6.</b> Geometric calibration of the image analysis system. ....	12
<b>Figure 7.</b> Scaling of a printed image. Printed images should contain a scale bar.....	13
<b>Figure 8.</b> Grey values for the head, the tail and the background region of the model comet image .....	15
<b>Figure 9.</b> The comet profile for a comet recorded from a microscope slide .....	18
<b>Figure 10.</b> The comet profile of the model comet image .....	19
<b>Figure 11.</b> The intensity centroids .....	20
<b>Figure 12.</b> The grabbed and the displayed image .....	22
<b>Figure 13.</b> Types of images with respect to colour.....	22
<b>Figure 14.</b> Look-up tables (LUTs).....	25
<b>Figure 15.</b> Effects of image acquisition settings on the information contained in an image of a comet.....	26
<b>Figure 16.</b> The histogram of image grey values .....	27
<b>Figure 17.</b> Image contrast and brightness.....	29
<b>Figure 18.</b> Image saturation.....	31
<b>Figure 19.</b> Viewing image histograms and grey values in <i>Photoshop</i> .....	33
<b>Figure 20.</b> Examples of indicators of the pixel grey value .....	34
<b>Figure 21.</b> Effects of uneven illumination of the field of view .....	35
<b>Figure 22.</b> Detection of image shading in <i>Photoshop</i> .....	37
<b>Figure 23.</b> Comparison of an 8-bit and a 10-bit camera.....	39
<b>Figure 24.</b> The effects of binning .....	41
<b>Figure 25.</b> Different versions of user interface for adjusting the camera settings .....	43
<b>Figure 26.</b> Examples of contrast and brightness settings of the displayed image .....	44
<b>Figure 27.</b> Measurement of the comet parameters: an overview .....	45

## Do you have a problem?

Here I will talk about the basic principles of image analysis used in the comet assay. This should help you spot and solve potential problems in your measurements, avoid measurement errors, understand the settings available in software packages and the language used in user manuals, and improve your communication with the software companies when you run into trouble. The rest of this chapter is crap and you should move to the next one or quit reading. Now, this is how all the problems began.

Not so long ago, my good friend Prof. Lite Microscope was elegantly standing on the real desktop in our lab. He lived a happy and respectful life. A fine screwdriver and a drop of oil fixed all his troubles, although he hardly had any. These days, he is choking under the weight of digital cameras, ambushed by monitors, keyboards, graphics tables and printers. An entanglement of cables ties him to his new master, Mr. P.C. Computer. The life of imprisoned Prof. Microscope has moved to the virtual desktop. Every day he is humiliated by software crashes, electronic mice, digital bugs, digital worms and digital viruses. The only one that ever looks down his shiny objectives is Ms. Camera Digital. His newly acquired, chip-controlled motors are desperately humming. Prof. Microscope suffers, and I suffer with him. I have no idea what the mean Mr. Computer might be doing to the digital results of my real experiments. I just sit in the dark for hours on end and click. Not so long ago, I used to understand my science.

Does this sound familiar to you? Do you actually know how the computer calculates the parameters of your comet assay? Do you understand the instructions and explanations in the handbook that came with your comet software? And is it important to know this at all?

Judge for yourself. I bought expensive software from a well-respected company, ran the computer and started clicking. I noticed straight away a problem with my data – the graphs that I plotted were somewhat, but not completely, in disagreement with previously published results from other labs. However, my results had remarkably small standard errors, so I concluded that the computer was doing a very fine job. I blamed the new method for staining of the slides that I used, and spent several months trying to perfect my staining skills. The strangest thing was that whatever I did, my results were always very similar, and always in disagreement with other labs. In the end I realised that it was not I, it was the computer that spoiled my results. So I sat down, learned the basics of image analysis and analysed the software algorithm step by step. In the end I found out that the software was surely measuring *something*, but that this something was not at all a meaningful parameter. The bad news was that the software could not be easily fixed – the whole measurement procedure was based on erroneous assumptions and at odds with the laws of physics. The next thing I did was learning how to program the computer for image analysis.

An annoying problem with computers is that, unless they crash, they will always give you a result. The above story is true, but it is not about the comet assay. Nevertheless, it is about measurement of the amount of DNA in nuclei on a microscope slide (to determine genome size in different organisms). I have to admit here that I have never prepared one single microscope slide for the comet assay. But I recently tried to find out how the comet assay works and I talked to researchers and students from several comet labs. While I got excellent explanations about DNA damage and variations in the procedure for preparation of the slides, the image analysis part of the experiment was more obscure. In most cases, the researchers bought the software and hardware, and then started to click. They had no time to study image analysis. Still, it is not really sensible to treat the only part of the experiment in which parameters are measured as a tightly closed “black box”.

Here are some arguments from the comet world. A colleague called me on the phone one day saying that she noticed abnormal values for the comet parameters in recent experiments. None of the researchers that spent days and months with the comet machine knew what went wrong. I only saw this machine once for an hour or so. Nevertheless, I was able to diagnose the problem and also propose the solution after only a few minutes of phone conversation. How did I do it? Well, I previously borrowed the software user manual, and because I know the principles of image analysis, I also understood how the black box works.

On another occasion, I had the idea to write a macro program for the comet assay. So I borrowed a CD with comet images from another colleague, who was then a comet beginner using another comet software package than the one mentioned above. When I viewed the images on my computer, I realised that there was something wrong with image contrast. In addition, quite a few spots on the image were overexposed, which means completely white. With such images, you do measure *something* with your computer, but your results are either imprecise or wrong. Optimising the settings of the camera solved the problem.

Now let us open the “black box”.

# 1 Fluorescence microscopy

We are running a comet experiment. We prepared the slides and ran electrophoresis. Now we have to measure the amount of DNA that migrated out of nuclei. The most important thing to remember is this: **Measurement is based on a known relationship** (e.g. linear, logarithmic) **between the amount of DNA in the sample and the parameter used for detection of DNA**. In all steps of the measurement procedure, we have to take care **not to compromise this relationship**.

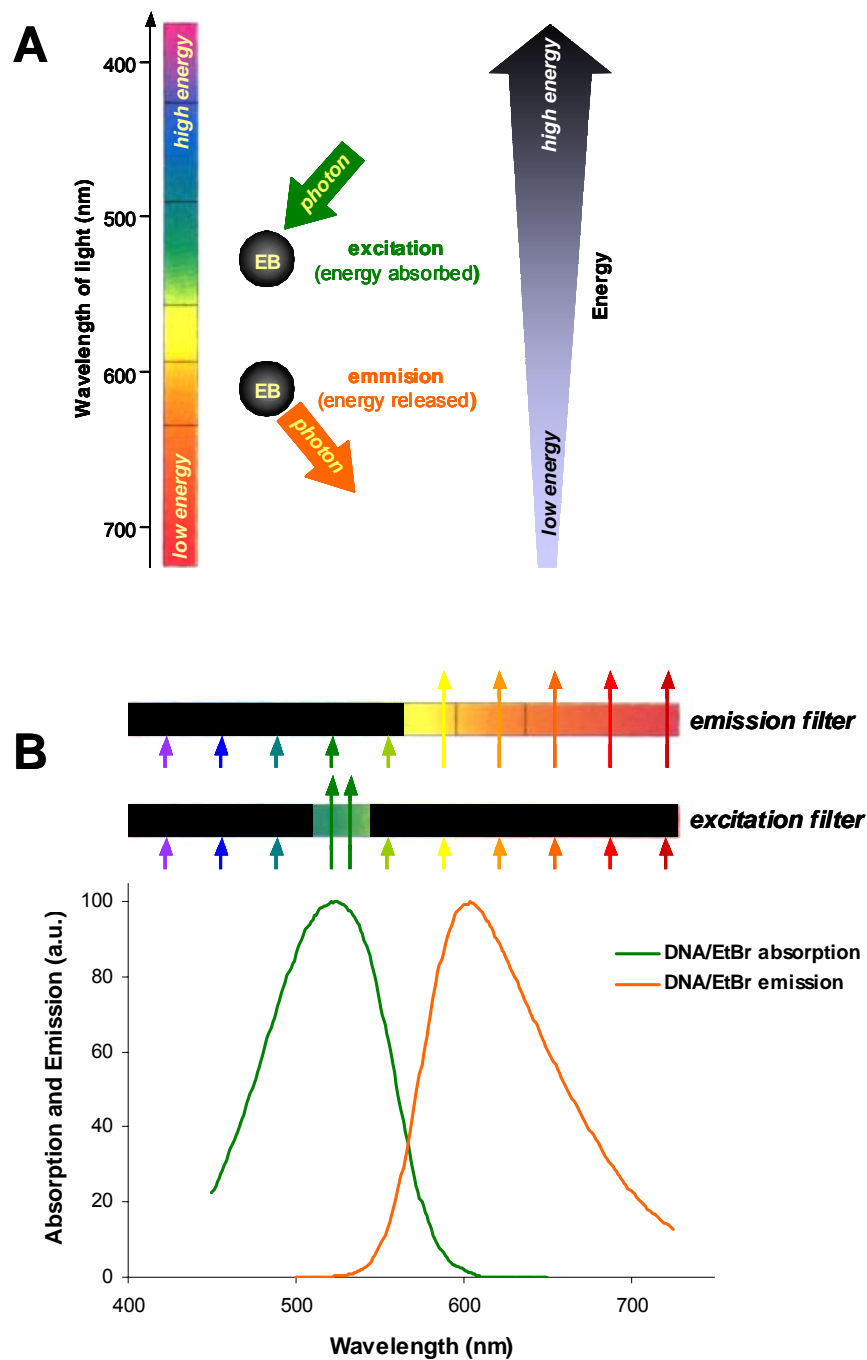
The first thing we need to do is stain DNA to make it detectable. Staining has to be quantitative: the amount of stain that binds to DNA must be stoichiometrically proportional to the amount of DNA present in the sample. In other words, **the relationship between the amount of DNA and the amount of the staining substance bound to DNA must be linear**.

The quantitative DNA dye usually used in the comet assay is ethidium bromide. During incubation in the staining solution, ethidium bromide binds to DNA on the comet slide. Ethidium bromide is an intercalator with little or no sequence preference and binds to DNA at stoichiometry of one molecule per 4-5 base pairs of DNA. Hence, **the relationship between the amount of DNA and the amount of ethidium bromide bound to DNA is linear**.

Ethidium bromide is a molecule with special properties - a fluorophore. **Fluorophores** are molecules that absorb the energy of light (photons) of specific wavelengths. This process is called **excitation**. Fluorophores then release the absorbed energy in the form of light energy (photons). This process is called **emission**. Because some energy is “lost” during the excitation-emission cycle, the energy of emitted photons is lower than the energy of absorbed photons. In other words, the wavelength of emitted light is longer than the wavelength of absorbed light. **Fluorescence** is thus emission of light or other electromagnetic radiation at longer wavelengths by matter as a result of absorption of shorter wavelengths. Emission lasts only as long as the stimulating irradiation is present.

Each fluorophore absorbs and emits light of specific wavelengths. For example, DNA-bound ethidium bromide absorbs green light and emits orange light (Figure 1A). Once ethidium bromide is bound to DNA, its fluorescence increases about 10-fold (compared to free molecules).

To be able to see the comets under a microscope, we need to illuminate the ethidium bromide-stained slide with intense green light. The source of high intensity light is usually a mercury lamp. Because a mercury lamp emits light of different wavelengths (including UV), we need to place an **excitation filter** between the lamp and the slide. This filter only passes green light and blocks all other wavelengths (Figures 1, 2A).



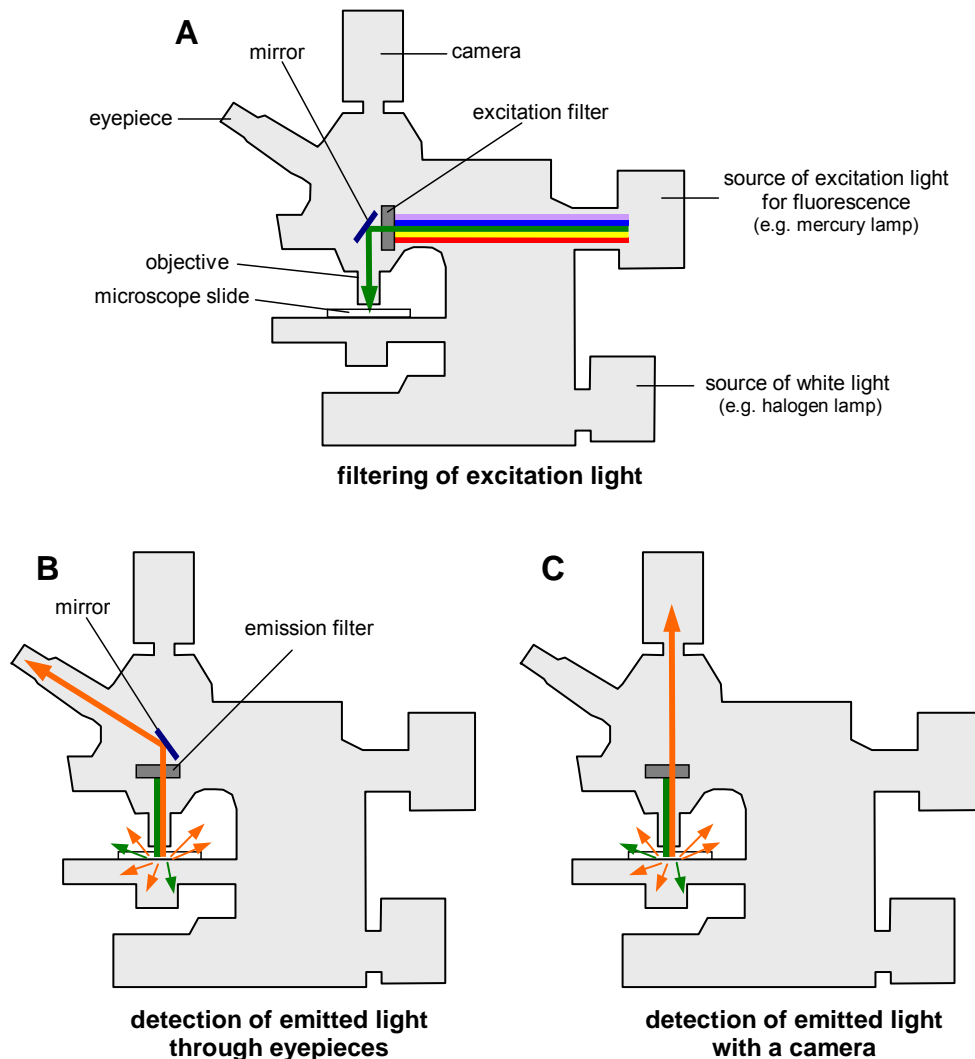
**Figure 1.** The principle of fluorescence and absorption and emission spectra of DNA-bound ethidium bromide. **A:** DNA-bound ethidium bromide (EB) absorbs green light (high energy) and emits orange light (low energy). **B:** DNA-bound ethidium bromide absorbs green light (peak at ~530 nm) and emits orange light (peak at ~610 nm). The excitation filter on the microscope passes green light and blocks all other wavelengths. The emission filter passes wavelengths longer than the excitation light wavelength (orange and red). Note that ethidium bromide has another absorption peak in the UV range of the light spectrum (not shown) – this peak is exploited for visualisation of DNA on agarose gels with a UV transilluminator. See Figure 2 for the position of filters in the microscope light path.

When a slide is illuminated with green light, DNA-bound ethidium bromide on the slide emits orange light in all directions (Figures 2B, 2C). During observation of an object through eyepieces of a microscope, we see only light that is collected by the objective. We now have two problems: the emitted light is very weak and some of the strong excitation green light is reflected from the slide. To prevent excitation light disturbing our observation of fluorescence, we need to place an **emission filter** between the object and the eyepieces. This filter only passes light with longer wavelengths than the excitation light. The emission filter for ethidium bromide passes orange light and blocks green light (Figures 1, 2B, 2C). The second problem is easier to solve: because the intensity of the emitted orange light is very weak compared to the intensity of daylight, we have to observe the slide in a dark room.

Both the excitation and the emission filter are usually mounted together in a cubical filter holder, which is typically located in a filter wheel just above the objective revolver. Each fluorophore requires different combinations of excitation and emission filters to match its specific absorption and emission properties.

