

## Spatial quantitative analysis of fluorescently labeled nuclear structures: Problems, methods, pitfalls

O. Ronneberger<sup>1§</sup>, D. Baddeley<sup>2§</sup>, F. Scheipl<sup>3</sup>, P. J. Verveer<sup>4</sup>, H. Burkhardt<sup>1</sup>, C. Cremer<sup>2</sup>, L. Fahrmeir<sup>3</sup>  
T. Cremer<sup>5,6</sup> & B. Joffe<sup>5\*</sup>

<sup>1</sup>Department of Pattern Recognition and Image Processing, University of Freiburg, 79110, Freiburg, Germany;

<sup>2</sup>Applied Optics and Image Processing Group, Kirchhoff Institute of Physics, University of Heidelberg, 69120, Heidelberg, Germany; <sup>3</sup>Department of Statistics, Ludwig-Maximilian University, Munich, Germany;

<sup>4</sup>Max Planck Institute of Molecular Physiology, 44227, Dortmund, Germany; <sup>5</sup>Department of Anthropology and Human Genetics, Ludwig-Maximilian University, Grosshaderner Str 3, 82152, Munich, Germany;

Tel: +49-89-2180-74332; Fax: +49-89-2180-74331; E-mail: boris.joffe@lrz.uni-muenchen.de;

<sup>6</sup>Munich Center for Integrated Protein Science, 81377, Munich, Germany

\*Correspondence

§These authors contributed equally.

**Key words:** confocal microscopy, image quantification, nucleus, preprocessing

### Abstract

The vast majority of microscopic data in biology of the cell nucleus is currently collected using fluorescence microscopy, and most of these data are subsequently subjected to quantitative analysis. The analysis process unites a number of steps, from image acquisition to statistics, and at each of these steps decisions must be made that may crucially affect the conclusions of the whole study. This often presents a really serious problem because the researcher is typically a biologist, while the decisions to be taken require expertise in the fields of physics, computer image analysis, and statistics. The researcher has to choose between multiple options for data collection, numerous programs for preprocessing and processing of images, and a number of statistical approaches. Written for biologists, this article discusses some of the typical problems and errors that should be avoided. The article was prepared by a team uniting expertise in biology, microscopy, image analysis, and statistics. It considers the options a researcher has at the stages of data acquisition (choice of the microscope and acquisition settings), preprocessing (filtering, intensity normalization, deconvolution), image processing (radial distribution, clustering, co-localization, shape and orientation of objects), and statistical analysis.

### Abbreviations

CLSM	confocal laser scanning microscope	HSA	<i>Homo sapiens</i> , abbreviation used for human chromosomes
FFT	fast Fourier transform	ICA	intensity correlation analysis
FWHM	full width at half maximum (parameter characterizing the width of a peak in a curve)	ICQ	intensity correlation quotient
		MRP	median radial position
		NA	numerical aperture of an objective

---

### Electronic supplementary material

The online version of this article (doi:10.1007/s10577-008-1236-4) contains supplementary material, which is available to authorized users.

PALM	photoactivated localization microscopy
PCA	principal component analysis
PDM	product of the differences from the mean, a parameter used in ICA
PSF	point spread function
SNR	signal-to-noise ratio
STED	stimulated emission depletion microscopy
STORM	stochastic optical reconstruction microscopy
TIRF	total internal reflection fluorescence

## Introduction

Recent progress in nuclear biology has shown an inherent connection between the spatial organization and function of nuclei. Spatial arrangements of chromatin loci in the nucleus have been considered as critically important in transcriptional regulation (Fraser & Bickmore 2007, Lanctot *et al.* 2007, Sexton *et al.* 2007, Soutoglou & Misteli 2007). Chromosome territories and chromosomal subregions have been shown to have non-random radial nuclear distribution, i.e., a more central or more peripheral location in the nucleus itself (Croft *et al.* 1999, Cremer *et al.* 2001, 2006, Bolzer *et al.* 2005, Kosak *et al.* 2007, Neusser *et al.* 2007). There are indications that some chromosome territories also have a non-random localization with respect to their neighbors that persists from one cell cycle to the next (Parada *et al.* 2002, Gerlich *et al.* 2003, Walter *et al.* 2003, Thomson *et al.* 2004, Kosak *et al.* 2007). Individual genes, depending on their transcriptional status, also tend to take specific positions in the nucleus (Kosak & Groudine 2004, Ragozcy *et al.* 2006, Chuang & Belmont 2007, Dundr *et al.* 2007) and, in particular, can associate with nuclear pores (Taddei *et al.* 2006, Akhtar & Gasser 2007). Transcriptional activity of genes also correlates with their positions in chromosome territories or possibly on loops expanding from them (Mahy *et al.* 2002, Chambeyron & Bickmore 2004, Kupper *et al.* 2007) and proximity to centromeric heterochromatin (Cobb *et al.* 2000). Several recent studies have discussed the importance of transient contacts between loci situated on the same chromosome or on different chromosomes (Spilianakis *et al.* 2005, Brown *et al.* 2006, Lomvardas *et al.* 2006, Fraser & Bickmore 2007, Simonis *et al.* 2007), as well as association of loci with transcription factories (Osborne *et al.* 2007), speckles (Shopland *et al.* 2003, Brown *et al.* 2006), and other nuclear bodies (Dundr *et al.* 2007).

Although molecular methods for studying the positioning of DNA sequences have recently been developed (Simonis *et al.* 2006, Zhao *et al.* 2006, Hagege *et al.* 2007), they demand a large number of nuclei from a single cell type and at a given stage of the cell cycle, postmitotic terminal differentiation, or physiological state and cannot replace the analysis of single cells in tissues. The development of new molecular technologies has thus stimulated interest in microscopic studies of nuclear biology (Murray *et al.* 2007, Shiels *et al.* 2007).

The vast majority of microscopic data in cellular and nuclear biology are now collected using fluorescent microscopy, and most of these data are then analyzed using quantitative methods. Each step of this work, from image acquisition to statistical analysis, may crucially affect the conclusions of the whole study: therefore expertise in the fields of physics, computer image analysis, and statistics is necessary. It is hoped that this article will help biologists to cope with this problem. It was written by a team uniting expertise in biology, microscopy, image analysis, and statistics. We consider the methodical options a researcher has at different stages of the work. For the methods discussed, we briefly explain the essence of the method (where necessary), discuss how to use this method correctly, and then consider the typical problems for which this method is useful and note the errors that may occur when using the method. When discussing image acquisition and preprocessing, we have tried to explain the available options in a language clear to biologists. The part dealing with image analysis also includes explanations of important established methods but concentrates on methods that have just started to be applied in nuclear biology, as well as on some original new approaches designed to solve problems that have only very recently been raised by biological studies. Although the focus of this article remains nuclear biology, the methods discussed are applicable to many other fields of cell biology.

## Imaging and image analysis: Stages of a single process

Biologists often consider the recorded images as a 'starting point' for the analysis which results in some measured quantities such as object sizes or positions. In this article we suggest a more general view. The

raw images are only an intermediate representation of the data in a long processing pipeline. This pipeline accumulates all effects that occur before the digital image arrives in the computer. Understanding all of these effects will help to design better experiments and to decide which image processing steps (and which parameters) can improve the validity and reproducibility of quantitative image analysis.

The information contained in the raw image depends on various parameters such as (1) illumination light; (2) the orientation and position of the object under the microscope, if the object is anisotropic; (3) interaction of the light with the object, e.g., fluorescence (and autofluorescence), absorption, reflection, refraction, diffraction; (4) the ‘optical transformation’ of the emitted light in the microscope, resulting in magnification, blurring of the image according to the point spread function (PSF), and filtering of certain wavelengths; (5) recording of the light intensity—for each pixel in the case of 2D evaluations of single optical sections or for each voxel in the case of 3D evaluations of entire image stacks—and conversion of the light intensities into electric charges and the conversion of these charges into digital numbers, e.g., to gray values from 0 to 255. It is therefore necessary to ensure that variation in these parameters does not affect the final—biological—results of the observations.

Conditions important for this, first of all, are:

- Proper calibration of the system (e.g., homogeneous illumination, reproducible setup of all microscopic parameters)
- If necessary, preprocessing of the raw images to remove (or reduce) alterations induced during image acquisition (e.g., compensation for light absorption in thick samples or for bleaching)
- Selection of measurements that are invariant to arbitrarily chosen parameters of experimental and analytical procedures and/or estimating the errors associated with them
- Adequate statistical analysis of the data obtained.

### Which microscope for which task

When choosing the type of a microscope, one must consider several factors, the most important of which

are resolution, sensitivity, rejection of out-of-focus signals, photo-induced specimen damage, and speed. Table 1 summarizes these parameters for several types of microscopes. When talking about resolution, we mean the smallest distance at which two features in an image are seen as separate objects. This is a function of the microscope optics and the wavelength of light used, and should not be confused with the popular use of ‘resolution’ to describe the number of pixels on a CCD camera chip. The sensitivity is a function of both microscope optics and the detector system. A good sensitivity will typically allow small amounts of label (potentially even single fluorophore molecules) to be detected and minimize the amount of photo-damage to the specimen during image acquisition. Systems with the smallest number of optical components, given a good detector, tend to have the best sensitivity.

For sensitivity, it is thus hard to beat a widefield microscope equipped with a good CCD camera. A major disadvantage of the widefield microscope, however, is the lack of optical sectioning. This means that, rather than being rejected, light from out-of-focus objects is simply spread out over a larger area. This is a significant problem when trying to extract 3D information from extended objects and some form of optical sectioning is thus often desirable. The established way of performing optical sectioning is to use a confocal laser scanning microscope (CLSM), where the sample is illuminated point by point. As only one point is illuminated at one time, a pinhole and a photomultiplier can be substituted for the CCD detector. This combination of spatial selectivity in both excitation and detection gives good rejection of out-of-focus light (for a manual on confocal microscopy see Pawley (2006).

A significant disadvantage in classical confocal systems is that they are comparatively slow, even though modern instruments have greatly gained in speed. Although modern confocal systems are sufficiently quick for routine studies of fixed material, their speed is still not sufficient for the *in vivo* imaging of structures that move and/or change shape quickly. Spinning disk confocal microscopes greatly increase the speed of scanning at the expense of a little resolution. Modern spinning disk systems equipped with em-CCD technology and micro-lens arrays are often more sensitive than their beam-scanning counterparts. New structured illumination techniques such as OMX (Gustafsson 2000, Carlton 2008) also offer optical

Table 1. Typical characteristics of several microscope types

Type	Resolution $x$ - $y$ (nm) <sup>a</sup>	Resolution $z$ (nm) <sup>a</sup>	Sensitivity	Photodamage	Optical sectioning	Speed
Widefield	250	650	Good	Good	None <sup>b</sup>	Good
Confocal	210	550	Fair	Fair	Good	Poor
Two-photon	250	650	Poor	Out of focus: good In-focus: poor	Good	Poor
Spinning disk	230	600	Fair	Fair	Fair	Good
TIRF	250	<200 <sup>c</sup>	Good	Good	Good <sup>c</sup>	Good
Structured illumination <sup>d</sup>	130	350	Good	Good	Can be very good	Fair
4Pi-A (two-photon)	220	120	Poor	Poor	Very good	Very poor
STED <sup>e</sup>	90 (16)	550 (33)	Poor	Poor	Good (very good)	Very poor
PALM/STORM <sup>f</sup>	<30	~60	Good	Fair	None/fair	Very poor

<sup>a</sup>Values are 'typical' values for a well adjusted system, and thus slightly worse than the theoretical values. When commercial systems are available, the values for these are given, rather than the best laboratory results.

<sup>b</sup>A good approximation to sectioning can be obtained for objects that have a constrained lateral extent by using deconvolution (see the deconvolution section in this paper).

<sup>c</sup>Whether TIRFs ability to constrain imaging to a small region adjacent to the coverslip can really be considered  $z$ -resolution or sectioning is moot.

<sup>d</sup>Based on figures for OMX (P. Carlton, personal communication, and Gustafsson 2000, the latter for lateral resolution only).

<sup>e</sup>Best laboratory results (Dyba & Hell 2002, Westphal & Hell 2005).

<sup>f</sup>PALM/STORM is new and rapidly moving field, and these values are likely to change soon. Some sectioning ability has been shown through the use of either TIRF (Betzig *et al.* 2006) or 2-photon photoactivation (Fölling *et al.* 2007). The newest advance—true  $z$ -resolution (Huang *et al.* 2008)—has yet to be combined with optical sectioning. Several efforts are underway to increase the acquisition speed (e.g. Geisler *et al.* 2007), and although the acquisition speed for highly resolved images is still slow compared with other forms of light microscopy, this has not prevented the technique from being used to good effect *in vivo* (Hess *et al.* 2007, Manley *et al.* 2008).

sectioning, and an additional resolution increase, while retaining most of the advantages of traditional widefield techniques.

Two-photon confocal microscopy (Denk *et al.* 1990) builds on confocal microscopy by using a pulsed infrared laser source and the 2-photon absorption effect to excite the fluorescence. Two-photon excitation has the advantage that the probability of exciting a fluorophore is negligible anywhere other than in the focus of the laser. Owing to the longer wavelength, scattering effects are also reduced and 2-photon microscopy finds its most common applications in imaging deep within thick biological specimens.

If the resolution obtained using widefield, confocal, or structured illumination techniques is not adequate, one might wish to use a technique such as 4Pi (Hell & Stelzer 1992, Hell *et al.* 1994), STED (Klar *et al.* 2000, Willig *et al.* 2006, 2007), or PALM/STORM (Betzig *et al.* 2006, Hess *et al.* 2006, Rust *et al.* 2006). While these advanced techniques are not necessarily intrinsically more damaging than other methods, achieving the same signal-to-noise ratio over a smaller region requires a larger overall number of photons and it is probably fair to say that a

brighter and more stable labeling is required. These methods are also typically slower and more sensitive to effects such as sample-induced aberration.

The resolutions given in Table 1 are for high NA (63 $\times$  or 100 $\times$ ) oil-immersion objectives as these offer both the best resolution and the best sensitivity. For quantitative work in fixed specimens, these objectives should be used if available. Glycerol objectives have a long working distance and are optimal for thick samples (see, e.g., Martini *et al.* 2002). In living cells, water immersion lenses (also 63 $\times$  or 100 $\times$ ) with a slightly lower NA and accordingly lower resolution are usually used. These offer a better match to the sample refractive index and therefore better imaging deep in the sample.

In short (considering the instruments that are currently in the market), for simple work not needing 3D resolution, for work with weakly labeled specimens, and for most *in vivo* work, the widefield microscope remains the microscope of choice. For imaging where 3D information is important, a standard confocal is a good option for fixed cells and a spinning disk confocal is good for *in vivo* work. TIRF (total internal reflection fluorescence) and

2-photon microscopy are useful for membrane objects and thick specimens, respectively.

## Image acquisition

### Sampling

Once one has decided which microscope to use, several aspects must be considered when taking the actual images. The most important of these is probably sampling. In order to correctly recover all the information in the image, the effective voxel size must be small enough that the smallest possible features are properly sampled. Failure to sample properly results in the loss of resolution, as well as the introduction of aliasing artifacts (spurious signals appearing, and/or small signals disappearing completely). Sampling theory (often called the Nyquist, or Nyquist–Shannon theorem) states that the sampling frequency must be greater than twice the highest frequency contained in the signal (a factor of  $\sim 2.3$  is often used in signal processing). When applied to imaging, this corresponds to a constraint on the maximum voxel size, namely that the voxel size must be less than half the smallest possible feature size. The smallest possible feature size is usually equivalent to the resolution (depending on the definition of resolution being used). For normal confocal imaging, voxel sizes less than or equal to  $80 \times 80 \times 200$  nm are usually perfectly acceptable.

Owing to the low speed of confocal microscopes, biologists have often tended toward under-sampling (voxel size too large for the objective used). Modern confocal instruments have a much higher speed that solves this problem. If high resolution is not necessary, one can rather use an objective with smaller magnification and gain in the size of the field of view and the depth of focus. While under-sampling is generally inexcusable when the images are going to be subjected to quantitative analysis, a small amount of over-sampling (voxel size smaller than necessary) is acceptable. Over-sampling followed by averaging can in some circumstances be used to improve the detection dynamic range (e.g. 4Pi with avalanche photodiode (APD) detectors). Excessive over-sampling makes data acquisition slower than necessary, increases photobleaching as well as photodamage *in vivo*, and results in excessive data volumes.

### Signal-to-noise ratio and dynamic range

The second most important aspect of image acquisition, with reference to subsequent image processing, is the signal to-noise-ratio (SNR) and the dynamic range. Both these quantities should be maximized, although constraints on labeling stability and acquisition time will normally require some form of compromise. SNR is the squared ratio of characteristic intensities of signal and noise (the standard deviation of signal intensity is also important in this context). SNR can be increased by increasing illumination intensity, or integration time, or by averaging. Dynamic range is the range of discrete signal levels available in the image data. Different detectors have different dynamic ranges, for example, confocal microscopes manufacturers often suggest an 8-bit (256 values) dynamic range, though modern instruments also allow 12 bits or more, whereas CCD cameras more often use 12 bits (4096 values) or even 16 bits (65 536 values). A higher dynamic range will result in larger file size. In any case, one should try to make the best use of the available dynamic range. In practice this means choosing the laser power and/or photomultiplier voltage (in the case of confocal microscopy) or integration time (for CCD cameras) such that the maximum signal value (i.e., intensity in the brightest portions of the studied sample) is around 80% of the available dynamic range (to avoid clipping, see below).

One should note that, for confocal measurements, increasing the photomultiplier gain also increases the noise. Therefore, given a sufficiently photostable labeling, increasing the laser power is preferable to increasing photomultiplier gain from a signal-to-noise standpoint (the same argument applies for modern electron-multiplying CCD cameras; an electron multiplication gain that is too high can be detrimental to the overall signal-to-noise level). While increasing illumination intensity is normally safe in widefield imaging, increasing the laser power must be approached with caution when using a confocal microscope as there is a very real risk that the fluorophores will be driven into saturation, resulting in a dramatically increased rate of photobleaching for a smaller than expected gain in signal. Averaging or accumulation of several repeated acquisition steps, an option always provided by confocal software, is thus one of the main means for increasing SNR, though photobleaching and, in case

of live observations, the acquisition time should also be taken into account. Both line averaging and frame averaging are normally offered; which one is preferable varies from microscope to microscope. Stack averaging is also possible but may cause misalignment of the channels in instruments with mechanical control of the stage position. In most cases it is preferable to set the direction of stack acquisition so that the focus plane moves towards the objective lens. In this case the deepest layers of the sample that lose more emitted signal are imaged first, when bleaching is minimal.

One common problem encountered with microscopic imaging is clipping or saturation. This occurs when the magnitude of a signal exceeds the available dynamic range and results in out-of-range pixels being assigned the maximum/minimum possible value and, therefore, all information contained in the real intensity values of these pixels being lost. Clipping at the top end of the range (saturation) is typically caused by either the photomultiplier gain (confocal), integration time (CCD), or illumination power being too high. At too high illumination intensities (note that such intensities can be achieved with moderate laser power when using a beam scanning confocal), saturation of the fluorescence transition can occur, leading to similar problems. Clipping at the bottom of the range is typically due to an incorrect photomultiplier offset setting and is usually restricted to confocal modalities. Image acquisition software of both confocal and widefield systems usually allows additional manual setting of the offset value, that is, setting all gray values below the offset to zero. This setting reduces background noise seen on the screen and written to the file. Such setting is nothing other than a threshold that also clips low-intensity signals. It should therefore be used with care and kept to minimum; some background signal should be retained (see also Flat field correction and background subtraction, below). Most modern systems have lookup tables which use contrasting colors (e.g. green/blue pixels in the Leica 'glow' colormap) to indicate clipping—it is very advisable to use such lookup tables to choose acquisition parameters.

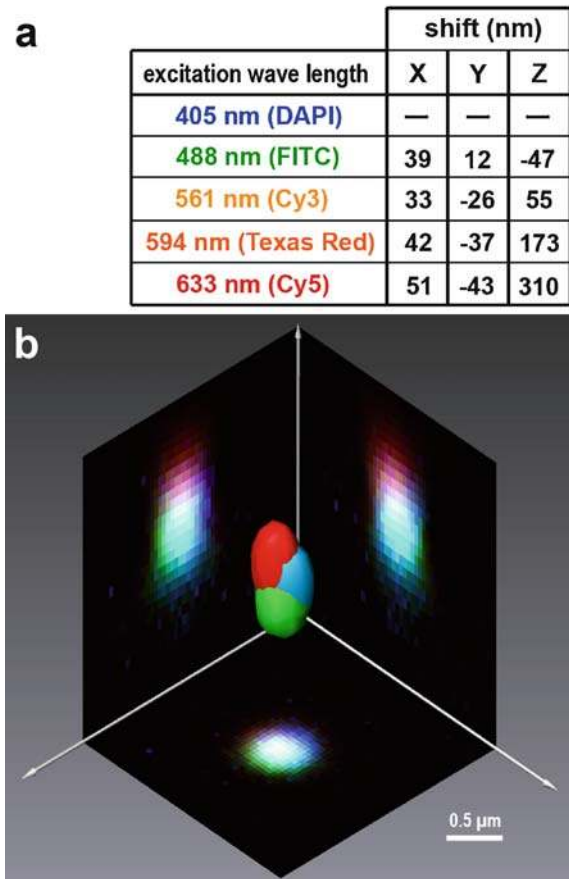
Several image acquisition packages offer additional postacquisition steps, for example, correction for bleaching and variations in the excitation intensity. While useful in principle, such features should be approached with caution, and only used when one

is aware of all the assumptions involved and is confident that they are met. The bleaching/excitation power correction, for example, is usually based on the assumption that the integrated intensity in all slices should be the same, a condition that is satisfied only very rarely. We generally advise collecting images without any corrections and correcting them, if appropriate, later (see Preprocessing, below).

#### *Chromatic shift: measuring and correcting for it*

A systematic error that is almost always present in microscopic data is chromatic shift: light of different colors will be focused at slightly different positions. The effect is particularly relevant along the  $z$ -direction, where the chromatic shift is typically worse than in the  $x$ - $y$  plane. The amount of chromatic shift depends on the microscope optics and on the optical conditions within the sample itself. Typical values of chromatic shift for a confocal microscope using a good high-NA oil objective are of the order of 10–20 nm laterally and 100–200 nm axially, worsening considerably with increasing depth into the sample (Figure 1). For poorly adjusted systems and/or thick specimens, shifts of more than 100 nm laterally and 500 nm axially are not uncommon. There is also some variation across the field of view (although this is of a lesser magnitude, and only really important for very precise distance measurements). These chromatic shifts will lead to errors, particularly in high-precision distance measurements in 3D or in co-localization analysis (see Co-localization, below). Luckily it is possible to correct for them. The first step is to measure the shifts, which can be done using multicolored fluorescent beads (see Walter *et al.* 2006 for a detailed protocol). Importantly, shift should be measured for each optical path used. As the shift is sensitive to temperature as well as any changes in instrument alignment (even at the level which would be induced by, e.g., removing and replacing the objective), measurements should be performed regularly. How the correction is done then depends on the analysis being performed. If measuring distances, it is trivial to add/subtract the measured 3D shifts to each of the components of the measured distance vector before calculating the absolute distance.