Let us return to our task to measure the amount of DNA on the slide. We already know that the relationship between the amount of DNA and the amount of ethidium bromide bound to DNA is linear. The good news is that also **the relationship between the amount of DNA-bound ethidium bromide and the intensity of emitted light (the number of emitted photons) is linear**. We can thus conclude that **the relationship between the amount of DNA and the intensity of emitted orange light is linear**. So all we need to do now is to somehow measure the amount of emitted light, and we will know how much DNA there is on the slide. This is where we need to start talking about images.





**Figure 2.** The path of light through a microscope during observation of fluorescence of ethidium bromide. A typical microscope setup is shown, where the excitation light is passed to the slide through an objective. For ethidium bromide, excitation light is green, emission light is orange. The slide is illuminated by light from a mercury lamp (**A**). A mercury lamp emits light of different wavelengths across the spectrum of visible and UV light. Only green light passes through the excitation filter. DNA-bound ethidium bromide is excited by the green light and emits weak orange light, which can be detected either by looking through the eyepieces (**B**) or by a camera (**C**). Because relatively much of the excitation green light is reflected in all directions, an emission filter is used that absorbs green light but passes light with higher wavelengths. See also Figure 1B for explanation about the properties of filters.

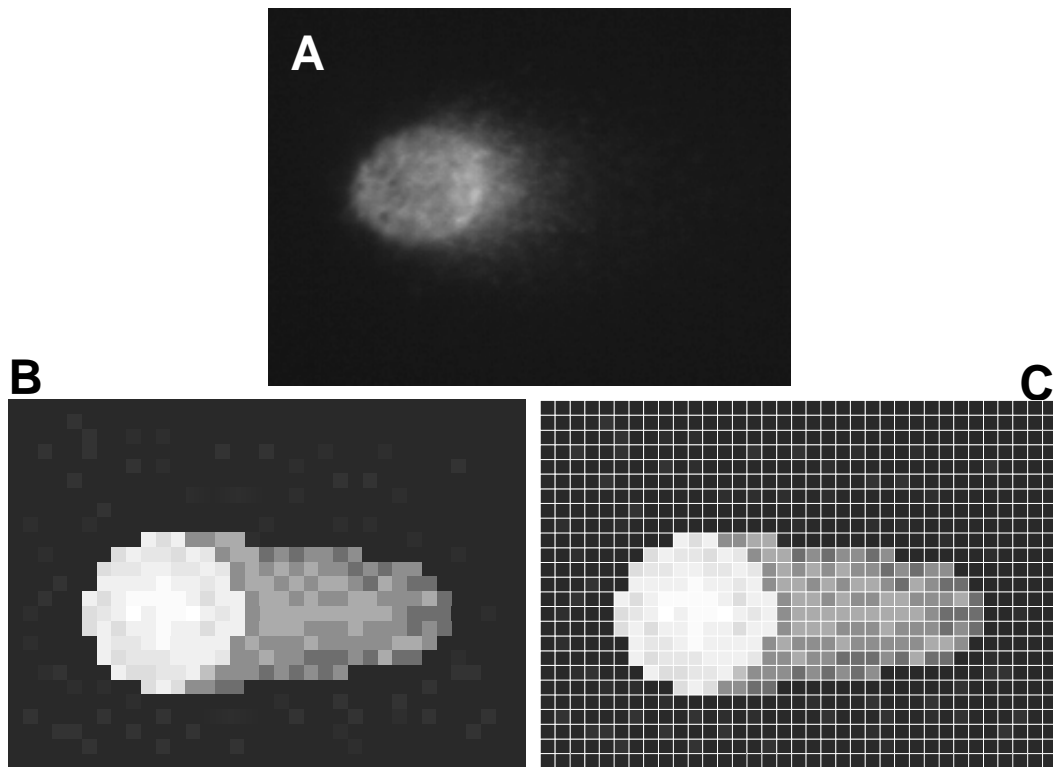
## 2 Detection of emitted light

Our eyes, supported by our brain, are amazing instruments. We process incredible amounts of visual information and store a lot of what we see in our memory. We can see very small details. We can focus to near and far objects. There is one problem with our eyes, though – we are not very good at judging the quantitative aspects of all the information we receive. By only looking, we cannot tell whether two similar pencils located at different parts of a table are of the same length or not, or how many centimetres they have. We are even worse at measuring light intensities with our eyes. When we walk into a dimly-lit room from sunshine, it seems to be very dark. A few moments later, we see perfectly well. When observing a comet slide under a microscope, we cannot say how much light comes from the head and how much from the tail. Our eyes are also not sensitive enough to see very weak light.

We need to use technology for precise measurement of the intensity of light emitted from comets. We can direct the emitted light through a microscope to a device that can detect light intensity - **a camera** can do this (Figure 2C). A detector located in the camera is composed of a rectangular array of small light-sensitive elements. The size of each detector element is only a few micrometers across. Each element converts the measured light intensity to a digital format, i.e. to numbers. By definition, small numbers represent a low light intensity (few photons hit the detecting element), and high numbers a high light intensity (many photons hit the detecting element). If we now connect the camera to a computer and record the image, the camera sends the numbers representing each detector element together with its coordinates on the detection array to the computer. The camera sends all numbers representing one image as a package of digital data (an image file). The computer decodes the received digital (numerical) information and displays it as an image on the monitor screen (Figure 3A).

An image is composed of small dots called pixels. **Each pixel represents one light-sensitive element of the camera, and the numbers corresponding to light intensity detected for each pixel are stored in a computer file.**

We will use a simplified model image of a comet (Figures 3B, 3C) to explain how the comet intensity is measured. This image was generated artificially and has extremely large pixels for easier explanation of the discussed procedures.



**Figure 3.** Light emitted from a comet on a slide is detected as an image. An image of a real comet on a microscope slide is shown in **A**. A simplified model image of the comet (**B**) will be used to demonstrate measurement of the comet parameters. An image is composed of separate pixels (**C**). The size of this model image of the comet is 35 pixels  $\times$  25 pixels. Images that cameras record have many pixels (e.g. 750  $\times$  550, 1300  $\times$  1030), so individual pixels cannot be discerned in **A** unless the image is considerably zoomed up.

## 3 Measurement of the comet parameters

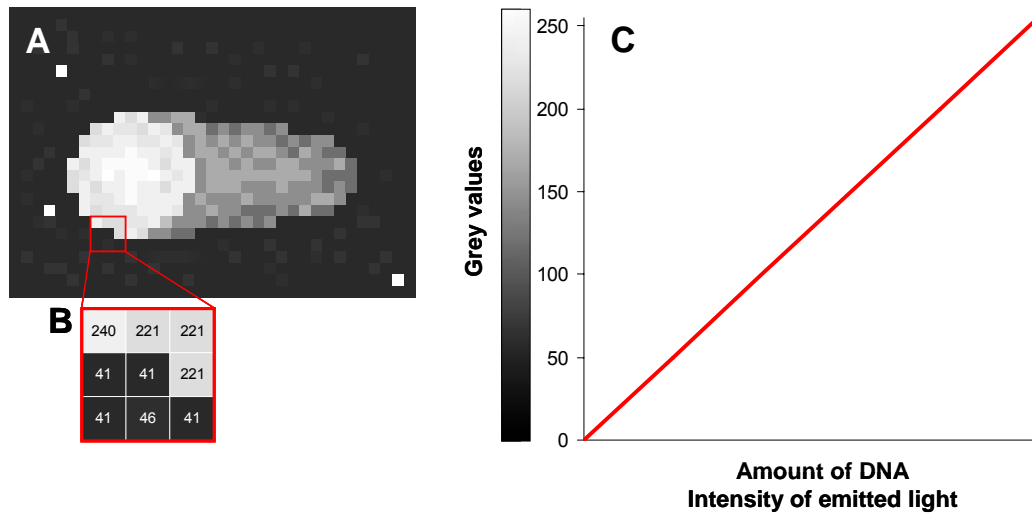
Now we recorded an image, and finally we come to **image analysis**. Many software packages can display the information they receive from a camera on the screen and save image files, but this is not image analysis. The (relatively vague) definition states that **image analysis is a tool for extraction of quantitative (!) information contained in an image**. And this is what we intend to do – we will use the information contained in the image to measure the intensity of light emitted from the comet, and hence the amount of DNA.

Although the general principles of image analysis used in different comet software packages are the same, there is some variation in the exact procedures, so **you need to check the user guide of your own comet software for details**.

### 3.1 Grey values

Our measurement is based on the relationships between different parameters. So far we learned that the relationship between the amount of DNA and the intensity of emitted orange light is linear. The intensity of light (and consequently the amount of DNA) is now coded as numbers in our image file. But what kind of numbers do we have in the file? How are they related to light intensity?

Not so long ago, there would be only one answer to the first question. Lately things got a bit more complicated. For clarity of presentation, let us pretend to go back in time a little and deal with additional problems later (see chapter *Why a 10-bit camera beats an 8-bit camera*, page 38). For the present purposes, the answer is: Light intensities are represented as whole (integer) numbers from 0 to 255 (together 256 values), where 0 means a black pixel on the image and 255 a white one (Figure 4). Because they are represented as shades of grey in the image, these numbers are called **grey values**. The answer to the second question is: **The relationship between light intensity that a camera element detects and the grey value of the pixel on the image is linear**. Hence, **the relationship between the amount of DNA present in one small dot in the comet and the corresponding pixel grey value is linear** (Figure 4).

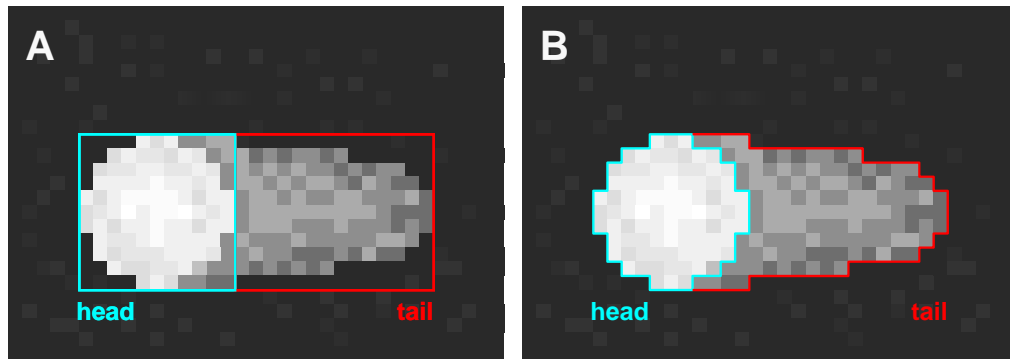


**Figure 4.** Information on an image is coded as grey values. An image of a comet is composed of separate pixels (A), where each pixel has a grey value, as shown in B for the small image area outlined in red in A. The model image is an 8-bit image with available grey values 0-255. The relationship between the grey value of a pixel on the image and light intensity that a camera element detects (emitted from DNA present in one small dot in the comet) is linear (C).

## 3.2 Segmentation

Next we need to determine which parts of the image represent the objects that we wish to measure. This step of image analysis is called **segmentation**, and the regions of the image that are measured are called **regions of interest or ROI**. With segmentation, our image is divided into segments: objects (ROI) and background (all area outside ROIs).

For measurement of the comet parameters, we need to define two regions of interest in the image: the head and the tail. There are several ways to define ROIs. We can draw boxes around the regions, as shown in Figure 5A (box segmentation). Some comet software packages can do a closer fitting of the head and the tail region using more sophisticated (semi-)automatic detection of the head and the tail (Figure 5B; close-fitting segmentation).

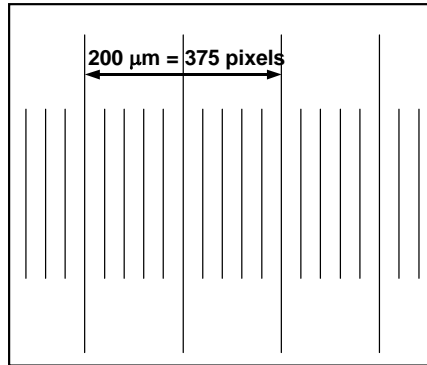


**Figure 5.** Segmentation of the comet image. During segmentation, the regions of interest (ROI) for measurement of the comet parameters (the head and the tail) are defined. The user can interactively draw rectangles that define the head and the tail ROI (**A**; box segmentation). In some software packages, the head and the tail are detected (semi-)automatically and outlined (**B**; close-fitting segmentation).

### 3.3 Geometric calibration

Let us measure one comet parameter for the simple case shown in Figure 5A – the length of the tail. All we need to do is think with computer logic: we count the number of pixels in the tail ROI, along the x-axis. The tail length is 14 pixels – and now we have another problem. We usually express lengths in micrometers, not in pixels. To be able to do this, we need to know how many micrometers there are per each pixel. The computer also needs to know this. So we will now perform a **geometric calibration** of the image analysis system.

What we need to do is record an image of a calibrated microscope slide with an engraved “ruler” in micrometer units, and then measure how many pixels there are per 100  $\mu\text{m}$  (or another known length on the ruler; Figure 6). We only need to perform the geometric calibration once – the **scaling factor** (the number of micrometers per pixel) is saved in the memory of the computer. Geometric calibration has to be done **for each objective** separately – do not just calculate the scaling factor of your 100 $\times$  objective from the measured scaling factor of your 10 $\times$  objective. The objective magnifications are not exact. In our system, the scaling factor for the 10 $\times$  objective is 1.06  $\mu\text{m}$  per pixel, and for the 100 $\times$  objective 0.104  $\mu\text{m}$  per pixel (and disagreements can be much larger). Correct geometric calibration is essential for measurement of those parameters that are expressed in absolute units ( $\mu\text{m}$ ,  $\mu\text{m}^{-2}$ ). We will see later that also the Olive tail moment has absolute units.



**Figure 6.** Geometric calibration of the image analysis system. We record an image of a micrometer ruler under a microscope with a camera and then measure the distance between the micrometer markings on the recorded image. In this case, the scaling factor is  $200\mu\text{m} / 375 \text{ pixels} = 0.53 \mu\text{m}$  per pixel.

Note that the objective magnification is not the only parameter that affects the scaling of images. In some microscopes, **additional magnification lenses** can be either inserted or removed from the optical path. Another important parameter is **binning**, a camera setting that we will discuss later (see *chapter Binning*, page 40). A separate geometric calibration is needed for all **combinations of settings of the image analysis system (on the microscope, the camera or in the software) that affect image scaling**.

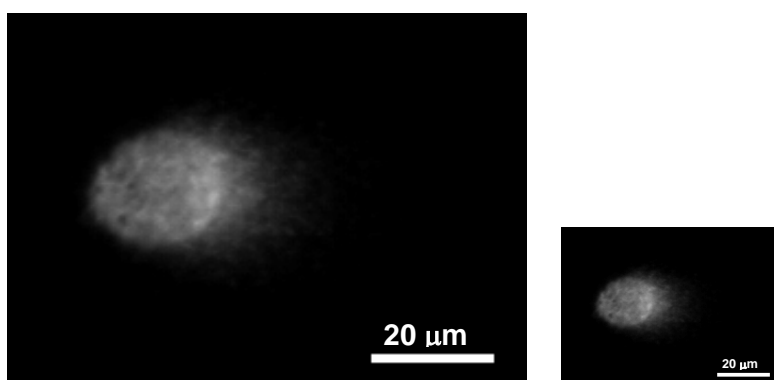
### 3.4 Measurement of geometric parameters

Let us assume that the scaling factor for our model image is  $1.6 \mu\text{m}$  per pixel. We can calculate that the tail length of 14 pixels corresponds to  $22.4 \mu\text{m}$ , so the tail is now measured in **absolute units**. Other geometric parameters of the comet, such as the length of the comet, the length of the head, the area of the comet, the area of the head and the area of the tail can be measured using the same procedure of counting the pixels with a consideration for the scaling factor of the image. These parameters describe lengths and areas and are called **geometric parameters**.

### 3.4.1 And I measured my comet at 400× magnification...

While we are talking about lengths, scaling factors and related topics, let us discuss a common misconception about “magnification”. We recorded an image of a comet and now we want to publish it. For publication, we need to quote the scaling of the image. Let us say that we observed the comets under the microscope, using a 40× objective and 10× eyepieces. We looked at the comets under a microscope at 400× magnification, because the laws of optics state that the final magnification of the image observed equals the product of the magnification of the objective and the magnification of the eyepiece ( $40 \times 10 = 400$ ). Then we recorded an image. Can we now write in the caption of the printed (published) image “at 400× magnification”? No, because this is *wrong*. In a microscope, the optical path that leads from the objective to the eyepieces is different to that from the objective to the camera (Figure 2C). There may be several lenses of different magnification inserted in the objective-camera optical path. A camera is mounted on a microscope with an adapter, and the magnification of commonly used adapters are 0.63× or 1×. Hence, to describe the scaling of an image, we need to insert a scale bar (Figure 7).

The same principle applies to writing the Materials and Methods section of scientific publications. We can mention the magnification of the objective that we used (40×), but not write “Images were recorded at 400× magnification”.



**Figure 7.** Scaling of a printed image. Printed images should contain a scale bar. This image was recorded using a 40× objective, at the scaling factor of the original image  $0.36 \mu\text{m}$  per pixel. In an image editing program (e.g. *Adobe Photoshop*), we can add a scale bar to the original image, which is exactly 56 pixels long to represent  $20 \mu\text{m}$  ( $20 \mu\text{m} / 0.36 \mu\text{m pixel}^{-1} = 56 \text{ pixels}$ ). We then save the image containing the scale bar. Now, we can enlarge or shrink the size of the image - the scale bar always represents the exact current magnification. If printed on A4 paper, the size of the scale bar for the image on the left is 16 mm and for the one on the right 6.5 mm. Hence, in the printed version, the left image is shown (printed) at magnification 800× and the right one at 325× ( $16\text{mm} / 20 \mu\text{m} = 800$ ;  $6.5 \text{ mm} / 20 \mu\text{m} = 325$ ).



## 3.5 Measurement of light intensity parameters

Several parameters measured in the comet assay describe the intensity of light emitted from the comet slide – they are called **the light intensity parameters**. Their measurement is based on information contained in the grey values of an image and they describe the quantity and distribution of DNA in a comet. The very good news is that we only need to measure **the relative amount of DNA** (e.g. light intensity in relative units, the percentage of DNA in the comet tail) and not the absolute amount of DNA (in picograms or the number of basepairs).

### 3.5.1 Integrated light intensity

Let us first take a look at **the integrated light intensity** – the sum of all light intensities in a region of interest (ROI; the head or the tail). This parameter is shortly called the **integrated intensity** or the **total intensity**. Because there is a linear relationship between the intensity of light emitted from a comet and the grey values in an image, the integrated light intensity should be equal to the sum of grey values for all pixels in a ROI.

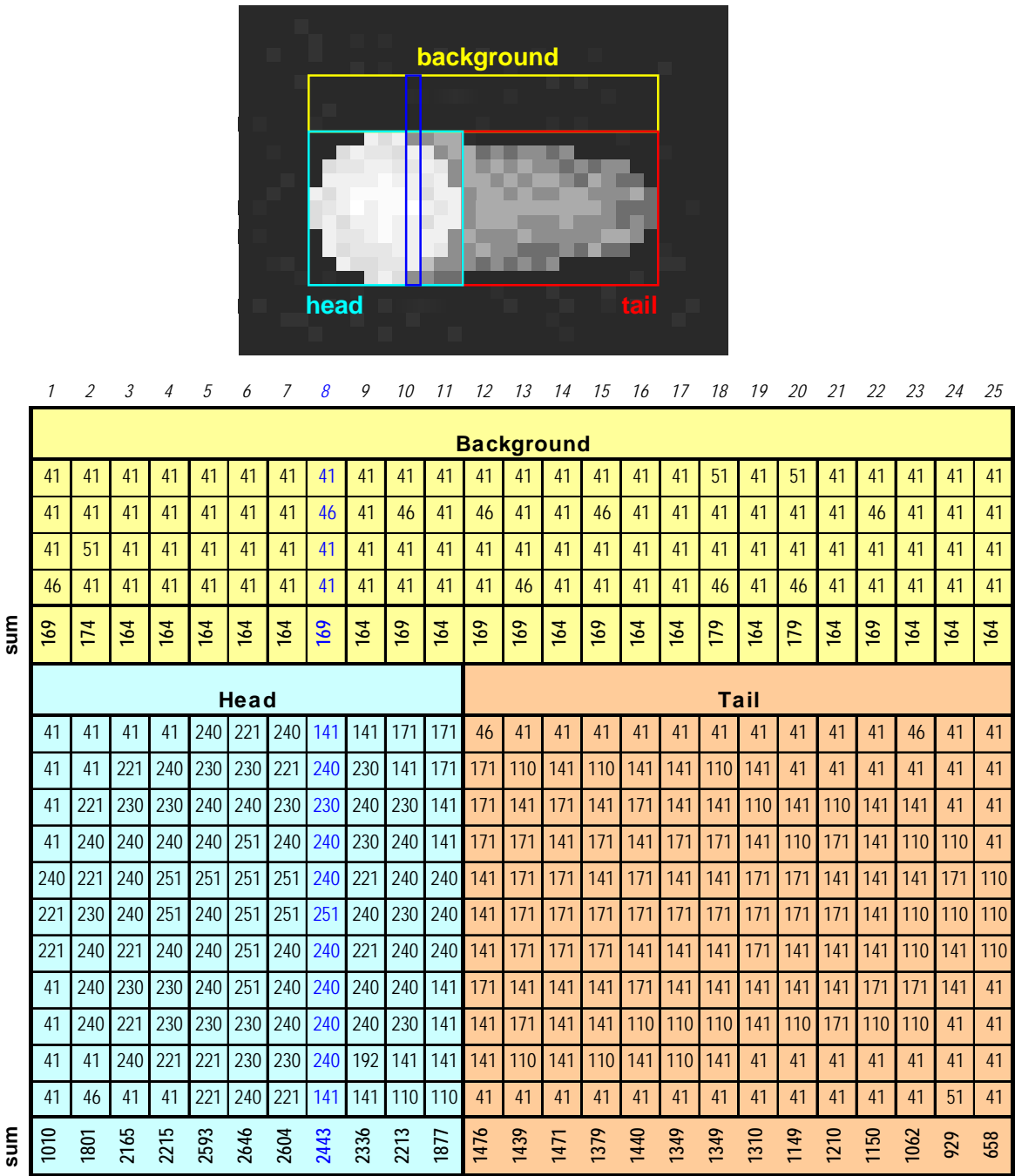
To measure the integrated light intensity for the head of the model comet image, we need to know the grey values of all pixels in the head, and then add the grey values together (Figure 8). The result is 23903. Calculated with the same principle, the integrated light intensity of the tail is 17371.

However, these results are not accurate. The problem is that **the background**, the part of the image where there is no comet, is not completely black. In other words, the grey values of the background are not 0; in the model comet image they range from 41 to 51. The camera is detecting a weak signal in the background, and we will see that **the background signal is as important for our measurement as the comet signal**.

Several factors may contribute to the background signal. Ethidium bromide around the comets may emit light. The camera always has some electronic noise (it “randomly” records grey values higher than 0 when there is no light). The mercury lamp may flicker a little, all parts of the equipment are plugged in the electricity network, which does not have a completely stable voltage, and this adds some more noise to the image. As we will learn later, the background on the recorded image should never be completely black anyway (see chapter *Image saturation*, page 30).

So we have to figure out how to take the background signal into account in our measurement. In the parts of the image where there is no comet, we detected some light intensity (background intensity). Physics teaches us that light intensities are additive. We can thus deduce that in the parts of the image where there is the comet, the signal that we recorded was the light intensity of the background *plus* the intensity of light emitted from DNA-bound ethidium bromide. Hence, **we need to subtract the background light**

**intensity from the total light intensity that we recorded.** If we do not do this, we will compromise the linear relationship between the amount of DNA on the slide and the grey values on the image.



**Figure 8.** Grey values for the head, the tail and the background region of the model comet image. Box segmentation is shown, with the head indicated in **blue**, the tail in **red** and the background in **yellow**. The pixel column number eight, which is used as an example in the text, is outlined in dark blue on the image and shown in blue numbers in the table. The number of each pixel column is indicated in *italic* above the table.

**Table 1.** Subtraction of the background signal for each pixel column of the comet on the model comet image. The data are for box segmentation (see Figure 8 for grey values of individual pixels). The pixel column number eight, which is used as an example in the text, is shown in blue numbers. The background-corrected comet intensity is used for all further calculations. The cumulative comet intensity is calculated from the background-corrected intensities (a separate cumulative frequency for the head and the tail).

ROI	Number of pixel column	Recorded comet intensity (sum of 11 pixels)	Recorded background intensity (sum of 4 pixels)	Average background intensity (per pixel)	Scaled background intensity (for 11 pixels)	Background-corrected comet intensity (sum of 11 pixels)	Cumulative comet intensity
Head	1	1010	169	42	465	545	545
	2	1801	174	44	479	1323	1868
	3	2165	164	41	451	1714	3582
	4	2215	164	41	451	1764	5346
	5	2593	164	41	451	2142	7488
	6	2646	164	41	451	2195	9683
	7	2604	164	41	451	2153	11836
	8	2443	169	42	465	1978	13814
	9	2336	164	41	451	1885	15699
	10	2213	169	42	465	1748	17447
	11	1877	164	41	451	1426	18873
Tail	12	1476	169	42	465	1011	1011
	13	1439	169	42	465	974	1986
	14	1471	164	41	451	1020	3006
	15	1379	169	42	465	914	3920
	16	1440	164	41	451	989	4909
	17	1349	164	41	451	898	5807
	18	1349	179	45	492	857	6664
	19	1310	164	41	451	859	7523
	20	1149	179	45	492	657	8179
	21	1210	164	41	451	759	8938
	22	1150	169	42	465	685	9624
	23	1062	164	41	451	611	10235
	24	929	164	41	451	478	10713
	25	658	164	41	451	207	10920

Subtraction of the background signal seems to be an easy task. However, the problem is that the background signal is not completely even across the image - we have some lighter and some darker pixels. So how can we know whether to subtract 41 or 51 from the grey value for each individual pixel in the comet? Because we cannot know this, we will average the background and then subtract it from the comet image.

Now we need to decide which part of the image we will take as the background. One approach is to take a narrow rectangular region of image just above the comet (Figure 8, region inside the yellow box). In our case, the height of the background region is 4 pixels.

However, we will now not average the grey values for the whole background region – we will instead calculate the average background grey value *for each column of pixels* separately. Why we need to do it this way will become clear later – it is to minimise the measurement error (see chapter *Shading*, page 35).

Let us take the pixel column number eight (outlined in dark blue in Figure 8) as an example. For this column, the sum of grey values for the 11 pixels in the head is 2443. In this column, the sum of grey values for the 4 pixels in the background is 169, so the average grey value of background pixels is 42.25. Therefore, we need to subtract 42.25 from each pixel in the head column. But because there are 11 pixels in the head column, we can subtract  $11 \times 42.25 = 464.75$  from the sum of grey values of the column:  $2443 - 464.75 = 1978.25$ . This is the background-corrected integrated light intensity for pixel column number eight in the head. After we repeat this procedure for all pixel columns in the comet region, we can present the results in a table (Table 1). **The background-corrected comet intensity is our best estimate of the actual intensity of the light emitted from the slide and is linearly proportional to the amount of DNA.**

Now that we have a good estimate of the true intensities recorded by our camera, we can calculate the total intensity of the head and the tail, by adding up intensities for all pixel columns in each region. For the model comet image, **the total intensity of the head** is 18873, and **the total intensity of the tail** 10920. The total intensity of the whole comet is thus 29793, the sum of the head and the tail intensities.

### 3.5.2 Percentage of DNA in the tail

One of the widely used comet parameters is **the percentage of DNA in the tail**. It is calculated as the ratio between the total intensity of the tail and the total intensity of the comet (head and tail together). For the model image of the comet, the percentage of DNA in the tail is  $10920 / 29793 = 0.37 = 37\%$ .

To illustrate how important it is to eliminate sources of methodological error such as the background signal, let us calculate the percentage of DNA in the tail *without background correction*:  $17371 / 41274 = 42\%$ .

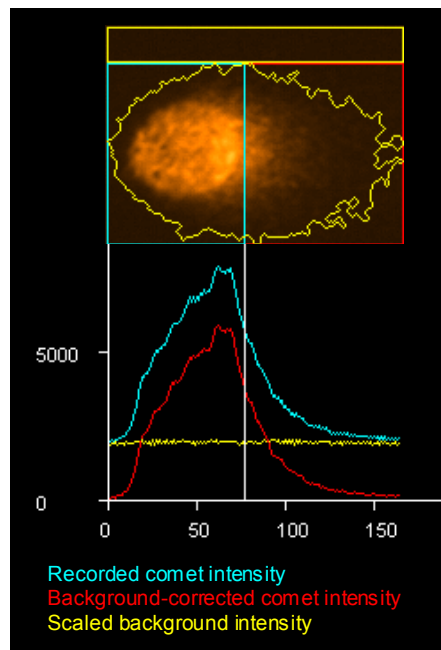
### 3.5.3 Tail extent moment

We can determine another comet parameter - the **tail extent moment**, defined as the product of the tail length (also called the tail extent) and the percentage of DNA in the tail.

We already know the tail length for our model comet image, so we can now calculate the tail extent moment:  $22.4 \mu\text{m} \times 37\% = 8.3 \mu\text{m}$ . Note that the tail extent moment is measured in absolute units - micrometers.

### 3.5.4 Comet profile

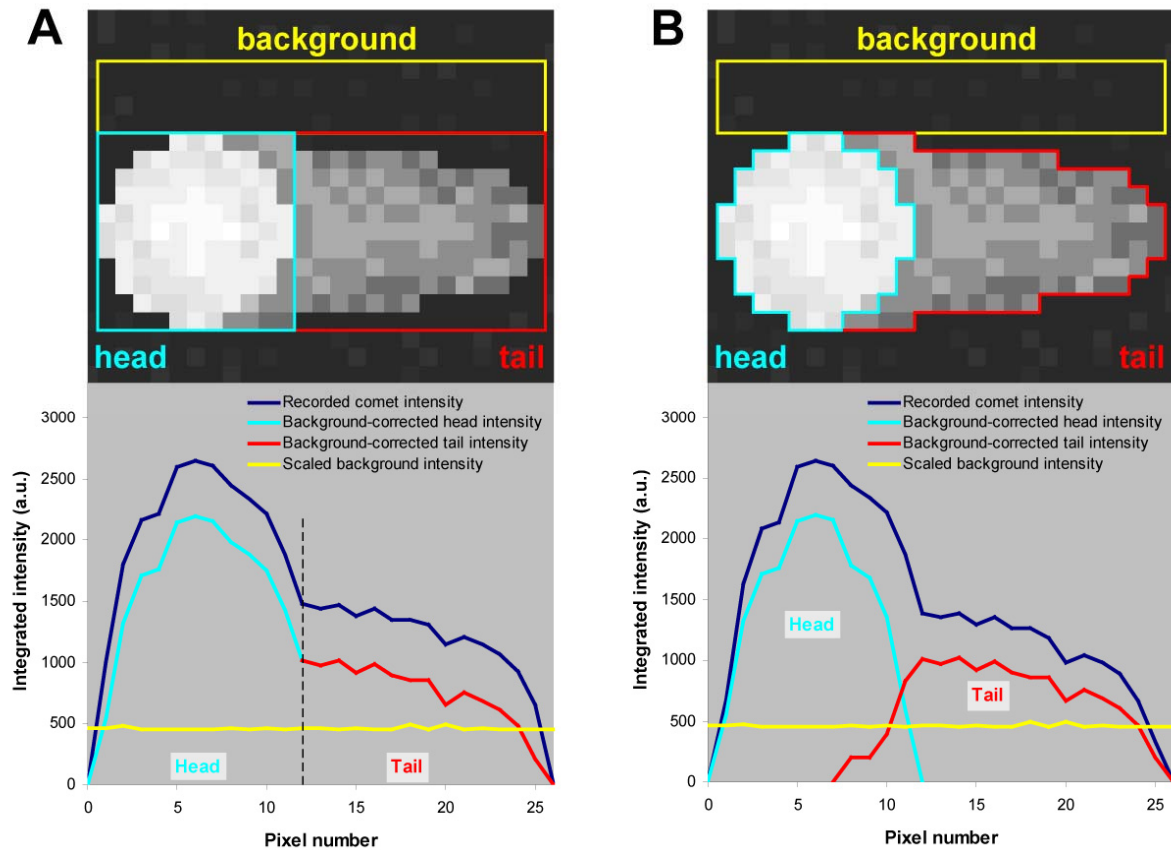
We have now done all the necessary things to tackle a new task – the **comet profile** of light intensities. An example of a comet profile is shown in Figure 9. How does the computer calculate and draw the profile?