A common technique to obtain *an image* coarsely corrected for the  $z$ -chromatic shift (rather than corrected coordinates of certain points in the image as



*Figure 1.* Chromatic shift. (a) Images of Tetraspec beads (0.5  $\mu\text{m}$ ) were acquired with a Leica SP5 confocal microscope using excitation wavelengths and emission filters optimized for the fluorochromes shown in the table. High-quality modern objective lenses compensate for chromatic shift for the wave lengths corresponding to visible part of the spectrum (FITC to Texas Red). The 405 nm/DAPI channel in modern microscopes may usually be fitted to the other channels at hardware level (adjustment should be performed regularly by a professional). Maximal axial shift is therefore observed for the 633 nm/Cy5 channel. The table shows mean shifts in relation to the 405 nm/DAPI channel calculated from images of 10 beads. (b) A 3D Amira reconstruction of a bead shows the effect of chromatic shift for 405 nm/DAPI, 488 nm/FITC and 633 nm/Cy5 channels.

discussed above) is to cull slices from the top or bottom of the stack, effectively moving the channels by an integer multiple of the voxel size with respect to each other. While this technique is often adequate for removing the worst of the visual aberration, it is usually not sufficient for precise distance measurements or precise co-localization studies. If one wishes to obtain a fully corrected image, the

procedure is a little more complex. One must resample the channels, interpolating to the correct voxel positions.

### Preprocessing

The main goal of preprocessing is to reconstruct the true fluorophore distribution as well as possible. This includes compensation (or at least reduction) of random and systematic errors that are caused by the imaging process. A typical random error is intensity noise. A typical systematic error is the point spreading caused by the objective. Below we provide a short overview of several of the most commonly used variants of preprocessing, their application areas, and the situations in which they should be avoided. It is crucial to understand which preprocessing may be applied to which experimental setup and how it influences the measurements carried out on the image. Changes induced by preprocessing are an important reason for differences between results obtained from the same biological material. Preprocessing is essentially calculating modified intensities; more complex calculations usually result in a stronger propagation of errors (see e.g., Wolf *et al.* 2007 for simple illustration of this propagation) and inappropriately applied preprocessing (or inadequately selected parameters) will unpredictably bias the results. In short, everything that is not really needed should be avoided.

#### *Flat field correction and background subtraction*

Flat field correction (also called ‘shading correction’) is typically unnecessary for confocal microscopy, which is of primary interest for this article, but we will briefly discuss it because it may be of crucial importance for widefield microscopy. It corrects for systematic errors of the imaging process such as bias and non-uniformity of illumination, or those caused by the optics, the CCD-sensor, or by the conversion of the electrical signal to gray values. Flat field correction is usually performed by imaging of a calibration sample with a uniform intensity and then computing the gray value offset and scale factor for each pixel; the latter are used to correct intensities in other images. General non-uniformity of illumination is often the case in widefield microscopy, especially when people try to adjust for maximum brightness

going away from the optimal Koehler illumination. In the typical single-cell analysis setup (the cell is near the optical axis and it spans only a small portion of the maximal field of view possible with the used lens) and with a well-adjusted microscope, high-end CCD sensors, and no dust particles in the optical path, such non-uniformities are usually negligible.

While a non-uniform background can usually be avoided, there are several factors causing a constant background throughout the image and, correspondingly, a constant gray value offset ('additional' intensity) for all pixels. Here we address the offset caused by the image acquisition system that does not depend on sample and can therefore be determined prior to imaging of the samples studied. Correct determination and subtraction of this offset is essential for all algorithms that strongly rely on the proportionality of gray value to the light intensity (e.g., fluorescence resonance energy transfer (FRET), and deconvolution). The best way to determine the background gray value is to record two calibration images (A and B) of the same (non-bleaching) sample, where the shutter time for the second image (B) is halved. The background pixel gray value  $g$  is then computed as  $g = \text{mean}(2B - A)$ , which is easy to perform, e.g., using the popular free image processing software ImageJ (this software may be downloaded from ImageJ website: <http://rsb.info.nih.gov/ij/>; a convenient installation of ImageJ for Windows with a useful selection of plugins is also available at the WCIF website: <http://www.uhnresearch.ca/facilities/wcif/imagej/index.htm>). Furthermore, to avoid clipping of low intensities, the offset in the analog-digital converter (which converts electrical signals to gray values) is typically set so that zero intensity corresponds to some low positive value. This produces a constant 'background' which should be determined as described above and subtracted from the pixel gray values of the images. Slightly negative gray values may appear in the image after background subtraction, owing to positive and negative contributions of the electronic noise in the sensor. Even though negative intensities do not exist in reality, for many image analysis processes (e.g., all linear and most of the nonlinear filters, in particular, for deconvolution, etc.), they should be retained and it is therefore important to use a proper data type for the resulting image: the 16-bit integer (which can store integer gray values between  $-32\,768$  and  $+32\,767$ ) or 32-bit float (which can store arbitrary gray values, e.g. 42.3

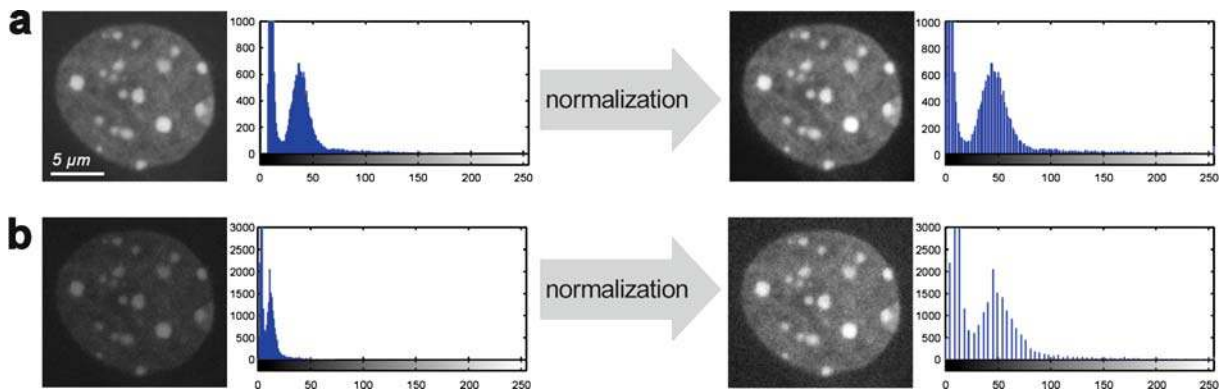
or  $-7.5$ ). In addition to hardware offset setting as described above, image acquisition software of both confocal and widefield systems allows the user to manually increase offset for images shown in the screen and saved to files. Quite clearly such online thresholding should be kept to minimum (see also Signal-to-noise ratio and dynamic range).

- Software for image analysis that is strongly dependent on the proportionality between pixel gray values and light intensity usually performs background subtraction based on the image itself, sometimes invisibly for the user. It should be taken into account that some such programs do not handle negative intensities: in this case background subtraction should not be done.
- Background subtraction is also unnecessary if the next processing step includes manual thresholding or if only the differences between gray values are used in the further processing.

#### *Intensity normalization*

The goal of intensity normalization is usually to correct for systematic errors of the imaging process that vary from experiment to experiment and can therefore not be determined by a prior calibration. This includes, for example, varying illumination along the depth of the sample, absorption, bleaching due to image acquisition, etc. Such corrections are primarily important for fully automated systems to ensure the reproducibility of results. In nuclear studies, where CLSM images of individual cells are usually processed separately with interactive tools and thresholds are set manually, intensity normalization is usually not needed. It may be used to facilitate visual inspection, but for further processing it is preferable to preserve the original gray values. On the other hand, intensity normalization may improve images for further processing if the real intensity distributions of the images of different experiments are nearly identical, except for a linear scaling of the gray values between images (e.g. due to bleaching of the fluorophore, Figure 2). These conditions are fulfilled, for example, if the same object is recorded at different times. Therefore, intensity normalization may be very important for *in vivo* time-lapse studies. Intensity normalization may also be useful for visual analysis of the co-localization of signals.





**Figure 2.** Intensity normalization. (a) Confocal image (8-bit format) of a mouse fibroblast nucleus stained with T0-PRO3 to show chromocenters. (b) The same image after simulated bleaching. The blue graphs show intensity histograms. Owing to nearly proportional fluorophore intensities in these images (as is also the case for natural bleaching e.g. due to image acquisition), intensity normalization improves both images. After this normalization, corresponding structures have similar gray values. This allows, for example, the same threshold to be applied to all images of the series.

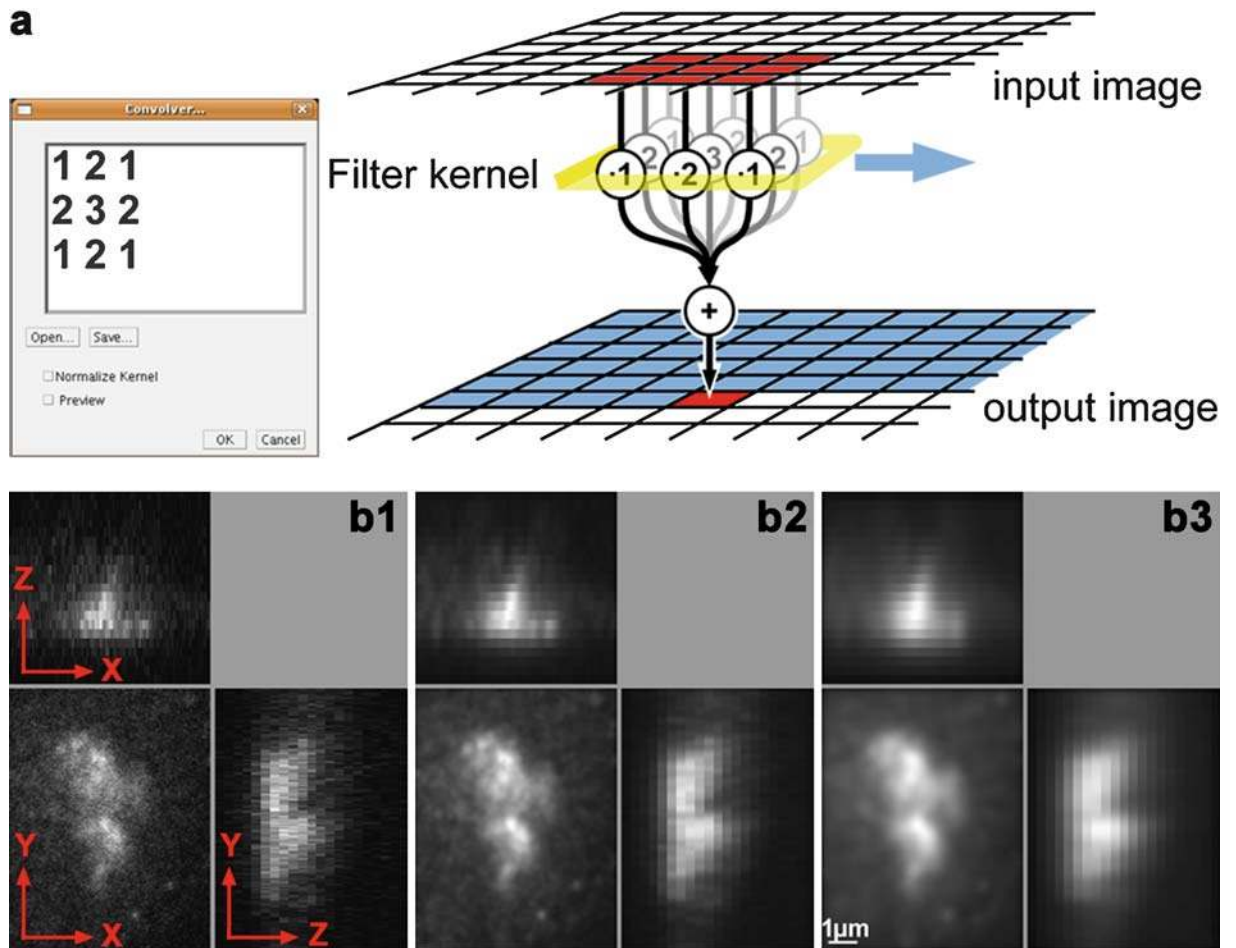
A very common normalization procedure is to use the minimum and maximum gray value of the image, to assign them the values of 0 and 255 (or another arbitrarily selected maximal gray value), and change other intensities proportionally. With this procedure no information is lost, but the result will be strongly affected if image contains outliers (very dark or very bright pixels that are not part of the considered structure). To avoid the effect of outliers, one can set to 0 and 255 a small percentage of darkest and brightest pixels (an option suggested by the ImageJ software). A more robust alternative is to use the median and the interquartile range, so that the normalized image has a median of, e.g., 100, and an interquartile range of, e.g., 80 (of course, the ratio of the 2nd and 3rd quartiles should be retained). A danger of the two last-mentioned approaches is a reduction in the dynamic range due to creation of over- and undersaturated pixels, and the loss of information due to the requantization of the gray values (e.g. a scaled gray value of 42.3 will be rounded to 42). The best way to avoid these problems is to use 16-bit integer or 32-bit float (see Flat field correction and background subtraction). If 8-bit gray values must be used (e.g., owing to limitations of the programs to be used subsequently for image analysis), one can try to avoid over-/under-saturated pixels, by a proper selection of the desired median and interquartile range, for example. A good starting point for this could be the median and interquartile range of an appropriate image of the series.