**Figure 9.** The comet profile for a comet recorded from a microscope slide.

The comet profile is a graph showing the **total intensities of pixel columns along the x-axis of the comet**. So all we need to do to get a comet profile for our model image is to plot a graph from the data in Table 1 (column *Background-corrected comet intensity*), taking into account the background signal (Figure 10A). The same principles can also be applied to close-fitting segmentation (Figure 10B).

For accurate measurement, **the background profile should have similar values along the x-axis** (Figures 9, 10). Any “bumps” in the background profile indicate that there is a part of a comet or some other fluorescing object included in the background ROI. In such cases, we need to re-position the background box or cancel measurement of this particular comet. A gradual upwards or downwards **slope in the background profile** indicates that a part of a comet is most likely included in the background ROI. If *all* comets recorded at the same position on the image have a sloping background profile, the problem is probably severe shading of the image (see chapter *Shading*, page 35). Comets with sloping backgrounds should not be measured.



**Figure 10.** The comet profile of the model comet image. **Box-segmentation (A):** The **recorded comet intensity** is the intensity on the original (grabbed) image. The **scaled background intensity** is the mean intensity in the background region (4 pixels) scaled to the height of the comet ROI (11 pixels; the *scaled background intensity* from Table 1). The **background-corrected intensity** is the recorded intensity minus the scaled background. The **dashed line** separates the head from the tail. See Table 1 for data shown in A. **Close-fitting segmentation (B):** With this segmentation, a separate profile is created for the head and the tail. The principle of calculation of the profile is the same as in A. Grey values of the image used for calculations of the profiles are shown in Figure 8.

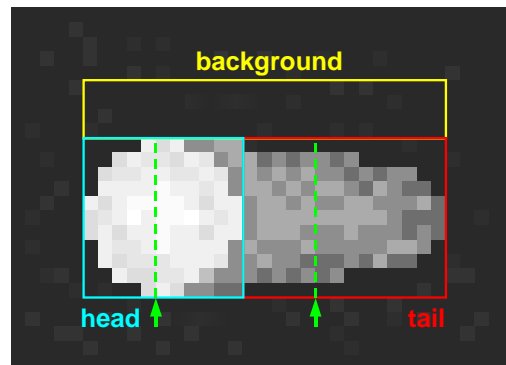
### 3.5.5 Olive tail moment

The **Olive tail moment (OTM)** is the last comet parameter for us to determine. OTM is calculated as a product of two factors: the percentage of DNA in the tail and the distance between the **intensity centroids** of the head and the tail along the x-axis of the comet. We already know the first factor, but what is an intensity centroid?

The centroid of an object is also called the centre of mass. However, in our image of the comet, the comet has no mass! Instead of physical mass, we will use grey values as the weights to determine the position of the intensity centroid.

We only need to determine the position of the intensity centroids along the x-axis of the comet. We also have to do calculations for the head and the tail separately.

Let us look at our model comet image (Figure 11). First, we will determine the position of the intensity centroid of the tail. The centroid is the **x-coordinate at which half of the total intensity of the tail is to the left and half to the right of this coordinate** (half of “grey value weight” to each side of the coordinate). The total intensity of the whole tail is 10920 (Table 1), so we are looking for the x-coordinate at which the total intensity to either side will be  $10920 / 2 = 5460$ . We start adding intensity weights at the left side of the tail. We already know that the total intensity of the left most pixel column of the tail (pixel column number 12) equals 1011 (Table 1, column *Background-corrected comet intensity*). Because this is less than 5430, we can add the weight of the next column to the right, and so on. We calculate  $1011 + 974 + 1020 + 914 + 989 = 4909$  (see column *Cumulative comet intensity* in Table 1). This is lower than the limit 5460. If we now add the intensity weight of the next column (pixel column number 17), the sum of weights will be 5807 - higher than the limit 5460. Hence, we can conclude that the intensity centroid of the tail lies between pixel columns 16 and 17 (Figure 11). Using the same procedure, we can determine the position of the centroid of the head between pixel columns 5 and 6. The distance between the intensity centroids of the head and of the tail is 11 pixels (Figure 11).



**Figure 11.** The intensity centroids. Positions of the intensity centroids are shown as green arrows on the model comet image with box segmentation. The positions of the centroids can be determined from data in Table 1. See text for details.

We measure distances in micrometers, hence we have to convert pixels to micrometers. The scaling factor of the model image is  $0.36 \mu\text{m}$  per pixel, so the distance is  $0.36 \mu\text{m pixel}^{-1} \times 11 \text{ pixels} = 3.96 \mu\text{m}$ .

Now we can also calculate OTM as the percentage of DNA in the tail multiplied by the distance between the intensity centroids:

$$\text{OTM} = 37\% \times 3.96 \mu\text{m} = 0.37 \times 3.96 \mu\text{m} = 1.5 \mu\text{m}.$$

**OTM is an absolute parameter with a measurement unit – micrometer.**

Now that we are finished with calculations of the comet parameters, we can turn to some general topics of image analysis.

## 4 Recorded image and displayed image

We assumed at the beginning of our calculation of comet parameters that we recorded an image of a comet with a camera. We worked with the pixel grey values of the original recorded image that are stored in the image file. However, when we are measuring comets, we have no direct access to the grey value data – instead of the numbers, the computer shows us an image on the screen.

When an image of a comet is shown on the screen, what we really see is a *reconstruction* of the recorded image that the computer creates from the file where all grey values of the recorded image are stored. Recording of an image is also called **image grabbing** or **image acquisition**. We will call the image that we recorded with the camera and is now saved in the computer memory **the grabbed image** and the image that we see on the screen **the displayed image**. For our measurement, the important thing is the grabbed image, because it contains the grey values that represent light intensity that, in turn, represent the amount of DNA on the slide.

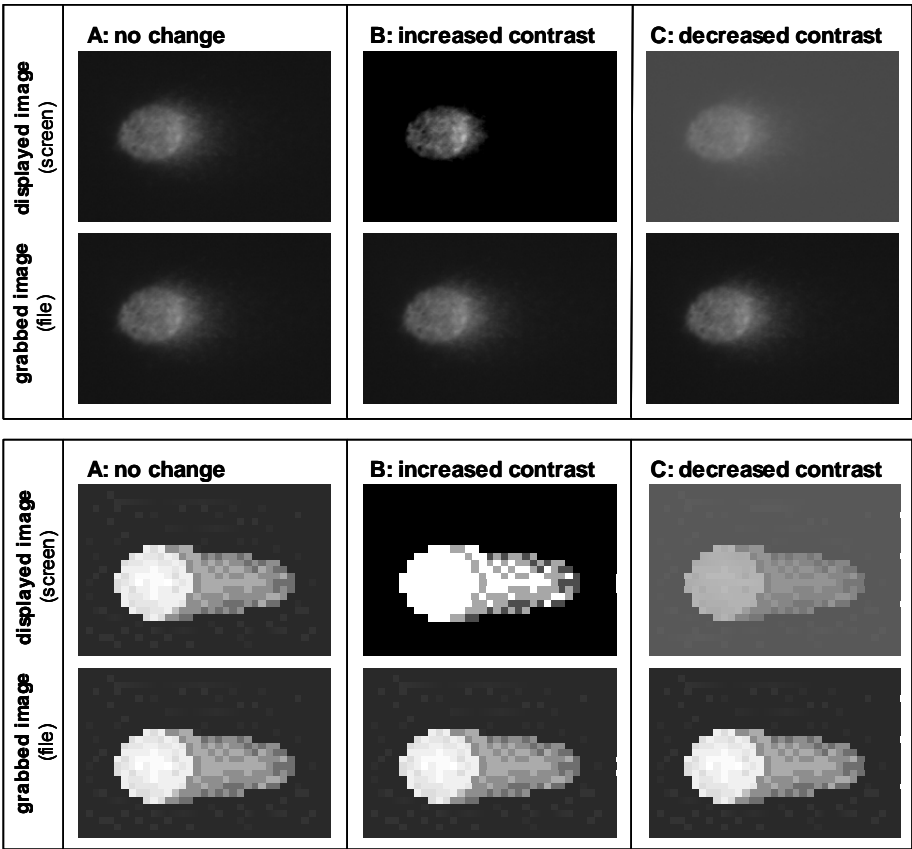
The grabbed and the displayed image are not the same thing: what we see on the screen is only a representation of the grabbed image. If we change the way that the image is displayed on the screen, the file where the grey values for the grabbed image are stored is not changed. If we enlarge or decrease the size of the displayed image by dragging its corners, the grabbed image file does not change. If we change the brightness or the contrast of the displayed image, the grabbed image does not change (Figure 12).

### 4.1 Look-up table (LUT)

Images of comets are composed of 256 different shades of grey; our eyes can distinguish only a few dozens. Often a comet looks relatively dark on the image, and we are very bad at differentiating dark shades of grey – we can hardly see the comet, although the computer can easily measure it. We are much better at distinguishing colours than shades of grey. We just found out that the displayed image can be different from the grabbed image. Can this help us see?

Before we continue this debate, let us look at types of images with respect to colour. **Greyscale images** are composed of different shades of grey (Figure 13A). In everyday life, these images are called “black-and-white”. However, in image analysis we actually have a special type of images that are truly only black (grey value 0) and white (grey value 255). These images are **binary images** (Figure 13B; also called masks) and are used for segmentation - to tell the computer which parts of the image are objects (ROI) that we wish to measure (by definition shown in white) and which are other non-interesting parts of the image (by definition shown in black). And we also have **colour images** (Figure 13C), sometimes also called RGB images (RGB – red green blue).





**Figure 12.** The grabbed and the displayed image. Changing the settings of the displayed image that we see on the screen (**top row**) does not affect the grabbed image, stored as an array (table) of grey values (numbers) in the computer memory (**bottom row**). The following settings of the displayed image are shown: the grabbed image is displayed with no changes to image settings (**A**); the grabbed image is displayed with increased contrast (**B**); the grabbed image is displayed with decreased contrast (**C**). See chapter *Image contrast and brightness*, page 27, for explanation about image contrast.



**Figure 13.** Types of images with respect to colour. **A** – a greyscale image; **B** – a binary (black-and-white) image with the object (a comet) shown in white and the rest of the image in black; **C** – a colour (RGB) image.

Images of orange comets on a black background (Figure 13C) are very popular. Notably, most systems for comet measurement use monochrome cameras (called “black-and white” cameras in ordinary life) that only detect the intensity of light, but not colour (wavelength). Monochrome cameras can thus only record greyscale images, not colour images. So how does the computer construct the orange image of a comet when all it received from the camera were grey values (information about light intensity) and no information at all about the colour of the comets?

To know how to display an orange image on the screen, the computer looks into a special table, which is, not surprisingly, called a **look-up table (LUT)**. A LUT is a table similar to a translation dictionary; each grey value (0-255) is paired up with its colour translation. The LUT for the orange images of comets says that black (grey value 0) is translated to black, white (grey value 255) is translated to bright yellow, dark grey (grey value 72) to very dark orange, and so on for all 256 grey values.

An image analysis software may contain many different LUTs. Some translate a greyscale image to a colour image, some to another (different) greyscale image. Using different LUTs, we can display a grabbed image of a comet in many different ways (Figure 14).

When we want to see what the original grabbed image actually looks like, the computer uses the LUT where each of the grey values on the grabbed image is translated to the same grey value on the displayed image (Figure 14B), so that the displayed image is an exact copy of the grabbed image saved in the computer memory.

LUTs are in many ways a wonderful thing. LUTs can bring out the details in an image that we would otherwise not be able to see. However, it is clear from Figure 14 that the type of LUT that we use to display an image of a comet affects the way we perceive a comet. For example, with some LUTs, the comet tail looks more prominent than with others, hence using different LUTs during measurement of one comet experiment can affect our decisions about the size of the head and the tail if we do interactive and not automatic measurements.

If your comet software uses LUTs to display images, there should be a menu from which you can select the LUT you want to use.

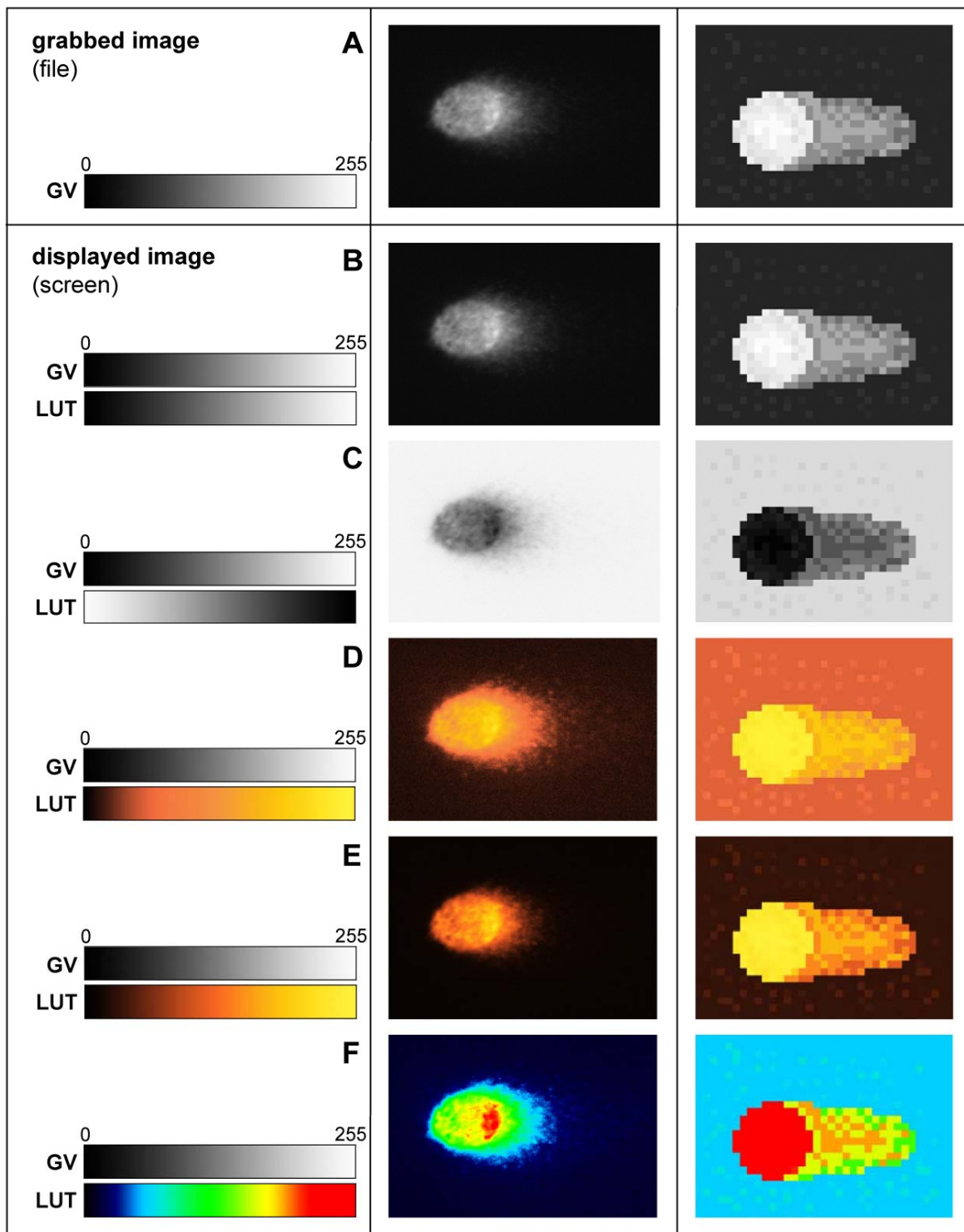
## 4.2 Live image and grabbed image

When we place a comet slide under a microscope and switch on the camera to record images, we see the image of the slide displayed on the screen and the comets on the screen move as we move the microscope stage. This kind of communication between the camera and the computer is called “live mode”, and we say that we see a **live image**. The camera is grabbing subsequent images at very short time intervals and sending the image files to the computer. The computer saves each image file in temporary memory, quickly decodes it, displays it on the screen and deletes it. Then it displays the next one. Hence we have the

same illusion of continuous motion as we have when watching movies, which are also composed of separate images.

When we click to record an image, we say that the image is grabbed or acquired. Unlike the live image, the **grabbed image** is saved in more permanent memory of the computer for later use in image analysis and measurements.

We can read in the instructions for some comet software packages that measurement is possible on live or imported (previously recorded) images. Measurement of a truly live image is not possible, because such an image does not exist in digital technology. The computer can only perform a measurement on a grabbed image, i.e. a saved image file where all grey values for all pixels are strictly defined numbers. There are, however, essentially two approaches to measurements. We can have a slide on the microscope, find a comet, record an image and measure it immediately (live mode). Or, we can record images of comets, save them in the computer memory, on a hard disk or on a CD, and do the measurements later. If we have the possibility to do “live” measurements, our camera is integrated in the software so that the computer can communicate with it, and we can both record and measure images without exiting the program. Therefore, **if we want to do “live” measurements, we have to make sure when buying the equipment for the comet assay that the software package supports our camera.**



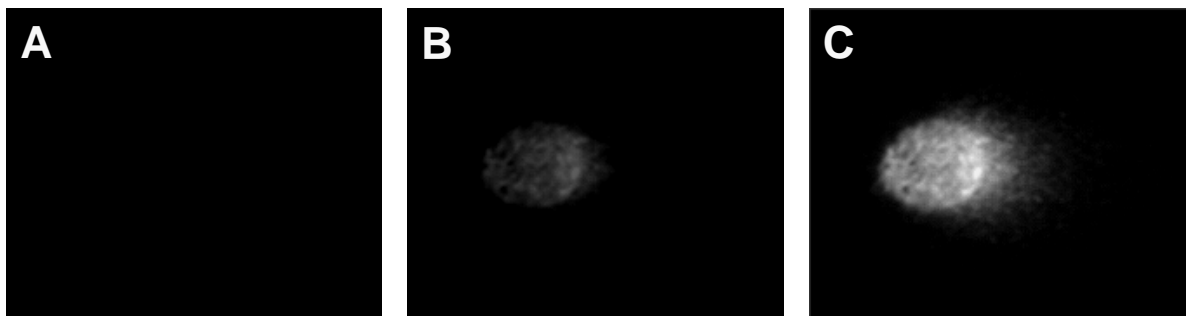
**Figure 14.** Look-up tables (LUTs). A LUT contains instructions for translation of the grey values on the grabbed image recorded by a camera (**A**) to the displayed image on the screen (**B-F**). If a LUT translates each grabbed grey value to the same displayed grey value, we see on the screen the actual grabbed image (**B**). A LUT is also used to construct a negative (inverted) greyscale image (**C**). The LUTs that were used to obtain the displayed images are shown in the left column. Note that a very small difference in LUTs can drastically change the way we perceive a comet (compare LUTs and images **D** and **E**).

## 5 Image acquisition

The first rule of image analysis states that all care should be taken to **record images of the highest quality possible**. Information that is not contained in the grabbed (recorded) image cannot be restored later, no matter how we change and process our image.

An image analysis system has many settings that affect the properties (quality) of the grabbed image. It is extremely important to make sure that all settings are optimal for recording of the comet images. **Incorrect settings for image acquisition can invalidate the linear relationship between the intensity of light emitted from the comet slide and the grey values on the recorded image!** This results in inaccurate measurement of the comet parameters. The main principle of image acquisition is this: **The image for measurement of the comet parameters must contain as much information as possible.** How do we achieve this?

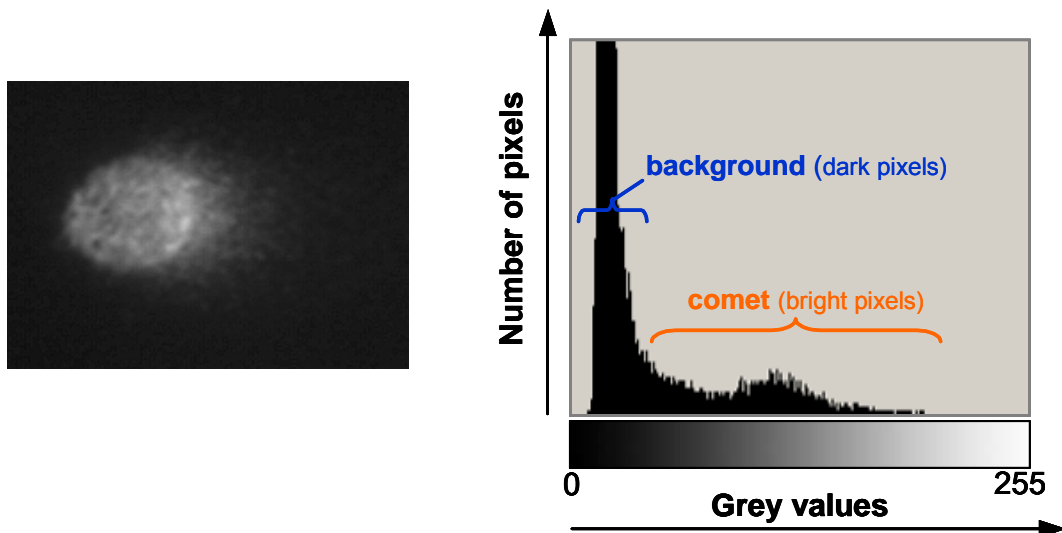
Our measurement of the comet parameters extracts quantitative information (numerical parameters) from an image. The more information the image contains, the more precise our measurement is. Let us say that we were careless when we were recording a comet image and we did not optimally adjust all settings that define image quality. So we grabbed a completely black image – the grey value of all pixels is 0 (Figure 15A). This image contains no information at all – we do not even know if there was a comet under a microscope or not. Now we try to grab a better image, and we adjust some settings. This time we have more “luck” – we can still hardly detect the comet, but at least we see that there is one (Figure 15B). Then we record a very good image – now we can see many different shades of grey in the comet head and tail (Figure 15C). This image contains sufficient information for accurate measurement of the comet parameters. Or, at least, this is what it looks like to our eyes.



**Figure 15.** Effects of image acquisition settings on the information contained in an image of a comet. A black image contains no information (A). Image B contains more information than image A, but does not show all details of the comet. Image C contains the most information of the three images.

## 5.1 Histogram of image grey values

It is, however, not important at all for our measurement whether the image “looks” fine to us. The only important thing is what the computer can distinguish in it. Because our eyes are a lousy tool for estimation of the number of shades of grey that an image contains, we will evaluate the quality properties of an image that are important for measurement of the comet assay by inspection of a **histogram of image grey values**. The histogram shows the frequency distribution of pixels along the whole range of grey values (0-255; Figure 16). The majority of the pixels in our comet image belong to the background, hence there is a high peak of pixel frequencies in the low grey value range. Because there is some variation in the background signal across the image, the background grey-value peak is not very narrow. The pixels of the comet have brighter shades of grey than the background and are represented by a long trail of lower pixel frequencies to the right of the background peak.



**Figure 16.** The histogram of image grey values. Because most of the image is very dark (background), the frequency of pixels with low grey values is high. To see also the distribution of brighter pixels (the pixels of the comet), the lower part of the y-axis has been scaled up and hence the top of the peak at low grey values is cut off. Note that in this image there are no pixels with very low or very high grey values.

## 5.2 Image contrast and brightness

The goal of image acquisition in the comet assay is to describe the brightness of the comet with as many different grey values as possible. The more grey values we use, the more accurately we can tell apart different intensities of light emitted from different parts of the comet, the more information the image contains and the more precisely we can measure the comet parameters. However, because the calculation of the comet parameters depends upon subtraction of the background signal, we also have to record **variation in the background signal**, not only in the comet.

The two basic properties of an image, related to the range of grey values that are used to store information, are image contrast and brightness. **Image contrast** is the difference between the highest (brightest) and the lowest (darkest) grey value in an image. **Image brightness** describes whether the range of grey values used in the image is in the darker part of the greyscale (at low grey values) or in the brighter part of the greyscale (at high grey values).

Let us compare the three images of the same comet shown in Figure 17. Image A has a high contrast, using grey values from 14 to 198, so that the total contrast range is 185. Image B has a lower contrast – it contains grey values from 6 to 85 and its contrast range is only 80. Image C has the same contrast range as image B, namely 80. However, in image C, the range of used grey values is in the brighter part of the greyscale (82 to 161), hence image C had a higher brightness than image B.

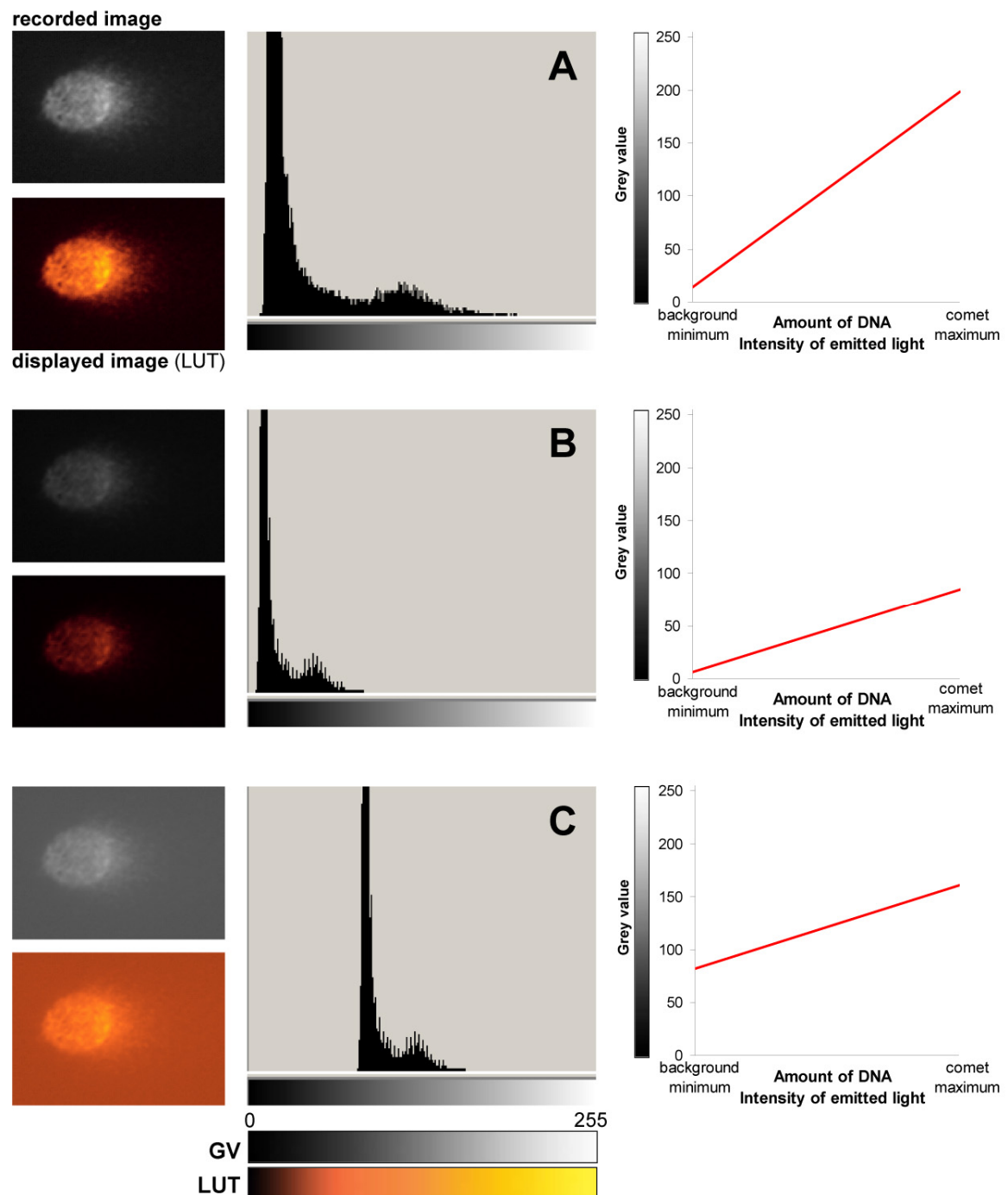
We must not forget that the grey values on the image are our measurement tool and that they must linearly relate to the intensity of light emitted from the comet (we postulated this general condition of precise comet measurement in Figure 4). The graphs on the right of the histograms of image grey values in Figure 17 show that in all three images this relationship is linear. However, there is a major difference between image A and images B and C. In image A, the intensity of the signal from the microscope is described with 185 different grey values. On the other hand, in images B and C, there are only 80 grey values. Therefore, the measurement of the comet parameters from image A would be more *precise* than the measurement from images B or C. Measurement from all three images would be, within limitations of different precision, valid.

To illustrate the problem of measurement precision, let us imagine that we are running an experiment in which we need to measure the height of a person. The precision of our result will largely depend on the instrument we will use for measurement. If we use a ruler that only shows metre marks, we will be in serious trouble – we can only say that this person is between 1 and 2 metres high. If we take a ruler that has also decimetre marks, we can do our job a sort of acceptably. If we use a ruler with centimetre marks, we are really getting the best out of our experiment.

If we return to the comet world, we are measuring the same comet in all three images in Figure 17. If we measure image A, we will have a “ruler” with 185 grey values to describe the range of light intensities emitted from the background and the comet. In images B and C, our ruler only has 80 different grey values.

We can conclude that **we have to use as many grey values as possible to describe different light intensities emitted from the comet and (!) the background**. However, there is an important **limitation**, as we will see in the next chapter – no pixels should be completely black (grey value 0) or completely white (grey value 255).

Of the three images shown in Figure 17, image A has the highest informational content and is most suitable for comet measurements.



**Figure 17.** Image contrast and brightness. Images of the same comet are shown in A, B and C. The grabbed images are shown as greyscale images (GV) and displayed also using an orange-scale LUT (shown at the bottom). Image **A** has a high contrast (grey values 14 - 198, contrast range 185). Image **B** has a lower contrast (grey values 6 - 85, contrast range 80). Image **C** has the same contrast range as image B (80, grey values 82 - 161). However, in image C, the range of used grey values is in the brighter part of the greyscale, hence image C is brighter than image B. The relationship between the intensity of light emitted from the slide and grey values on the image is shown in the right column. See text for details.