Importantly, if the conditions mentioned at the beginning of the section are not satisfied (relative intensity distribution in the images must be nearly identical, in the first place), intensity normalization, irrespective of method, will result in an unpredictable change of the gray values and will cause strong corruption of results if the further processing steps use absolute gray values or differences of gray values.

There are several other normalization techniques based on histogram analysis, typically nonlinear. Moreover, many common automated thresholding techniques (e.g., Otsu thresholding or maximum-entropy thresholding) are actually based on histogram analysis.

### Filtering

Filters are a big family of transformations, ranging from simple to very complex ones, that change intensity values of pixels (pixel values) based on those for a group of pixels, e.g., neighboring ones (Figure 3a). The most important class of filters (for routine use) are linear filters. In discrete (e.g., pixel-based) image processing, a linear filter replaces each pixel value by weighted average over the pixel values in its neighborhood (Figure 2a). The weights for each pixel depend only on the position relative to the considered pixel; the rule which determines these weights is called the filter kernel. The mirrored version of this kernel (i.e., the rule for setting values



**Figure 3.** Linear filtering and Gaussian filter. (a) Kernel of a 2D linear filter. Intensity in the pixel marked in red in the output image is determined from the intensities of this and the neighboring pixels marked in red in the input image. The matrix on the left (illustrating how kernels may be seen when using ImageJ) shows the weights for the pixels (in our case, they are symmetrical). The output intensity of the target pixel is the sum of the weighted intensities of the pixels covered by the kernel. Usually the weights are normalized so that total intensity in the image does not change after filtering (this is achieved by dividing the weights by their sum). (b) Applying a Gaussian filter. The radii for appropriate Gaussian filtering can be computed from the PSF. (b1) Original confocal image of a chromosome territory after FISH with a chromosome paint. An optical section ( $xy$ , bottom left),  $xz$  and  $yz$  sections (top and right). The voxel size  $60 \times 60 \times 325$  nm, the PSF size  $\sim 210 \times 210 \times 550$  nm. (b2) Smoothing with the appropriate radius: the standard deviation of the Gaussian distribution used for kernel is  $70 \times 70 \times 175$  nm (1/3 of the respective PSF sizes). Noise is reduced. (b3) smoothing with a radius that is too high; the standard deviations of the Gaussian distribution used for the kernel were  $140 \times 140 \times 350$  nm (more than half the respective PSF sizes). Fine details disappear.

in the neighborhood of a pixel from intensity in this pixel) is often called the ‘point spread function’ (PSF). When a linear filter is applied to an image that contains only a single bright point on a black background, the filtered image shows the PSF. Another name for the application of a linear filter is convolution. The physical effects during the optical transformation of the emitted light in a microscope can be well approximated by a linear filter or (synonymously) by a convolution. In many image

processing programs, filters based on the FFT (fast Fourier transform) are implemented: low-pass, band-pass, high-pass filters, etc. Without going into detail, we mention here that all these filters are just linear filters with a certain filter-kernel. Computation via the FFT only speeds up the processing, but the result is identical to the direct implementation of a linear filter.

As an example of a linear filter, we will consider the Gaussian filter which is very common in image

preprocessing and suggested by practically all software packages. Its kernel is just a Gaussian distribution. It improves images by reducing random noise that is usually caused by image acquisition, e.g., by noise in the photomultipliers of the microscope (Figure 3b). Gaussian filtering assigns each pixel a value ‘averaged’ across the neighboring pixels within a certain radius. The resulting improvement is based on the assumption of the normal distribution of light intensity from a point light source, and a random uniform (completely random) distribution of the additive noise (white noise). Gaussian filtering may be done in 2D (averaging in a single plane only) or in 3D. If available, 3D filtering is preferable. Radius (set by the user) usually has to be provided as the standard deviation  $\sigma$  for this Gaussian distribution or as the full width at half maximum (FWHM), which are related in a simple way:  $\text{FWHM} \approx 2.35\sigma$ . A radius that is too big causes averaging over too large a region and loss of information: sharp borders and small structures will be lost (Figure 3, b3). A radius which is too small may not result in the desired noise reduction. An appropriate radius value can actually be determined from the PSF of the microscope: as a rule of thumb, radius should be 2–3 times smaller, than PSF. Owing to the different characteristics of the PSF in  $x$ – $y$ - and  $z$ -directions, the radius of the Gaussian filter should also be different in  $x$ – $y$ - and  $z$ -directions. Note that the radius of the PSF and of the Gaussian have to be specified in real-world coordinates (e.g., in micrometers). If the image processing program needs the Gaussian radii in pixels, the values have to be converted according to the voxel sizes in  $x$ – $y$ - and  $z$ -directions. The resolutions in Table 1 were estimated as the FWHM of the characteristic PSF for a high-NA objectives.

- Gaussian filtering typically makes objects more homogeneous with regard to intensity; their borders will be smoother. It therefore appears ‘easier’ to set thresholds (though thresholds themselves do not become less arbitrary).
  - Gaussian filtering may be recommended for images that suffer from random noise. For instance, noisy background (outside image proper) is a good reason to apply Gaussian filtering.
  - Gaussian filtering does not improve the estimates of the positions of the centers of objects, the respective distances between them, etc.
- It also does not improve results of calculations based on massive averaging of values for individual pixels.
  - Gaussian filtering may strongly affect the borders of objects: their smoothness, their surfaces. If applications of this kind are used, one should be especially careful with the correct choice of radius, and the differences in the results that may be caused by differences in chosen radius values.

Beside the large class of linear filters, there exists an even larger class of nonlinear filters which replace each pixel value with a *nonlinear* combination of the pixel values in its neighborhood. A very common nonlinear filter, the nonlinear counterpart of the mean filter, is the median filter that may be applied for de-noising images but does not smooth away the edges. Other typical applications of nonlinear filters are enhancement or weakening of certain structures in the image, such as extraction of certain features (e.g., edges). To improve spatial adaptivity, so-called robust filters were introduced (see, e.g., Geman & Reynolds 1992), as well as adaptive Gaussian filters (see e.g., Brezger *et al.* 2007). Filters can be used to detect spots in noisy images (see, e.g. Olivio-Marin 2002, Genovesio *et al.* 2006). Other topics in recent research on filter design include holomorphic filters for detection of complex structures (Reisert *et al.* 2007; see also Supplementary Material S1), or interactively trainable ‘filters’ for the recognition of certain 3D textures (Fehr *et al.* 2005; Ronneberger *et al.* 2005).

Filters very easily trespass over the delicate border between preprocessing and processing of images. We would make two recommendations here. First, if complex filters are needed, it is preferable to consult an expert in the respective field. Second, we would always advise relegating any complex transformations of the input images to the image analysis procedure, in which case they will be applied equally to all data and will necessarily be tested together with other procedures, and there will be less chance to overlook their effect on final results and conclusions.

### Deconvolution

An optical image of a point-like object does not look like a point but shows a 3D distribution of intensities

described by the point spread function (PSF). This is an inherent property of optical microscopes that follows from the physical nature of light. As a result, microscopic images are always blurred: the image is contaminated by contributions from sources that are situated away from the point of interest, for instance in different optical planes. This image blurring can be described as a linear filtering with the PSF (see Filtering). Deconvolution applies inverse filtering methods to correct images for this linear blur. Although deconvolution is mostly associated with high-magnification images, it is also important for high-resolution imaging of large specimens using low-magnification systems (Verveer *et al.* 2007). Deconvolution is a broad topic and we will focus our discussion on the aspects that are important for quantitative measurements. Discussion of other issues may be found in several monographs and articles (Wallace *et al.* 2001, Conchello & Lichtman 2005, Pawley 2006, Swedlow 2007), and a very useful review by Wallace, Schaefer, Swedlow, Fellers and Davidson is available online at the Microscopy Primer web site (<http://micro.magnet.fsu.edu/primer/digitalimaging/deconvolution/deconvolutionhome.html>).

Two basic types of deconvolution methods can be distinguished, known as deblurring and image restoration algorithms. Deblurring algorithms handle single optical planes individually, rather than the 3D image (stack) as a whole. These algorithms attempt to remove out-of-focus light by subtracting the contribution of two neighboring image planes. The main advantage is high speed that is due to the relative simplicity of the calculations. However, the result is only a rough approximation and these algorithms should therefore not be used if the result is to be interpreted quantitatively. Restoration algorithms take into account the full image formation process of the complete 3D stack. Since the blurring with the PSF leads to a loss of information, the inverse process is non-trivial, and as a result, these algorithms are generally nonlinear and iterative in nature. These algorithms repetitively calculate an improved estimation of the true object that reproduces the observed image when blurred with the known PSF. Modern algorithms use additional information about the object such as non-negativity of the intensities and smoothness assumptions to obtain a result that is close to the true object. An increasingly popular group are the so called blind deconvolution algorithms (Boutet de Monvel *et al.* 2003, Holmes

*et al.* 2006), which do not require exact knowledge of the PSF but rather attempt to estimate both the object and the PSF simultaneously from the data.

Nowadays, since computers have become sufficiently fast, nearly every package for image acquisition and analysis offers a deconvolution option. However, deconvolution changes the raw data strongly and the results may be very different depending on how deconvolution was performed. Therefore, a well-balanced practical approach is crucial to assuring reproducibility of the results. In this aspect there is a big difference between widefield and confocal 3D images. Widefield images are so strongly affected by blur that shortcomings of the deconvolution are less important than for confocal images (see Swedlow 2007 for discussion of deconvolution of widefield stacks). The main application where stacks of widefield images are currently used is for live cell observations, which are not considered in this article. The only (or at least, the main) application where widefield stacks of fixed material are currently unavoidable is multicolor 3D FISH because widefield instruments are still necessary for more than 5–6 colors (Bolzer *et al.* 2005, Walter *et al.* 2006). Confocal microscopy strongly reduces blur at the hardware level, but noise levels in confocal images tend to be much higher compared with widefield systems. Therefore, although confocal images are obviously not free of blur, deconvolution of such images needs a more detailed discussion of the factors that affect the result.

The first of these factors is the PSF, which should be determined as precisely as possible. In practice one can use either a theoretically calculated PSF or an experimentally measured PSF. The theoretical PSF is calculated from the parameters of the optical system (microscope type, refractive index of the medium, numerical aperture (NA) of the objective, etc.). The experimental PSF is generally extracted from images of fluorescent beads. This is preferable, especially for high-resolution studies, because the optical parameters of individual microscopes and high NA objectives can vary strongly (Swedlow 2007). In addition, PSFs may be different between samples and even within a single sample if the cells contain bodies with refractive indices that are significantly different from those of their surroundings (yolk granules, chromocenters, etc.), as well as within thick samples (e.g., Holmes *et al.* 2006; von Tiedemann *et al.* 2006). Theoretically the use of

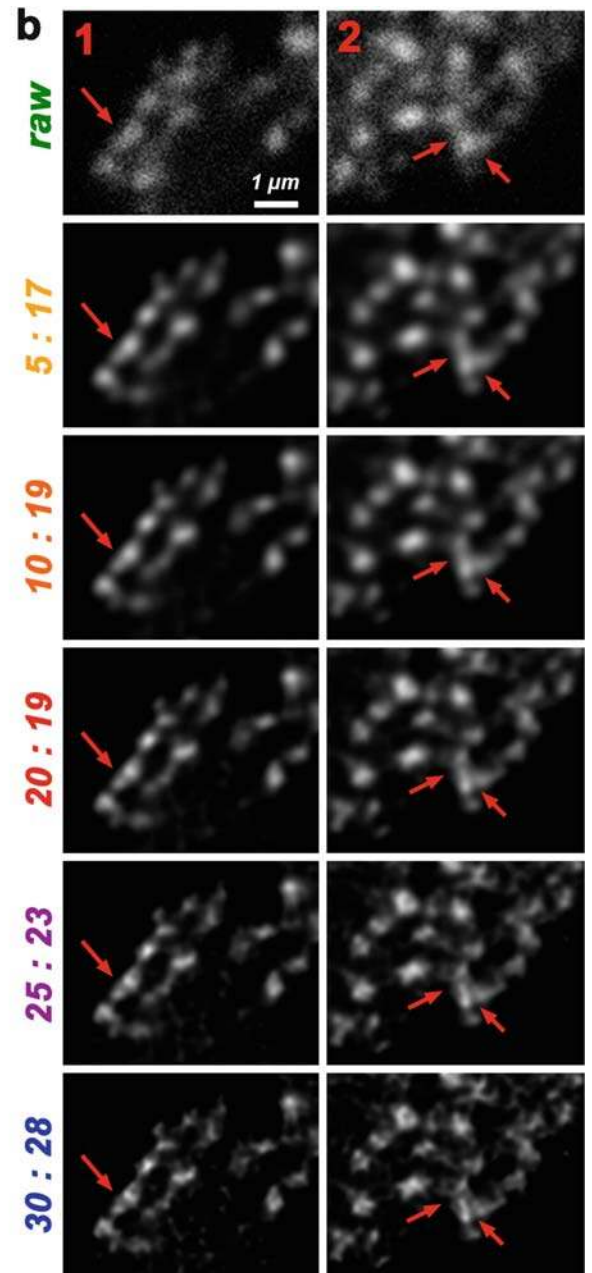
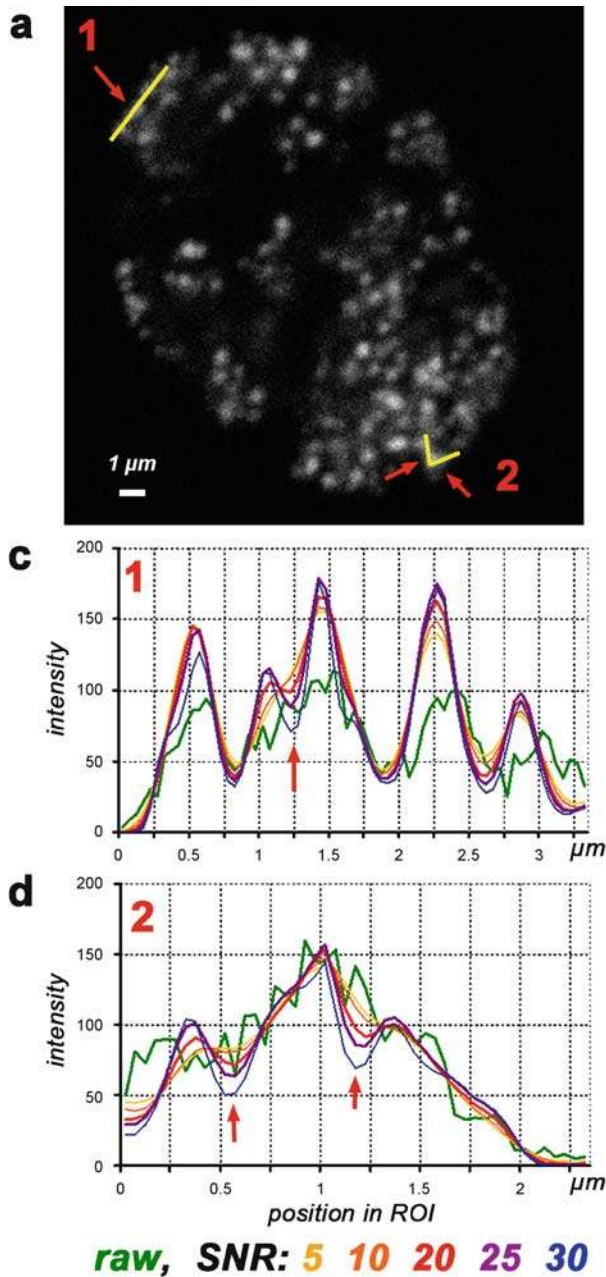
blind deconvolution algorithms could alleviate these problems. It should be noted, however, that the quality of a blind deconvolution depends strongly on the input image. Sparse images, with point-like objects, are more suitable for this approach than dense, complex images. For visualization tasks, blind deconvolution may be an appropriate choice, but quantification of heterogeneous images using such algorithms should be done with great care.

The second cardinal factor is the actual deconvolution algorithm that is employed. Various algorithms are available that are based on different assumptions about the properties of the data and of the object. Statistical iterative algorithms such as maximum likelihood estimation (MLE), maximum entropy (ME), or expectation maximization (EM), are somewhat more effective than simpler iterative algorithms that do not take into account the statistical properties of the data (see the Microscopy Primer web site). Such algorithms are implemented by a number of commercial vendors and some free programs are also available (e.g., *xcosm* at <http://www.essrl.wustl.edu/~preza/xcosm/>; plugins for ImageJ realizing more simple algorithms may also be of interest for simpler tasks: see <http://rsb.info.nih.gov/ij/plugins/index.html>). It should also be noted that all deconvolution programs use various preprocessing routines (for instance, background subtraction) and may modify the PSF (e.g., induce its symmetry). Although not a part of the deconvolution algorithm proper, this may affect the result (Wallace *et al.* 2001, Swedlow 2007). Therefore, even different implementations of the same algorithm may yield different results.

The third factor is the selection of the parameters that control the deconvolution algorithm. These parameters must be selected carefully to obtain an optimal result. The software that implements the algorithm should give guidance for the proper settings, but since manual adjustment of the parameters is usually necessary, the effect of changing any parameter should be understood. Which parameters are important depends strongly on the type of algorithm that is used. However, two aspects are important for most modern deconvolution algorithms: the number of iterations and parameters that affect the smoothness of the result. Owing to the loss of information resulting from blurring, on one hand, it is difficult to recover small features; on the other hand, artifacts such as over-sharpened edges and fake

structures (features absent in the sample) are easily introduced by deconvolution. Moreover, existing noise in the data can be amplified. Constraining the result to sufficiently smooth solutions ('forbidding' physically unrealistic features mentioned above) can prevent this. Some algorithms do not do this explicitly, and in this case artifacts will arise if the number of iterations is too high. In these types of algorithms the choice of the number of iterations is critical for limiting noise amplification and obtaining a good result. A good example of an algorithm that critically depends on the number of iterations is the MLE algorithm that is found in many popular software programs. The optimal number of iterations depends on the signal-to-noise ratio (SNR) of the data and must be set by the user (Figure 4). This is done either directly or by monitoring some parameter that quantifies the difference between the results of successive iterations. Although such parameters are usually good rules of thumb, the number of iterations remains an arbitrary user-defined setting. As mentioned above, many modern algorithms impose smoothness constraints on the result (regularization), where features that are not sufficiently smooth with regard to intensity gradients and contours are not accepted. In this case the number of iterations is less critical: after a given number of iterations the result will no longer change much, and the algorithm can safely be terminated. Although the user can often control the number of iterations, in this case it is not a critical parameter provided that the algorithm is not terminated too early. Many software packages require the SNR as an input and use it to control the number of iterations and other parameters in an empirical fashion, alleviating the burden to the user of selecting proper parameters (Figure 4).

The SNR of the data is therefore an important factor that determines the quality of the deconvolution result. Data should be acquired at an SNR that is as high as possible; reasonable oversampling also improves deconvolution results (see Sampling; see also Cannel *et al.* 2006). In this respect it is useful to experiment with the instrument settings to optimize SNR without compromising other important features. Some estimate of the achieved SNR can be obtained from the image data. However, its assessment from the image itself is not easy and demands certain assumptions about what is signal and what is noise in the particular image. It will often still be necessary to manually tweak the SNR setting (or the relevant



algorithm parameters directly) to optimize the result (Figure 4). When the SNR of the data is estimated too high, parameters can easily be set such that the contrast in the result is too high and low-intensity objects are suppressed. Errors caused by spherical aberration also may become prominent in the decon-

volution result. To avoid such artifacts it is sometimes better to assume a lower SNR and decrease the number of iterations, or to impose more smoothing in the deconvolution. However, assessing the SNR too low leads to parameter settings that can cause insufficient removal of blur, leading, for instance, to a

Figure 4. An example of deconvolution. (a) Raw image of a fibroblast with replication foci labeled in early to mid S-phase; a section from a confocal stack. Yellow lines (arrows) mark two line ROIs (regions of interest (ROIs) 1 and 2). (b) Deconvolution of this stack using the maximum likelihood estimation (MLE) algorithm. The figure shows the results of deconvolution in the regions including line ROIs (arrows) marked as 1 and 2 in (a). Deconvolution was performed with iteration numbers (17 to 28) automatically determined by the user setting of the 'quality threshold' to 0.1 as recommended by the manufacturer of the software we used (Huygens, SVI) and different SNR settings. (c, d) Intensity profiles after deconvolution with different SNR settings for linear ROIs 1 and 2, respectively. This example illustrates some of the problems that one encounters when optimizing the deconvolution parameters. The implementation of the MLE algorithm used by us depends critically on both iteration number and SNR setting. Very similar deconvolved images may be obtained with different combinations of SNR and iteration number (data not shown). At higher SNR settings (and with higher iteration numbers) the resulting contrast between the objects is higher, and the peaks in the intensity profiles are sharper, while smoothness of the deconvolved image decreases. Transmission electron microscopic data indicate that replication foci are small (diameter ~100–120 nm, up to 200 nm) and situated at least 200 nm from each other (Koberna *et al.* 2005). Because of the small size of replication foci, their angular appearance after deconvolution with high SNR settings (25–30) can probably be interpreted as an artifact, suggesting that the optimal setting for this image (with respect to the number of iterations used) is 20 or slightly more. Note, however, that higher SNR settings often allow resolution of two parts in elongated foci (red arrows in b–d). Presenting them as two intensity peaks (two foci) at a distance of 400–500 nm from one other is feasible from the physical point of view and may reflect the real situation. Optimizing deconvolution parameters for objects with variable structure and a genuine lack of a sharp border (e.g. chromosome territories, cf. Figures 2 and 12) would be much more difficult than in this example.

poor resolution of neighboring bright structures and an incorrect estimation of their numbers.