## 5.3 Image saturation

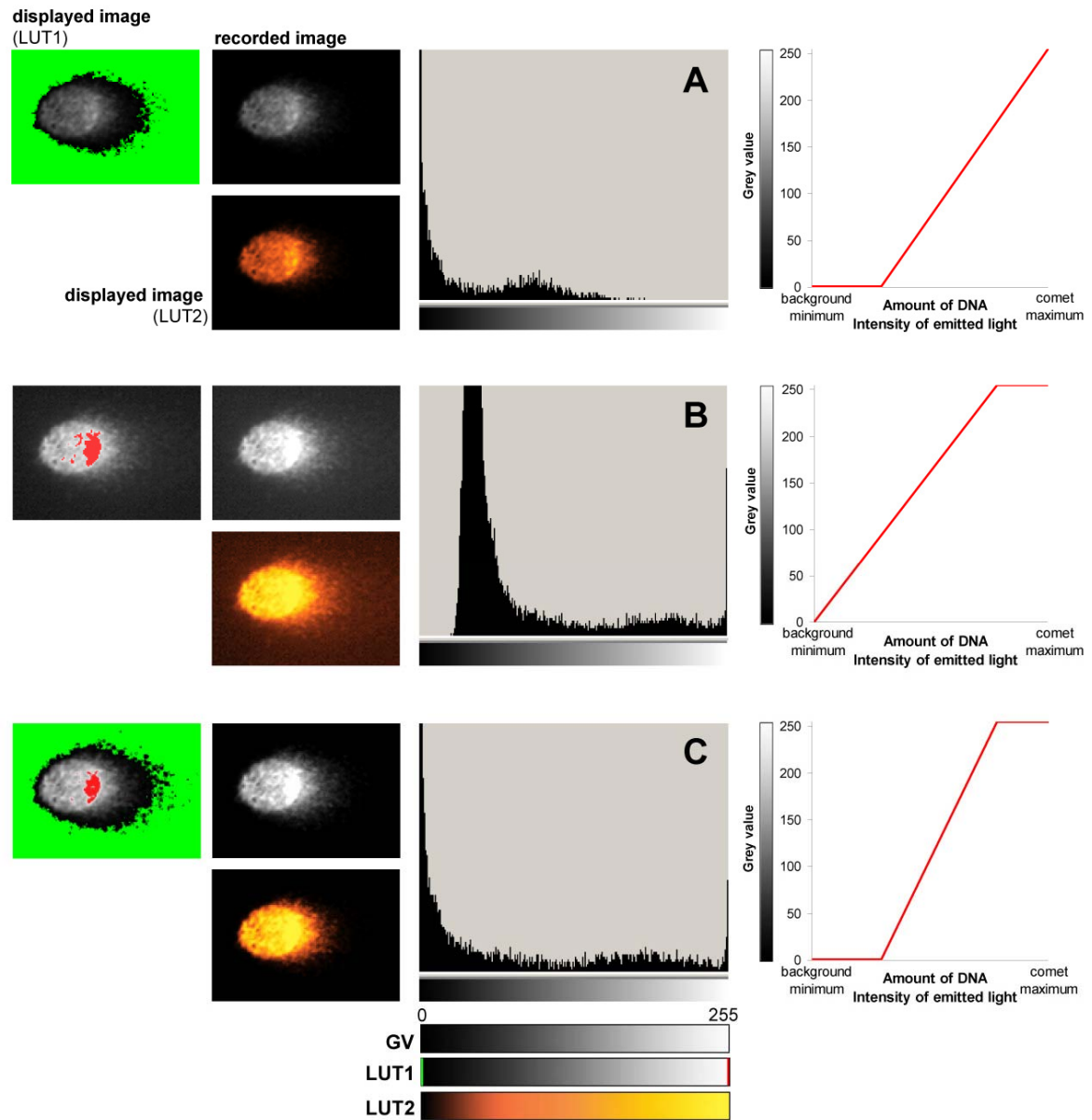
The recorded image is not suitable for measurement of the comet parameters if it is either overexposed or underexposed or both (Figure 18). This means that some pixels in the image have the minimum grey value (0) or the maximum grey value (255); we say that these pixels are **saturated**. Consequently, an image that contains saturated pixels is also saturated. Images containing pixels with grey value 255 (white) are **overexposed**, and those containing pixels with grey value 0 (black) **underexposed**.

The problem with **saturated pixels** is that they **do not contain enough information for a precise comet measurement**. Let us look at examples in Figure 18. Image A has a good contrast – it contains high and low grey values. However, a large area of the background has grey value 0 - in the histogram of image grey values, the frequency of pixels with grey value 0 is high – the image is underexposed. We can also visualise the saturated pixels by using a special LUT (LUT1 in Figure 18), where all grey values of the recorded image are shown as the original grey values in the displayed image, but the pixels with grey value 0 or 255 are displayed in colour (in this case, grey value 0 is green and grey value 255 red). While we know from the information contained in the image that the saturated regions of the background have a lower light intensity than the region of the comet, we do not know about the *variation of the signal* in the background itself (see the graph to the right of image histogram in Figure 18). If we now position the background ROI above the comet and measure the comet parameters, these parameters are *wrong*, because the subtraction of the background signal during calculation of the comet profile is incorrect. The background signal in Image A was not recorded appropriately, and such an image *should not be measured*.

Likewise, images B and C in Figure 18 should not be measured. Image B is overexposed, so all we know about some pixels in the head is that they are very bright, but not also how bright they are. Image C illustrates that an image can be both underexposed (no detailed signal from the background) and overexposed (no detailed signal from the comet head) at the same time.

We should only measure images that contain full information about the variation of light intensity in the comet head, in the comet tail and in the background. If we keep in mind the problem of contrast explained in the previous subchapter, **we should only measure images that have a good contrast and are not saturated**. Among examples shown in Figures 17 and 18, image A in Figure 17 is most appropriate for comet measurement – it gives us a *precise measurement of the comet parameters*. Images B and C in Figure 17 give us a less precise measurement of the comet parameters, because we can only crudely distinguish between shades of grey in the image (few grey values are used to record information). However, measurements from these images can be regarded as *rough estimates* of the true value of the comet parameters. Comet parameters measured from any of the images shown in Figure 18 are *wrong* (or uncontrollably inaccurate at best), because

they violate the linear relationship between the amount of DNA on the slide and the recorded grey values.



**Figure 18.** Image saturation. Images of the same comet are shown in A, B and C. The grabbed images are shown as greyscale images (**GV**) and displayed also using two LUTs (shown at the bottom). **LUT1** displays the grey values of the grabbed image, except for two grey values: grey value 0 (underexposed pixels) is shown in green and grey value 255 (overexposed pixels) in red. **LUT2** displays the recorded image in shades of orange. All images are saturated: image A is underexposed; image B is over exposed; image C is both underexposed and overexposed. The relationship between the intensity of light emitted from the slide and grey values on the image is shown in the right column. See text for details.

Some comet software packages feature an option to set the maximum number of allowed saturated pixels. A very small number of saturated pixels does not severely affect the measurement. In addition, some light-sensitive detector elements in the camera may not function properly; they may not be able to detect the signal or may misfire when there is no light. Even a new camera may have some dysfunctional pixels. In any case, we should **avoid systematically overexposing or underexposing images**.

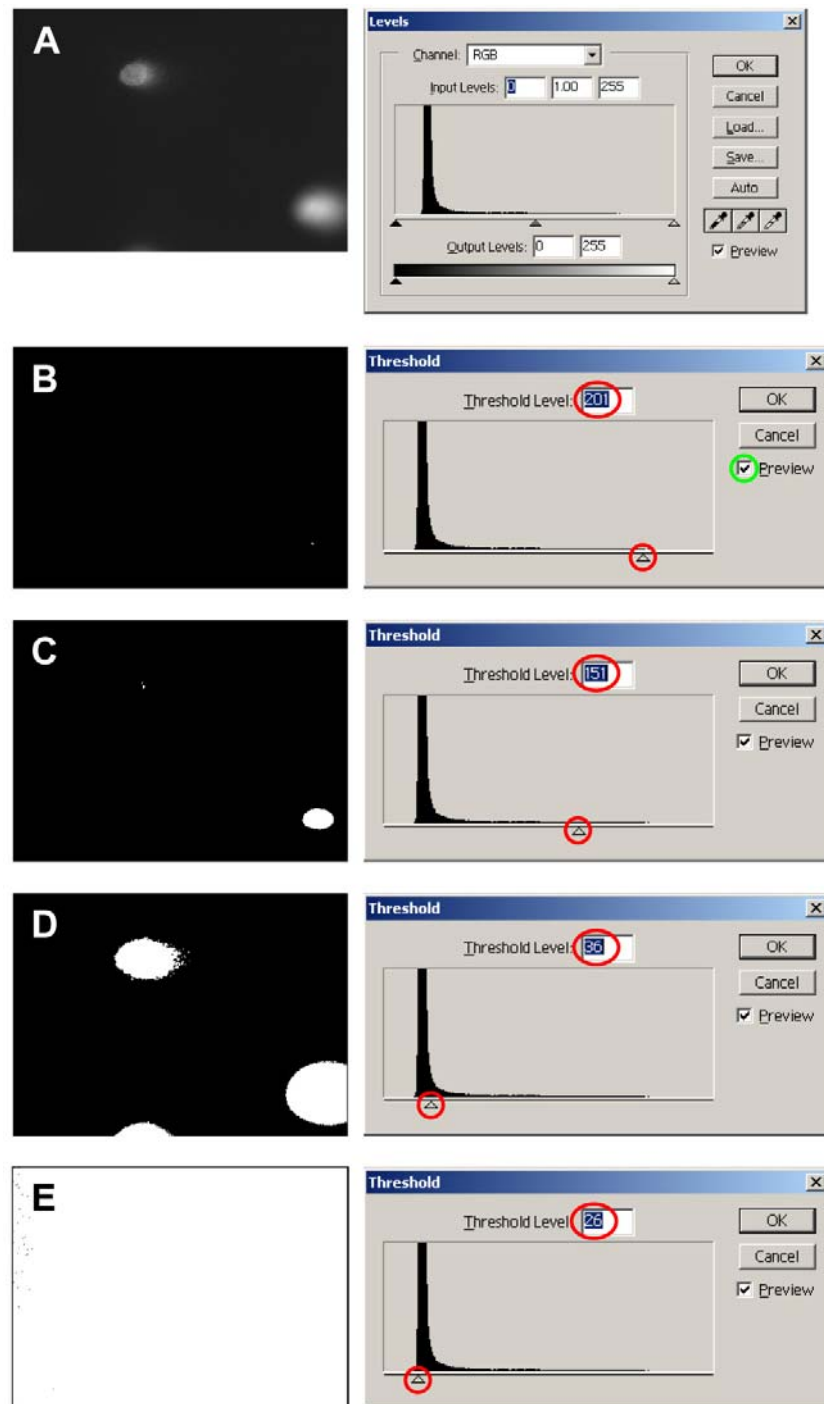
## 5.4 How can I evaluate the quality of my image?

A histogram of image grey values is an essential tool for evaluation of image quality and optimisation of the settings of image acquisition. Nevertheless, not all commercially available comet software packages have the option to display the histogram. In some cases, the histogram is included in the user interface for adjustment of the camera settings (see examples in Figure 25). If we want to display histograms of image grey values for our comet images but our comet software cannot do it, we can use another image processing software that features the histogram function. Here we will look at functions available in *Adobe Photoshop*.

In our comet program, we save the image in a lossless format, such as TIFF (the TIFF format preserves all information contained in the grabbed image without any distortion). We need to make sure that the image that we export (save) is the original grabbed image recorded by the camera, and not just a displayed version of the grabbed image. We are only concerned about the quality of the recorded images because these are measured; the quality of the displayed images is not important (for further comments see chapter *Which image does my software measure?*, page 41). Then we open our saved image in *Photoshop* (from the main menu File / Open). To see the histogram of image grey values, we select from the main menu Image / Adjust(ments) / Levels (Figure 19A). Because an image of the whole field of view usually contains a lot of background regions and just a few comet regions, it is difficult to see the frequency distribution of the signal that we are most interested in, namely the brighter regions of comets. (In Figures 16, 17 and 18, this was not a problem, because only the small part of the whole image that contains the comet was analysed.) In some image processing software packages, it is possible to scale up the y-axis of the histogram, which solves the problem.

For a more detailed analysis of image grey values in *Photoshop*, we can segment the image on the basis of grey values. We select from the main menu Image / Adjust(ments) / Threshold (Figure 19B-E). When we check the “Preview” checkbox, the original greyscale image is converted to a binary (black and white) image. If we now drag the slider under the histogram from right (grey value 255) to left (grey value 0), we see different details of the image emerging on the binary image. At each position of the slider, we see all regions of the image that have the grey value *higher* than the grey value selected with the slider as

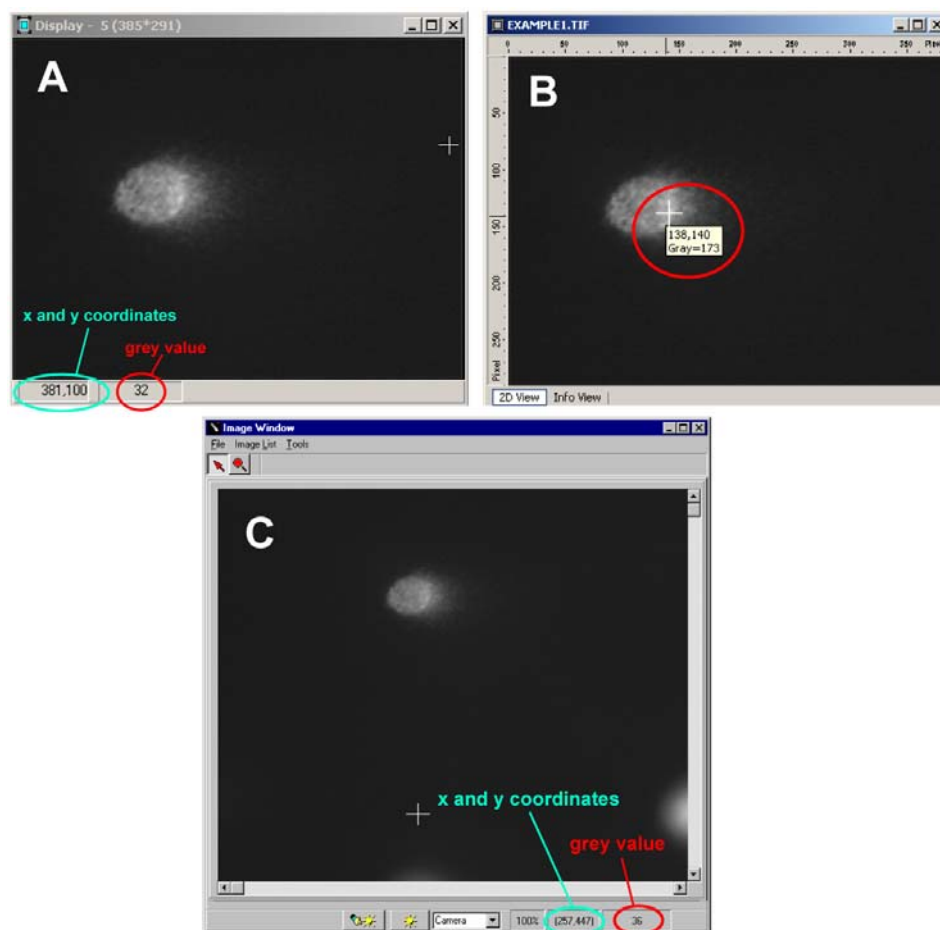
*white regions*, and those regions of the image that have the grey value *lower* than the grey value selected with the slider as *black regions*.



**Figure 19.** Viewing image histograms and grey values in *Photoshop*. Image histogram is displayed using the “Levels” tool (A). Using the “Threshold” tool to analyse the grey value range used in the image is illustrated in B-D. Note that the “Preview” box (green circle in B) has to be selected in the “Threshold” window to update image as we drag the slider under the histogram. The current position of the slider (current threshold grey value) is displayed in the “Threshold Level” box. The shown threshold grey values are: 201 in B, 151 in C, 36 in D and 26 in E. See text for explanation.

Let us now drag the slider from left to right for the image shown in Figure 19. At grey value 0, the whole image is black (no signal). At grey value 26, the first pixels appear as small white dots (Figure 19B) – the minimum grey value in our image is 26. At grey value 36, we can see the first pixels of our comet (Figure 19C), and at grey value 151 the whole comet (Figure 19D). At grey value 201, we see the last black pixels on a white image (Figure 19E) – this is the maximum grey value used in our image. So, we found out that the range of grey values used in our image is 26 to 201, and that the signal in our comet covers the grey value range 36 to 151.

In some software packages, the grey value of the pixel at the current position of the cursor can be displayed on the screen (Figure 20). By moving the cursor around the image, in particular to the very dark and the very light spots, we can roughly estimate the range of grey values used in the image. This approach is better than not checking the image grey values at all, but we really should find a way to consult a histogram of image grey values for adequate evaluation of the quality of the image.



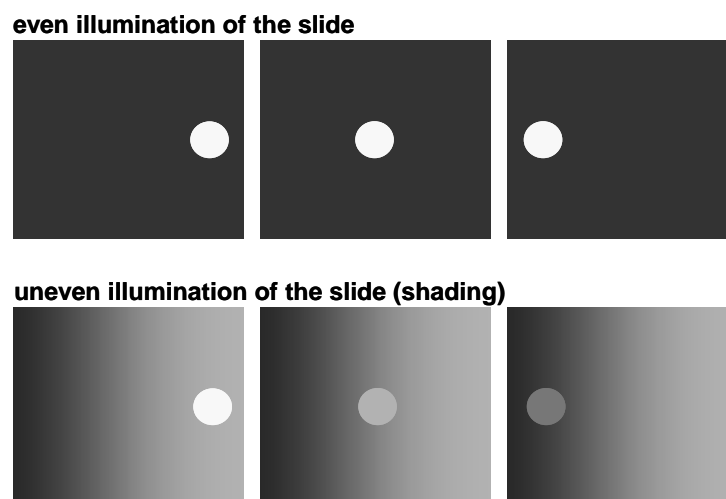
**Figure 20.** Examples of indicators of the pixel grey value. In some software packages, the grey value of the pixel at the current position of the cursor is displayed under the image (A, C) or close to the cursor (B). As we move the cursor around the image, we can get a general impression about the grey value range used in the image. In the shown case, the background is not underexposed (grey value 32 on A; 36 on B) and a pixel in brightest spot in the comet head has the grey value 173 (B).