Deconvolution can cause characteristic artifacts depending on such problems as spherical aberration, errors in the measured PSF, or inadequate setting of parameters (McNally *et al.* 1999, Markham & Conchello 2001, Wallace *et al.* 2001; also see the Microscopy Primer web site). We emphasize that the quality of deconvolved images is not a simple question of a 'good' or a 'bad' result but involves many trade-offs that depend on the particular application that one has in mind. For instance, small black circles in the gray background are obviously artifacts that appear with an increasing number of iterations. However, if the deconvolution is terminated before they appear, blur may not have been sufficiently removed and an incomplete separation of bright features is often observed. Thus, the parameters of the deconvolution algorithms should be adjusted according to the type of objects one is interested in.

Application of deconvolution to a particular problem with the goal of quantitative interpretation of the result requires a great deal of experimentation. In any case, if deconvolution of images is planned, enough information about the distribution of intensities has to be collected; therefore, images should preferably be acquired using a high dynamic range (12- or 16-bit format: see Signal-to-noise ratio and dynamic range for advantages and disadvantages of these formats). To optimize deconvolution results, different algorithms should be tested and the parameters of the algorithm should be optimized using

representative data. Changing several parameters makes the problem of robustness multidimensional (i.e. values for iteration number, SNR, and threshold after deconvolution must be explored in combination). We would suggest the following tentative advice for the deconvolution of confocal images:

- Deconvolution is an extremely important tool for qualitative exploring biological structures, even though usage of this tool requires some manual optimization of parameter settings.
- We do not encourage routine deconvolution of all confocal images irrespective of the method of further analysis (as suggested by some authors). In particular, this applies to images that will be analyzed based on object centers or by averaging intensities in pixels over large parts of the image. The results of such computations are not much affected by blur (see below, Measured parameters robust to threshold settings), and deconvolution could easily reduce reproducibility, rather than improve the results.
- Deconvolution may be very useful when quantitative image analysis involves determining the number of objects, or determining direct physical contact (or lack of it) between objects of the same or different type. However, in this case attention must be paid to assuring robustness of the results to settings made by a user or, at least, to the estimation of the error associated with the parameter setting (see Estimating the error associated with arbitrary settings, below).

## Robustness of measurements

Robustness of parameters estimated from images is strongly affected by all steps in the process of image acquisition and analysis. The obtained gray values are never an exact representation of the true fluorophore distribution. Whatever efforts one makes to calibrate the image acquisition system, it is usually impossible to ensure exact reproducibility of gray values in different experiments. Additionally, the unavoidable natural biological variation of the samples will cause significant variation in the finally observed gray values. On the other hand, for most purposes it is necessary to separate objects of interest from the rest of image. Most often this is achieved by specific staining and by segmentation of the object of interest by intensity thresholding. Intensity threshold is the most prominent, though not the only, arbitrarily set parameter in image processing (e.g., see Deconvolution, above). A reasonable threshold value is relatively easy to set when the structures of interest are known and their intensities are relatively constant in the image: in this case one chooses such a threshold value that the correct structures are selected. For the majority of applications, however, a single ‘correct’ threshold simply does not exist. This is the case with all objects that do not have a sharp border: chromosome territories

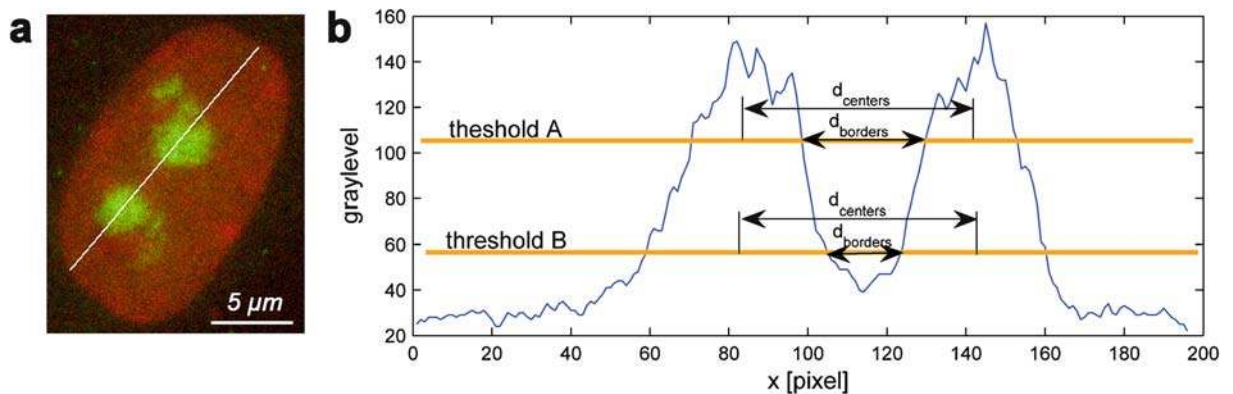
stained with chromosome paints are a clear example of this kind.

### *Estimating the error associated with arbitrary settings*

The best solution to this problem is to use measures that are not dependent on absolute gray values, or at least that do not change within a reasonably wide range of thresholds. If this is not possible, one usually still has the option to determine the final value of the measure with a certain (known) error. In particular, one can search for a threshold which is surely too low and a threshold which is surely too high, and compute, for example, volumes for a range of thresholds between these extremes. The resulting estimate is not a single value, but a range.

### *Measured parameters robust to threshold settings*

Here we mean, in the first place, using centers of objects and distances between them instead of borders and distances from borders. Provided that an object is not highly asymmetrical, the positions of the centers (geometrical or intensity centers of gravity) are reasonably robust against threshold settings (Figure 5). In particular, blur (out-of-focus light) is not an exception here: by and large, provided that PSF is symmetrical



**Figure 5.** Robustness of the positions of the centers of objects. (a) Human fibroblast with two chromosome territories visualized using FISH with the respective chromosome paint. (b) Gray value profile for chromosome paint along the white line in (a). Setting a certain threshold corresponds to cutting the peaks at certain gray level (orange lines). The distance ( $d$ ) between two chromosome territories measured between chromosome territory borders ( $d_{borders}$ ) depends strongly on the selected threshold. The distance between the chromosome territory centers ( $d_{centers}$ ) is very robust to the selected threshold. In this case the centers are taken to be the centers of the corresponding portions of peak width, i.e. the geometrical centers: intensities within the thresholded area are not taken into account. Intensity centers of gravity that weight thresholded pixels with their gray levels tend to be even more robust than geometrical centers.



and varies only slightly across the sample (which is usually true for nuclear biology studies using CLSM), the positions of object centers are only slightly affected by blur (see also Deconvolution).

#### *Combination of parameters robust to threshold setting*

Although absolute parameters (surface, volume, total fluorescence intensity) nearly always strongly depend on threshold setting, their combinations may be much more robust. This approach was used successfully, for example, for comparison of chromosome territories of active and inactive X chromosomes ( $X_a$  and  $X_i$ , respectively). While volume and surface measured for  $X_a$  and  $X_i$  clearly decrease as higher thresholds are used, the  $X_i/X_a$  ratios proved to be reasonably constant for both volume and surface (Eils *et al.* 1996). To find such parameters and prove their robustness, it is important to test them in a sufficiently wide range of threshold values. For example, one can start from a threshold value just above the level where background is apparently segmented together with the object and end with a value which obviously divides an object to several parts.

#### *Relative measurements (internal controls)*

One way to obtain robust results is to design experiments so that some reference structures (internal controls) are present in the biological sample. In this way one can reliably address two problems: first, to test whether certain parameters are different for the object of interest and the internal control (in this case the internal control should be maximally comparable to the object of interest); second, to test whether a certain parameter is different between two objects of interest when measured using the same reference. Selection of a useful internal control is not always straightforward. The main criterion is that unavoidable variations between experiments should change the desired measure and the internal control in the same way. The structures present in the same nucleus as the object of interest are usually the best option. For example, in many cases one can use active and inactive X chromosomes as reciprocal controls and simply measure differences in parameter values between them. For experiments designed in

such a manner, statistical analysis using tests for dependent samples (paired *t*-test or Wilcoxon signed-rank test, etc.) can be used to test for differences between the control and the object of interest or across several objects of interest. Exact description of shape and many other morphometric parameters is difficult, and their variation is usually high (see Shape and orientation of objects). Therefore, when shapes are considered, the use of reasonable internal controls is mostly more efficient than comparison of the observed distribution with some theoretical distribution.

#### *Normalization: advantages and pitfalls*

Normalization of any parameter is a very strong tool for reducing variation. Accordingly, any normalization should be justified in each case when it is used. (as stated, e.g., in the statistical checklist for authors of *Nature*: 'any data transformations are to be clearly described and justified'). The idea behind normalization is transparent: one transforms one of the two analyzed parameters so that the relation between them becomes linear or the effect of some third factor is excluded, which simplifies the analysis. Normalization is both useful and dangerous. An example of justified normalization was discussed above (see Intensity normalization). Unsuitable normalization can strongly alter the results:

- Normalization generally transforms data in a linear fashion. It makes sense only if the relation is indeed not too different from a linear one (nonlinear transformations are also possible, but their justification is much more difficult). Size normalization is the most common case of misuse, especially when objects of different shape are normalized by their linear size.
- After normalization, the estimates may lose their physical meaning. Clear examples are (1) transient contacts between chromatin regions and (2) distance from a gene to the peripheral heterochromatin.
- One finds normalization by maximal or minimal observed values surprisingly often in biological publications, in particular, by minimal and maximal size. For normalization one should always use some robust parameters: mean, median, or centile (e.g., 95% level; the median is the 50% centile).

### Choice of objects for image acquisition

The remaining part of the article considers several typical questions targeted by nuclear biology studies, the methods of image analysis, and statistical evaluation. Before we consider these questions, a remark should be made on a topic both very important and rarely discussed. Quite clearly, the results of statistical analysis and the final conclusions of a study can be reliable only if the material for analysis was chosen randomly: it is important to avoid unintentional bias toward, for example, expected results. Really random choice procedures are rarely possible for practical reasons, but there are rules of thumb that help to avoid bias:

- Criteria that determine which objects are suitable for the analysis should be formulated before image acquisition; in case of nuclei after FISH they are usually (i) well-preserved shape of the nucleus, (ii) presence of all targeted signals, and (iii) absence of clear artifacts
- A good option is to choose objects at a small magnification or observe nuclear counterstain and *then* check whether they satisfy the suitability criteria.
- Another option is to include in the analysis groups of objects, rather than individual objects. As an example of an appropriate rule: if a nucleus is chosen for analysis, all other nuclei observed in the same field of view (with such a magnification that a field usually contains several nuclei) and satisfying the suitability criteria should also be used.

### Radial distribution

#### *General*

The classic example for this problem is the distribution of chromosome territories in nuclei. It has been shown that chromosome territory positions within the nucleus are non-random and, depending on the cell type, correlate strongly with chromosome gene content (Croft *et al.* 1999, Bridger *et al.* 2000) and/or chromosome size (Sun *et al.* 2000, Cremer *et al.* 2001, Habermann *et al.* 2001, Bolzer *et al.* 2005, Mayer *et al.* 2005, Neusser *et al.* 2007); radial

positions of chromosomes are not identical in normal and malignant cells (Cremer *et al.* 2003) and are an important factor contributing to probabilities of chromosome translocations (Boei *et al.* 2006, Meaburn *et al.* 2007). Recently the interest also focused on the positions of individual loci (Dietzel *et al.* 2004) and groups of genes with certain properties, e.g., representing gene-rich and gene-poor regions (Kupper *et al.* 2007). Studies of the radial distribution of chromosome territories and chromosomal subregions also provided important information about the evolution of the nuclear architecture and its relation to the evolution of the karyotype (Habermann *et al.* 2001, Tanabe *et al.* 2002, Neusser *et al.* 2007).

Considering radial distribution, one wants (1) to determine whether, for example, objects tend to be situated in the central or the peripheral part of the nucleus or (2) to demonstrate a difference in the radial distributions of two types of objects. In some sense, radial distribution is a basic problem for measurements in the nucleus because it implicitly affects many parameters, such as characteristic distances between objects of different kinds (see Boei *et al.* 2006 with regard to positions of chromosome territories).

Two basic approaches exist to analyzing the radial distribution of the defined nuclear objects, depending on the reference used to measure distances: nuclear center or nuclear border (Figure 6a,b). Although the nucleus does not have a structurally marked center, technically it is easy to define nuclear center as the center of gravity of nuclear counterstain. Intensity of counterstain in individual voxels may be taken into account (counterstain center of gravity) or not (geometrical center). The two centers, however, will really differ only when nuclei are strongly and asymmetrically heterogeneous: if they contain a large nucleolus or chromocenters. All initial publications on radial distribution used a nuclear center-based approach. It is known, however, that active euchromatin predominantly occupies inner positions in the nucleus, while transcriptionally inactive heterochromatin tends to locate near the nuclear border (and around nucleoli). Moreover, it has been shown that some genes move away from the nuclear envelope in case of transcriptional activation and back to it in case of inactivation (Kosak & Groudine 2004, Chuang *et al.* 2006, Ragozy *et al.* 2006, Chuang & Belmont 2007). Therefore, a number of studies have

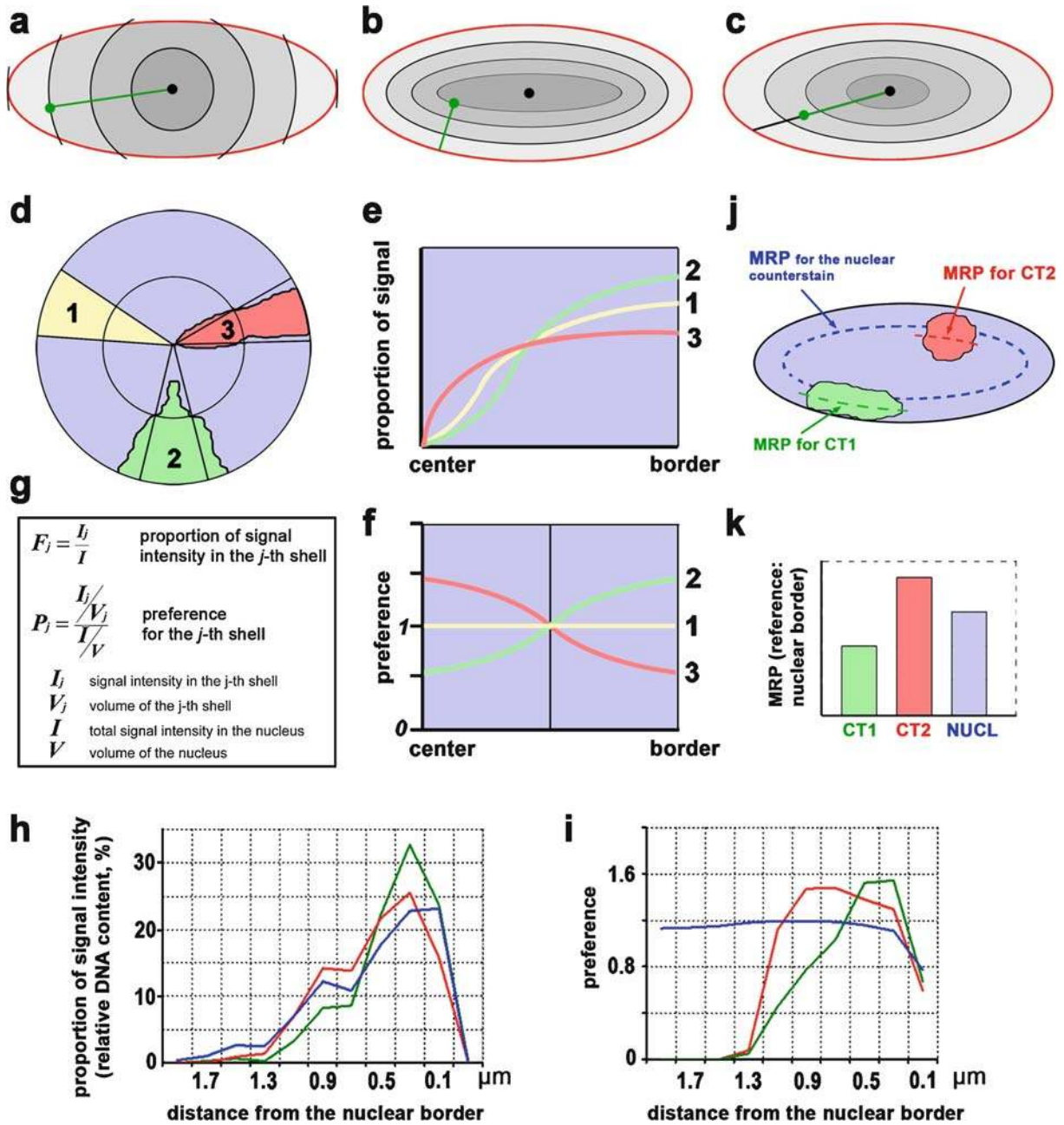
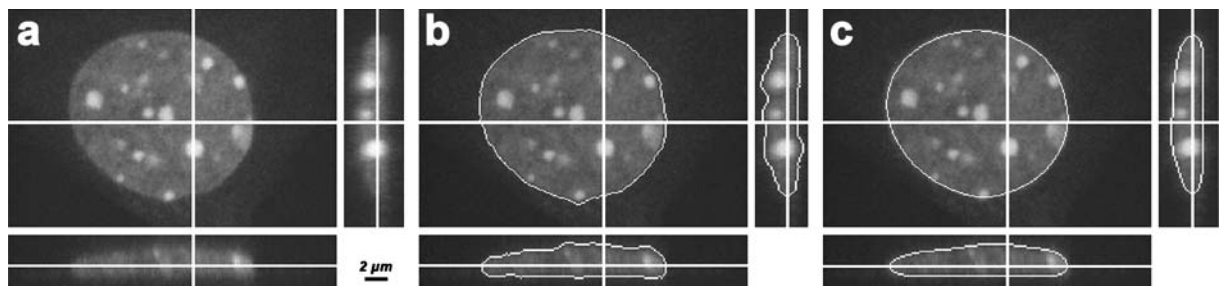


Figure 6. Radial distribution: general concepts. (a–c) References and distance measurements used for the analysis of radial distribution. Black lines show subdivision of the nuclear space into shells, often used for calculations. (a) Reference: nuclear center, absolute distance. (b) Reference: nuclear border, absolute distance. (c) Reference: nuclear center, relative distance measured as the ratio of the absolute distance to the nuclear center to the length of the respective nuclear radius. In this case shells have the same relative radius. (d–g) Graphs showing the radial distribution of signal exemplified by three ‘chromosome territories’ (1, 2, 3) with a different distribution of signal (d). For all three chromosome territories, more signal is present near to the nuclear periphery; i.e. the proportion of signal situated at certain radial distance decreases from the border to the center (e). If the volume available at the nuclear periphery and in the center is taken into account, a different preference of the three chromosome territories may be demonstrated: no preference, preferential peripheral positioning and preferential central positioning, respectively (f). If the nucleus is divided into shells, radial distributions may be calculated as shown (g). (h, i) An example (based on real material) showing different preferred radial positioning for two chromosomes (red and green); nuclear DNA counterstain (blue) shows no preference. (j, k) The median radial position (MRP) is determined by the radial position of the surface dividing the respective objects to two equal parts. CT, chromosome territory.

determined radial distributions using the nuclear border for reference (Kupper *et al.* 2007, Neusser *et al.* 2007). Even though the nuclear border (envelope) is a real structure, it is technically more difficult for practical use than the nuclear center, since one needs to threshold either the nuclear counterstain or the nuclear envelope, e.g., by staining the lamina. Therefore, the reference strongly depends on the respective threshold setting, whereas the position of the center of a segmented nucleus is robust against variations in threshold. Furthermore, segmenting the nucleus may be difficult because of bright chromocenters and, sometimes, dark nucleoli, which cause convexities and concavities ('bumps and dints') in the nuclear surface when using a straightforward thresholding procedure. A simple approach to obtaining a more appropriate estimate for the nucleus border was recently suggested (O. Ronneberger and M. Keuper, unpublished results) that can be applied if the curvature of the border is significantly lower than the curvature of its uneven regions (Figure 7). The necessity to study the spatial organization of genomes in nuclei of native tissues has recently raised a problem of segmentation of nuclei from tissue section images, where nuclei have irregular shape and often 'touch' one another. Such segmentation can be done, for example, using commercial AMIRA software (Mercury Computer Systems, Boston, USA), though doing it accurately for nuclei of irregular shape is quite time consuming. A dedicated program has also been developed recently (Gudla *et al.* 2008).

If the nuclear center is used as reference, for each voxel in the signal one determines the distance from this voxel to the nuclear center (Figure 6a). When the nuclear border is chosen for reference, for each voxel in the signal the distance to the nearest point in the nuclear border is determined (Figure 6b) in a currently available implementation (see Albiez *et al.* 2006). This approach (called distance transform) is a simple, but not generally optimal, solution: compare a voxel within a concavity of the border and a voxel situated at the same minimal distance from its convex region. Irrespective of the reference, the distance may be weighted or not weighted by intensity in this voxel. Three related pieces of information may be extracted from these data (Figure 6d–i). (1) Radial signal intensity distribution: the distribution of signal intensity in the nucleus (this is the distribution of the number of voxels with signals above threshold, if there was no weighting). This distribution shows which proportion of the total signal is situated at given radial position (Figure 6e,h). (2) Preferred localization: considering all voxels that constitute the total nuclear volume, central positions are less abundant than the peripheral ones. If one corrects for this different abundance (the proportions of voxels situated at a certain distance from the center; in other words, the difference in volume between central and peripheral shells of the same width), one learns which positions are preferentially taken by certain types of signals (Figure 6f,i). A disadvantage of this approach is a low precision for the central part of the nucleus (which has a small



**Figure 7.** Computation of a smooth nuclear border. (a) DNA counterstain of a mouse fibroblast showing bright chromocenters. (b) An attempt to define the nuclear border by straightforward thresholding yields inadequate results; in particular, basal and apical surfaces are compromised by blur from chromocenters. (c) A much more appropriate border may be defined using a simple smoothing procedure. First a thresholded binary image of the nucleus is prepared (background 0, foreground 255). This image will simply show a bumpy white nucleus on black background depicted by the white borderline in (b). This image is then smoothed with a 3D Gaussian filter with radius of 1/3–1/2 of the size of the bumps in both *xy* and *z* directions and then thresholded at the gray level of 128. This procedure is repeated several times. In the resulting binary image the 'bumps and dints' in the border are eroded, whereas the rest of the nuclear border remains practically unchanged.

volume). (3) For each stained structure in a nucleus, one can also determine the surface that has a constant distance from the reference (center or border) and divides the nucleus in such a way that half of the structure lies inside this surface and the other half is outside it (Figure 6j,k). This surface characterizes the median position of the signal, and its distance from the reference may be called median radial position (MRP), in our case, in the nucleus. The radial distribution and preference curves discussed above have some disadvantages. They are dependent on the threshold settings for the nuclear border. Furthermore, although they are sufficient to show the trend of the signal distribution, defining their errors is difficult because signal densities in neighboring shells are not independent. MRP is least dependent on the threshold settings for signals and allows sufficiently robust analysis of the data using routine statistical tests. Therefore, it is currently the most practically important piece of information.