## 5.5 Shading

Our measurement was based on the assumption that a linear relationship exists between the amount of DNA on the slide and the light intensity that the camera detected. One of the factors that can endanger this relationship is **shading** of the image. We talk about shading when there is a gradient in the intensity of the background signal across the image. We usually get this effect when the whole field of view under a microscope is not evenly illuminated during image acquisition.

The intensity of light emitted by DNA-bound ethidium bromide depends upon the intensity of the green excitation light. Because we only measure the relative amount of DNA, the actual excitation intensity is not so important, as long as we have a strong enough signal for detection. However, **the whole field of view has to be evenly illuminated**, so that all comets and all parts of one comet on an image receive the same intensity of excitation light. We need to optimally align all components of the optical system. In particular, careful centring of the mercury lamp is essential.

When shading occurs, the detected intensity of light emitted from a comet depends on the position of this comet in the field of view, and consequently in the grabbed image. In addition, if the field of view is not evenly illuminated, we can encounter strong gradients of the background signal (Figure 21). The combined effects of shading compromise the linearity and the accuracy of measurement, hence **images with a strong shading effect should not be measured**.



**Figure 21.** Effects of uneven illumination of the field of view. In an evenly illuminated field of view under a microscope, the background in the image is even and the recorded grey values of the comet (shown as a filled circle) are the same for all positions of the comet on the image (**A**). On the other hand, in an evenly illuminated field of view, the background in the image shows a gradient of grey values and the recorded grey values of the comet are different for different positions of the object on the image.

Uneven illumination of the field of view cannot be completely eliminated even in an optimally aligned microscope. Hence, it is good practice to place all measured comets in the same position on the image during image acquisition, close to the centre of the field of view, where shading is least severe in a well-aligned microscope.

A weak shading effect is partially corrected during calculation of the comet parameters. You will remember that to create a comet profile, we calculated the light intensities for each individual column of pixels, taking always into account the local background in each respective pixel column (Figure 10, Table 1). In this way, a small gradient of background grey values running from left to right (along the x-axis) can be corrected. However, severe shading and shading in the direction from top to bottom of the image (along the y-axis) is not accounted for and can lead to erroneous results.

When DNA-bound ethidium bromide is continuously exposed to strong excitation light, the intensity of the emitted signal gradually decreases. Excitation of the comets surrounding the current field of view can be minimised by closing of the diaphragm in the excitation light path. Note, however, that some shading can occur if the edge of the diaphragm aperture is too close to the edge of the field recorded by the camera. So we have to make sure that the diaphragm is not closed too tightly.

## 5.6 How can I detect shading in my image?

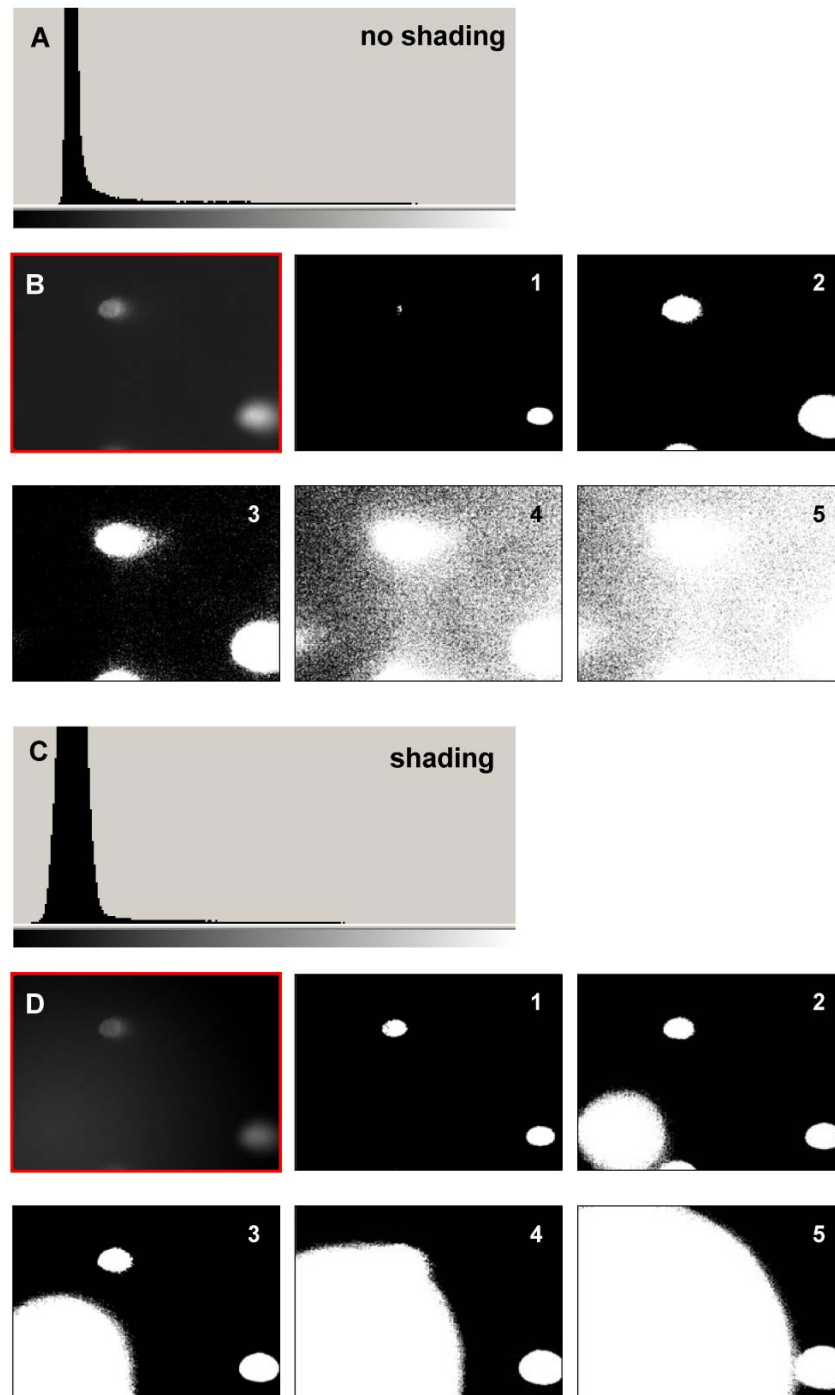
Shading cannot be detected by just looking at an image – our eyes cannot detect gradual changes in the shades of grey in the very dark background. One indicator of shading can be seen in the comet profile. In severe cases of shading (and where the gradient of the background signal runs from left to right), the background line of the comet profile is sloping (without shading, it is horizontal).

To detect shading in an image, we will use the “Levels” and “Threshold” tools of *Photoshop* (see chapter *How can I evaluate the quality of my image?*, page 32 for a description of these tools and explanation about import of a comet image to *Photoshop*). We will analyse an image with an acceptable level of shading (Figures 22A, 22B) and an image with a severe shading effect (Figures 22C, 22D).

One of the indicators of shading is a wide peak of background grey values in the histogram of image grey values (Figures 22A, 22C; *Photoshop*: from main menu Image / Adjust(ments) / Levels).

The best approach for visualisation of shading is application of the “Threshold” tool in *Photoshop* (from the main menu Image / Adjust(ments) / Threshold). If we then drag the slider for threshold grey values under the histogram, we can observe the distribution of grey values in the background. In an image with an acceptable level of shading, the whole background appears even within the range of a few grey values (Figure 22B: 1-5). This image has some shading – background in the left side of the image is slightly darker than

on the right side. In the image with severe shading (Figure 22D: 1-5), the background is clearly much brighter in the lower left part of the image than in the upper right part.



**Figure 22.** Detection of image shading in *Photoshop*. Two cases are shown: an image with a low level of shading (**A**, **B**) and an image with severe shading (**C**, **D**). The analysed images are indicated with a red border. Images with shading have a wider peak of background grey values in the histogram of image grey values (**C**) than images without shading (**A**; “Levels” tool in *Photoshop*). Using the “Threshold” tool of *Photoshop*, as the slider for threshold grey values is dragged from 255 (white) to 0 (black), the background is gradually revealed. Several successive steps of this procedure are indicated with numbers 1-5 in **B** and **D** (decreasing threshold grey values). See text for explanation.



## 5.7 Why a 10-bit camera beats an 8-bit camera

In the comet assay, a camera serves as a detector of light intensity and hence its properties affect the quality of the measurement. There are many different types of cameras out there, and we will only cover some topics essential for understanding the comet measurement.

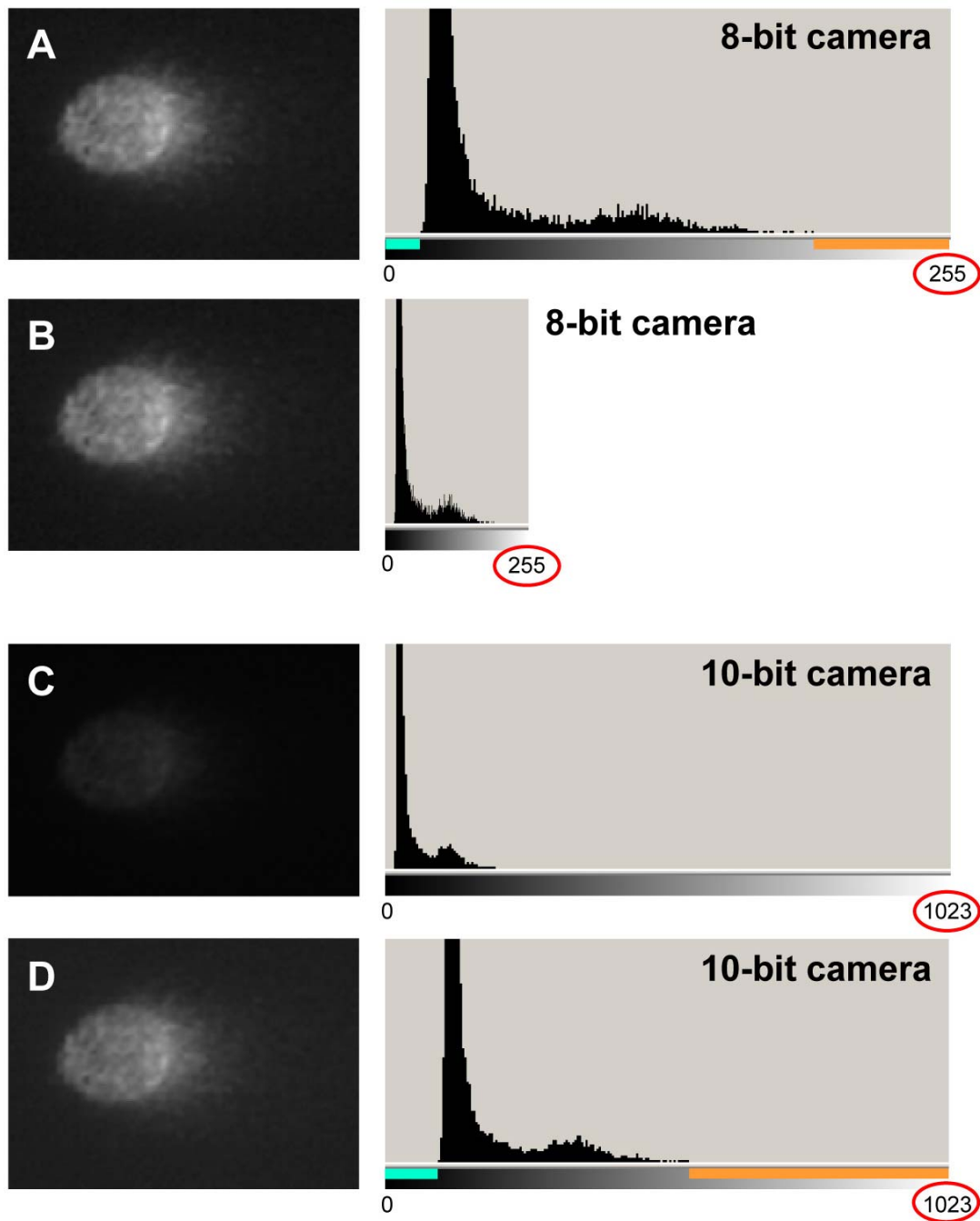
We have video cameras and digital cameras, and both can be used for the comet assay. They differ by the mode of data transfer from the camera to the computer, but this is of no concern here. Then, we have colour and monochrome (“black and white”) cameras. A **monochrome camera** is a better tool for detection of fluorescence – anyway, we are only measuring light intensity, not wavelength (our signal is a narrow band of orange to red wavelengths). The fluorescence signal is very weak, hence the camera has to have a **high sensitivity** – it has to be sensitive to low light intensities. It is also better if a camera has a low noise. Noise is a signal, generated by various components of the camera, which is not related to the true signal from the slide. A camera used for the comet assay must have a **linear response to light intensity**, according to our postulate in Figure 4. Some cameras show considerable deviation from linearity and are not suitable for the comet assay. Cameras also have different **resolutions** – the number of light-sensitive elements in the camera detector.

Finally, cameras differ in the way they record pixel grey values. So far we always assumed that the camera records **256 different grey values**. However, this is only true for an **8-bit camera**. Pixels can also be described with a higher number of grey values: 1024 grey values in 10-bit cameras, 4096 in 12-bit cameras, 16384 in 14-bit cameras and so on. The number of available grey values to describe each pixel is called **pixel depth**.

If your camera has pixel depth higher than 8-bit, then all of the examples described in previous chapters also hold true for your camera, except that wherever we used 256 grey values you have to imagine a higher number of grey values.

Since our eyes detect much fewer shades of grey than 256, we do not need to use pixel depth higher than 8-bit if we are only recording images to look at. However, there are some advantages in using cameras with a higher pixel depth when we are recording images for measurements of light intensity.

Let us compare an 8-bit and a 10-bit camera used for the comet assay. We have either 256 or 1024 available grey values. Let us record an image with an 8-bit camera (Figure 23A). In this image, a large part of the available grey values is used to record the signal from the slide. A 10-bit image has four times more available grey values, so let us scale up the x-axis in the histogram of image grey values to 1024 (Figure 23B). We see that if we record a 10-bit image that contains the *same* amount of information as our original 8-bit image, a very small part of the 10-bit range of available grey values is used (Figure 23C). So when we record a good quality image with a 10-bit camera, we have more grey values available to record the signal, and we can record light intensity with a higher precision (Figure 23D).



**Figure 23.** Comparison of an 8-bit and a 10-bit camera. While an 8-bit image has 256 available grey values (A, B), a 10-bit image has 1024 available grey values (C, D). Hence, a 10-bit image that contains the same amount of information as an 8-bit image uses a small range of all available grey values (C). When adjusting the image acquisition settings at the beginning of comet measurements, we leave some unused grey values to the left (blue) and to the right (orange) of the used grey value range to account for the possible lower background signal and higher comet signal in images that will be recorded later during the experiment, and hence avoid image saturation.

Having a higher pixel depth brings another advantage. When we start recording images for a comet experiment, we optimise the system to acquire images of good quality. We should also not change the settings of image acquisition during recording of one experiment. The problem is that when we are adjusting the settings of the system, we only see a few comets. We do not know whether on other slides we will have comets with much brighter heads than the ones we see. We also do not know whether some slides will have a lower level of background signal than our current slide. As we have seen, our image must never be saturated – it must always contain information on all variation in the intensity of the background and the comet signal. To account for a possible lower background signal and a higher comet signal in the comets that will be scored later, we adjust the camera settings at the beginning of measurement so that we leave some unused grey values lower than the current background values (Figure 23, marked in blue) and some unused grey values higher than the current comet values (Figure 23, marked in orange). With a 10-bit camera this is less problematic, since even if we leave about half of the range of available grey values unused, we are still describing the signal with 500 grey values. Nevertheless, it is possible to score comets with an 8-bit camera, as long as we are careful not to saturate the image.

## 5.8 Binning

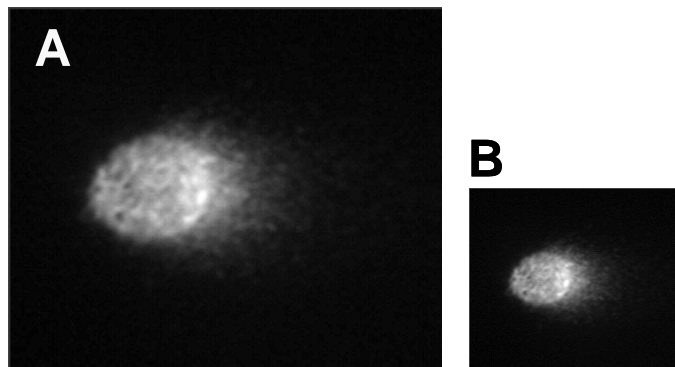
Some comet measurement systems offer an option of **binning** for image acquisition. Binning is a setting of the camera – it specifies how light intensity is detected in the camera detector. Without binning, each light-sensitive element in the camera records the intensity of light, which is then represented as one pixel in the grabbed image. With binning set to  $2 \times 2$ , the camera records the signal from 4 adjacent detector elements as one pixel (Figure 24). With binning  $3 \times 3$ , the signal from nine detector elements is pooled together.

Application of binning has two effects. The **sensitivity of the camera** increases – the camera can detect a weaker light signal. On the other hand, the **resolution of the camera** decreases – the grabbed image uses fewer pixels to record the same field of view, so we see fewer details in the image, the comet ROI contains fewer pixels and the line graph representing the comet profile is constructed from fewer intensity points.

Binning  $2 \times 2$  is usually recommended for cameras with a high resolution. A camera with full resolution  $1300 \times 1030$  pixels records images with size  $650 \times 515$  pixels at binning  $2 \times 2$ . A positive effect of binning for comet scoring is a higher sensitivity of the camera, so that a weak fluorescence signal can be detected. In addition, at binning  $2 \times 2$ , the comet ROI contains 4-times fewer pixels than in the full resolution image without binning (Figure 24), so the computer can do the calculations of measured parameters faster. In practice, we always make compromises between camera resolution and sensitivity – we need to detect

the signal at a high enough resolution to allow for reliable measurement of the comet profile. The optimal setting of binning for recording of comets depends on the properties of the equipment for comet measurement (the full resolution of the camera, the magnification of the objective etc.) and on the size of scored nuclei and comets (smaller comets require a higher resolution of the image, weaker comets require higher sensitivity).

Note that **binning affects the scaling factor of the image** (Figure 24).



**Figure 24.** The effects of binning. Only the part of the recorded image that contains the comet is shown here to illustrate the effects of binning. **A:** The image was grabbed at full resolution of the camera without binning. The size of the image is  $280 \times 230$  pixels and it contains 64400 pixels. The image scaling factor is  $0.36 \mu\text{m}$  per pixel. **B:** The image was grabbed at binning  $2 \times 2$ . The size of the image is  $140 \times 115$  pixels and it contains 16100 pixels. The image scaling factor is  $0.72 \mu\text{m}$  per pixel.

## 5.9 Which image does my software measure?

We have seen that after we grab an image of a comet we can change the way it is displayed on the screen. In some comet software packages, we can adjust the settings of the displayed image, using contrast and brightness tools and applying LUTs. This can help us see the comets and determine where the borders of the comet are in the *displayed* image.

In principle, any comet software should save the original grabbed image of a comet in a very safe place that the user cannot access. After the user confirms the positions of the comet ROIs, the computer should perform the measurements on this original image and on no other derivative of it. The grabbed image contains all available information on the comet, and information can only get lost or distorted by any later fiddling with the image.

However, this is not always the case in practice. The problem arises if our software uses the changed *displayed* image to *measure* the comet parameters, and not the original image. If this is the case, we have to use the display options *very carefully*, and make sure that

after we mark the comet borders, we return the image back to original prior to measurement.

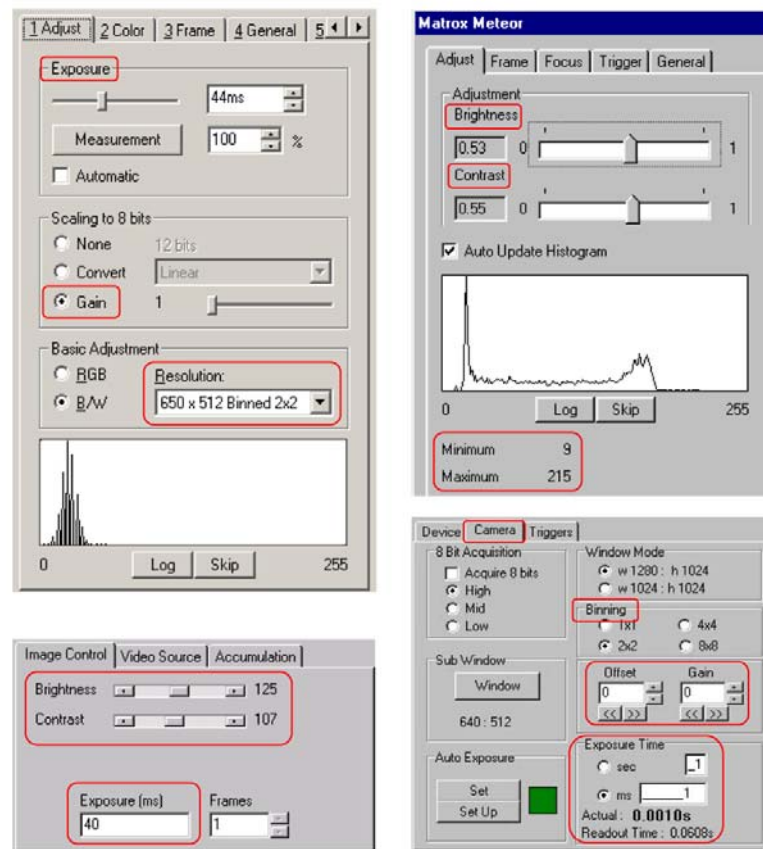
How can we find out how our software actually works? Possibly this is described in the user manual. We can also record a good image of a comet, save it and then import it to the comet software and try to measure this same image several times using drastic changes of the displayed image (e.g. very overexposed, very underexposed). We should keep the borders of the comet constant in all measurements. If the way that the image is displayed on the screen affects the values of the measured parameters, then our comet program does not measure the original recorded image.

To get the best out of the comet assay and avoid measurement errors, we should **know how our particular software works**. There are considerable variations in the exact procedures that different software packages use.

## 5.10 How do I get a good quality image out of my comet equipment?

The system for measurement of comet parameters with image analysis has several elements (a microscope, a camera, software), and all need to be optimally adjusted to get the best possible quality of a recorded image. The general setup of a microscope and possible problems with its alignment have been discussed before (see chapters *Fluorescence microscopy*, page 3, and *Shading*, page 35). Other parameters that determine the quality of the grabbed image are related to the settings of the camera and its control via the computer.

The aim of camera adjustment is to get an image with no saturation and with some unused low and high grey values (see Figure 23). There are many versions of the user interface for adjusting settings of the camera (Figure 25). Some versions even show the histogram of image grey values. The main settings that influence the histogram of image grey values are contrast / brightness or gain / offset and exposure (integration) time. Contrast and gain determine how much the signal is amplified; hence they adjust the range of used grey values. Brightness and offset determine the general brightness of the image – the position of the range of used grey values on the whole scale of available grey values. Exposure time defines for how long the camera collects the light signal – longer exposure times result in brighter images. To get to know the optimal settings of your system, you will need to experiment a little.

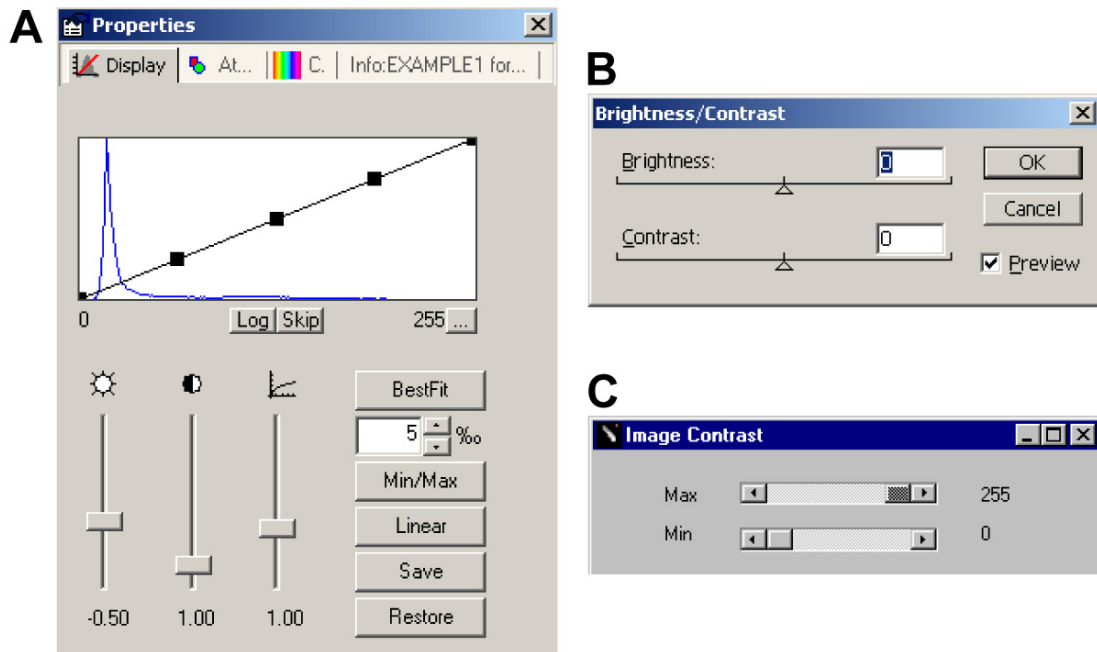


**Figure 25.** Different versions of user interface for adjusting the camera settings. The main settings are contrast / brightness, offset / gain, exposure time, binning.

When we are adjusting the camera and we have no image histogram available, we need to make sure that the image that we see displayed on the screen is a replica of the grabbed image and not changed by any display settings. In particular, we should check which LUT is used (the LUT for display of actual grey values usually has the name “grey” or “greyscale”). Another problem is that not only the camera, but also the display has contrast and brightness settings. To see the actual grabbed image, the brightness and contrast of the *displayed* image should be set to “neutral”(Figure 26). Because the way this is achieved varies in different software packages, we have to check how to set our own software correctly.

For accurate recording of comet images, any automatic corrections of the image that a camera may feature have to be switched off. For example, the setting “gamma”, if switched on (so that the gamma factor is not 1), distorts camera response away from linearity.

**Once the camera options are adjusted, it is best not to change them throughout the experiment.**



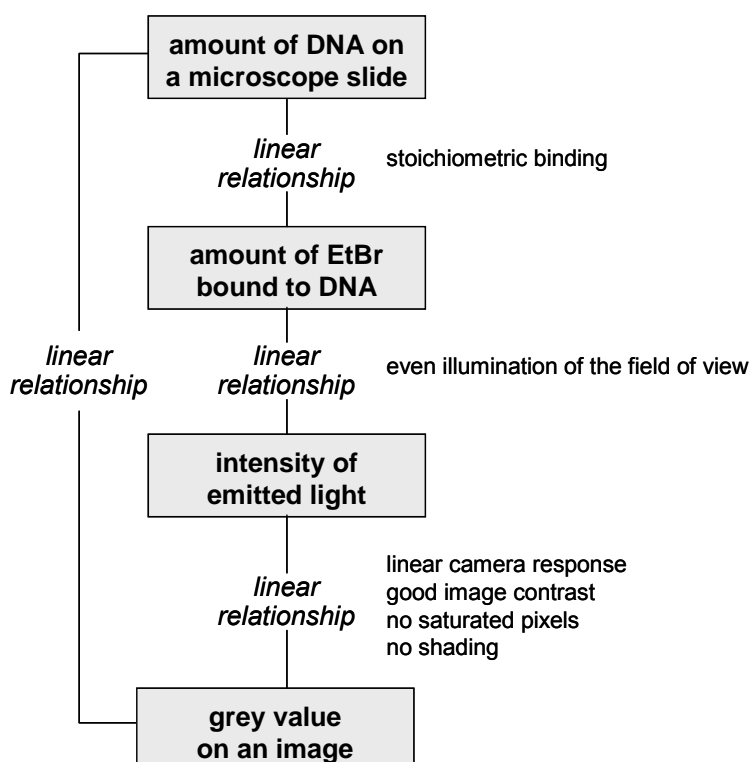
**Figure 26.** Examples of contrast and brightness settings of the displayed image. For adjustment of the camera settings, the contrast and brightness settings of the displayed image have to be set to neutral, so that the actual grey values of the grabbed image are displayed. In example **A**, the neutral settings of the displayed are brightness  $-0.5$  and contrast  $1.0$  (and gamma correction factor of the displayed image  $1.0$ ). In example **B**, the neutral display settings are brightness  $0$  and contrast  $0$ . In example **C**, a neutral contrast setting of the displayed image is achieved when the image contrast is stretched across the whole range of available grey values, from  $0$  to  $255$ .

Our final consideration goes to saving images of comets. If we are not using the “live” mode for comet scoring, we have to record and save images before the measurement. It is good practice to save images of measured comets even if we are working in the “live” mode, to keep a record of the experiment. There are many formats available for saving images (JPEG, TIF, BMP, GIF etc.). Images that are used for image analysis should be stored in the **TIFF format** – this format preserves all information contained in the recorded image, hence the image files are larger than in other formats and use up a lot of computer memory. However, note that images saved in other formats, e.g. in the popular JPEG format, get distorted during compression to make files smaller. Such distorted images should not be used for measurements.

## 6 Concluding remarks

I said “linear relationship” so many times that I would really like to avoid it in my concluding remarks. But, here it comes again, because there really is only one rule for measurement of comet parameters: **Do not violate the linear relationship between the amount of DNA and the grey values on the image of a comet** (Figure 27). Once we have a microscope, a good camera and a reliable software package, we have to make sure that we are **recording images of good quality**. So for each experiment, we have to optimise the settings of the microscope, the camera and any other settings that affect image quality. We also have to select the right geometric calibration (image scaling factor). Only then can we start to click.

When I started writing, I had no idea that I will need to explain so many things. I hope that after you went through all the trouble of reading this lengthy paper, your understanding of what your computer does to your comets has increased from grey value 0 (a completely incomprehensible black box) to at least grey value 113 (a slightly mysterious grey box).



**Figure 27.** Measurement of the comet parameters: an overview.



## Notes

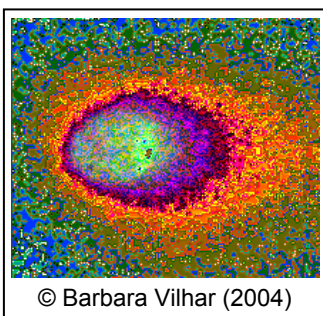
*I learned about image analysis and the comet world from many people, and I feel indebted to them all. In particular, I wish to thank Dr Aleš Kladnik for inspired image analysis debates and for helping me survive computer crashes. Janja Plazar and Irena Žnidar drew my attention to various “comet mysteries”. Janja Plazar also recorded the image of a comet used as an example throughout this work.*

*Adobe Photoshop is a trademark of Adobe Systems Inc. Images of windows from the following software packages are used as illustrations: Photophop (Adobe Systems), Komet 5 (Kinetic Imaging), KS400 (Carl Zeiss Vision) and AxioVision (Carl Zeiss Vision).*

*A hardcopy of this work is deposited at the Biology Library, Večna pot 111, SI-1000 Ljubljana, Slovenia.*

*Please cite this work as: Vilhar B (2004). Help! There is a comet in my computer!  
[http://botanika.biologija.org/exp/comet/comet\\_guide01.pdf](http://botanika.biologija.org/exp/comet/comet_guide01.pdf)*

*This work should not be distributed or reproduced without permission from the author.*



© Barbara Vilhar (2004)

**Dr Barbara Vilhar**  
University of Ljubljana  
Biotechnical Faculty  
Department of Biology  
Večna pot 111  
SI-1000 Ljubljana  
Slovenia  
email [barbara.vilhar@bf.uni-lj.si](mailto:barbara.vilhar@bf.uni-lj.si)  
<http://botanika.biologija.org>