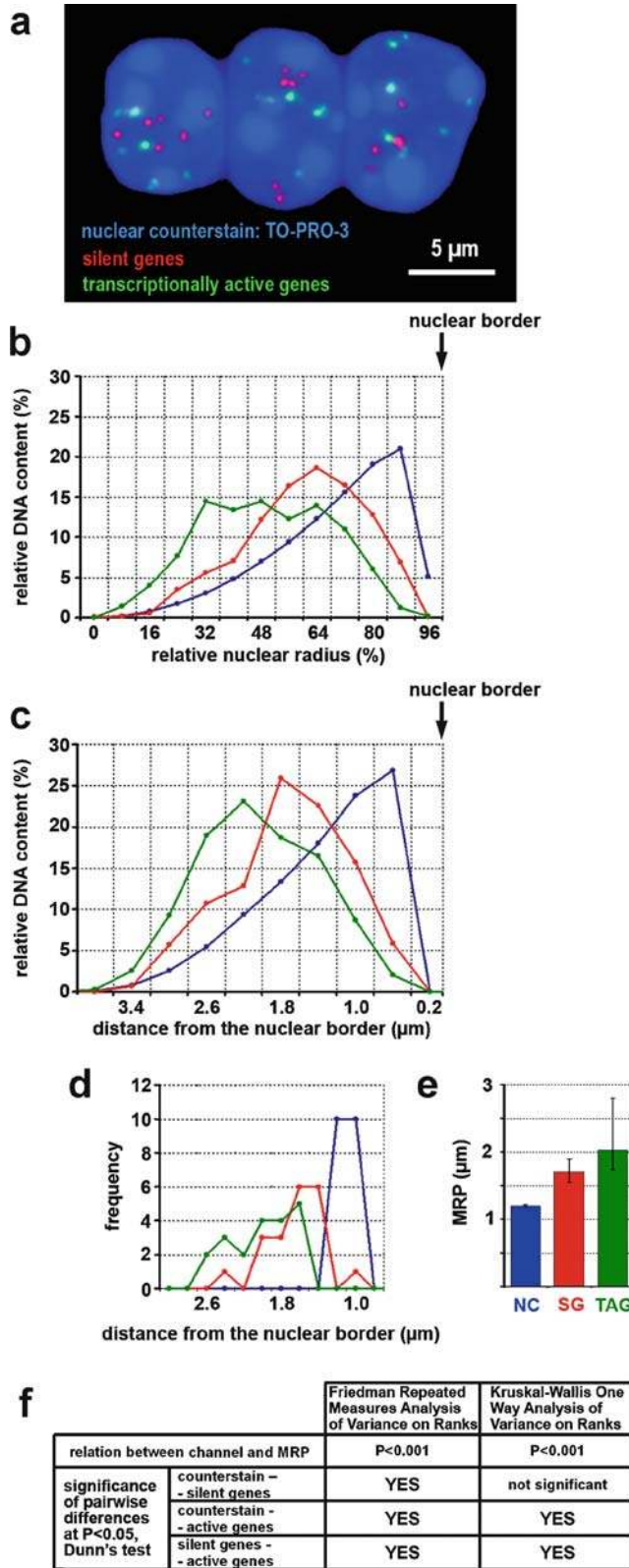
The statistical analysis follows the same scheme with any approach (it will be illustrated using an example shown in Figure 8). One has a sample of nuclei with MRPs for each signal in each nucleus. Note that MRPs for different signals from the same nucleus are not independent (e.g., they are affected in a similar way by the shape of the nucleus). MRP values in sufficiently big samples usually show a statistically significant deviation from normality (Figure 8d); we therefore suggest always to use non-parametric tests. Correspondingly, we recommend Wilcoxon signed rank test for comparison of a single signal with nuclear counterstain and Friedman repeated measures ANOVA on ranks followed by Dunn's test for pairwise differences in case of two and more signals. Assuming that they are implemented using some statistical software (e.g., SigmaStat or many others), it will first be tested whether the distribution of signals depends on signal type in a statistically significant way for the whole data set. Then, if it does, the significance of pairwise differences will be checked (Figure 8f). We recommend in all cases to include nuclear counterstain in the analysis. If there is no significant difference between at least one of the signals of interest and the counterstain, it may mean that sample size (the number of the nuclei studied) was simply too small.

An important issue in the analysis of the radial distribution is the decision about description of data

in 3D or 2D. Here we mean not the type of image acquisition (confocal stacks vs single widefield images), but the model for data description. A characteristic example here is the two variants of the radial distribution of chromosomes within the interphase nuclei expressed most pronouncedly in spherical and very flat nuclei, respectively (for review see Foster & Bridger 2005). In voluminous nuclei (such as spherical nuclei of human lymphocytes), radial positions of chromosomes correlate with their gene content. This is clearly a 3D distribution. By contrast, in flat nuclei (such as those of human fibroblasts) chromosome territories often expand from the basal to the apical nuclear surface. Therefore, chromosome territory distribution may reasonably be analyzed in 2D, considering individual nuclear sections or even projections of entire stacks (Bolzer *et al.* 2005). In this case, radial positions of territories were found to correlate with chromosome size but not with gene content: territories of small chromosomes were all located close to the nuclear center, while territories of large chromosomes were located at the nuclear edge (Figure 9). Nevertheless, accurate 3D analysis revealed a non-random distribution of gene-rich chromosome territory 19 and gene-poor chromosome territory 18 along the basi-apical axis of the nucleus (Neusser *et al.* 2007), which corresponds to the 3D gene content-dependent distribution of chromosome territories in voluminous nuclei.

#### *The relative radius approach*

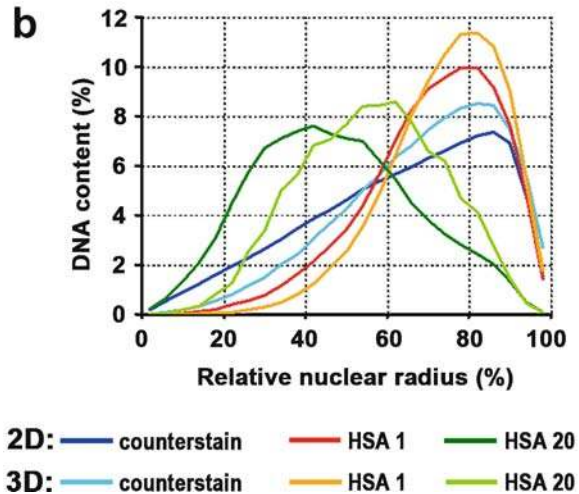
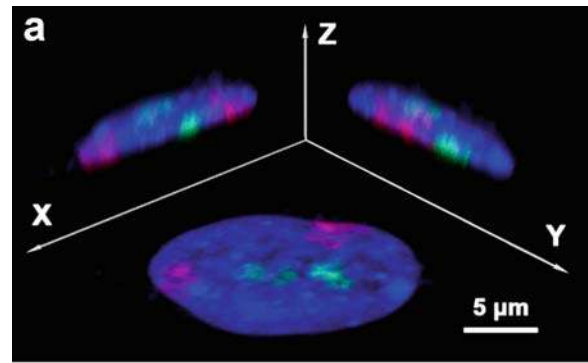
Several early studies on relative chromosome territory arrangements used lymphocyte nuclei that have a very regular spherical shape (Park & De Boni 1998, Croft *et al.* 1999, Boyle *et al.* 2001). To compensate for size variation, the authors determined the relative radius of signals (Figure 6c): the position of each voxel is presented as the proportion of the radius going through this voxel (note that radial position thus depends on segmentation of the nuclear border). The relative radius approach was later also applied to non-spherical nuclei. This approach transforms all shapes to spheres (3D) or circles (2D), which strongly distorts the distribution of physical distances in flat or spindle-shaped nuclei. The relative radius approach has been applied in studies of lymphocytes and similar cells (Kosak *et al.* 2002, Ono *et al.* 2007), often in combination with methanol-acetic acid fix-



ation following hypo-osmotic treatment. This treatment yields highly flattened nuclei with considerably increased  $x$ - $y$  diameters. While such nuclei are ideally suitable for 2D analysis, it is important to note that radial positions of signals do not change proportionally (Kozubek *et al.* 2000). Nevertheless, in some cases the loss of 3D information may be less important than quick image acquisition of large numbers of nuclei. The following consideration helps to establish preliminary correspondence between the radial positions calculated from projections and respective 3D positions in a sphere. For a signal uniformly distributed over the surface of a sphere with radius  $R$ , the MRP of the distribution resulting from projection to a plane ( $r_m$ ) is

$$r_m = \frac{R\sqrt{3}}{2} \approx 0.866R$$

The relative radial positions observed in biological studies using the 2D approach are typically in the range 0.5–0.7, which is understandable. On one hand, as mentioned above, even signals wholly restricted to the surface of a spherical nucleus cannot have radial position above 0.866. On the other hand, the whole internal sphere with the radius of  $R/2$  contains only 12.5% of its volume, while half of the



**Figure 9.** Radial distribution of two chromosomes (HSA1 and HSA20) in the nuclei of human fibroblasts: comparison of 2D and 3D models. This figure illustrates the importance of choosing a correct model for the analysis of the radial distribution. (a) Three projections of a representative fibroblast nucleus with HSA1 (red) and HSA20 (green) after FISH with respective chromosome paints. Note that more central position of HSA20 is observed only in XY projection: in flat nuclei most chromosomes contact the cell border at the apical or basal surface of the nucleus. (b) The distribution of the two chromosome territories was studied in a sample of 25 nuclei using the relative radius approach, either using a 2D model (the nuclear border was determined from the projection of the nucleus) or a 3D model (using the real nuclear border). In the latter case the more central position of HSA20 and the difference in the distribution of the two chromosome territories were notably underestimated.

volume is contained in the outer shell with width of only  $0.2R$ .

#### Reference determined by the studied signals

If the internal structure of the nucleus is not symmetrical, it may be reasonable, for the analysis

**Figure 8.** Radial distribution of genes transcriptionally active (housekeeping) and silent in mouse ganglion cells: from images to statistical analysis. (a) Three ganglion cell nuclei (projections of confocal stacks) showing the result of FISH with cocktail probes for three housekeeping genes and three genes silent in ganglion cells. The same color code (red, silent; green, active; blue, nuclear counterstain) was used for all parts of the figure. A small sample (20 nuclei) was analyzed. (b–d) Relative radius approach (b) and the analysis of absolute distances from the nuclear border (c) gave qualitatively similar results, indicating a more internal localization of the active genes. (d, e) The distributions of the MRPs (d) and median MRP values (e) are also different. The shape of ganglion cell nuclei is rather uniform, therefore MPR values for counterstain show only small variation. MPR values for genes are more variable and their distributions are asymmetrical. Error bars show 95% confidence interval for the median MRP values. (f) Statistical analysis shows a statistically significant relation between the type of signal and MRP; pairwise comparison confirms statistically significant differences between radial positions (MRPs) of signals. This example illustrates that the Friedman test (which takes into account that signals from the same nucleus are dependent) should have more power for this analysis, than the Kruskal–Wallis test (which does not use this information).

of DNA loci, to use the center of gravity of all studied signals as a reference point for their radial distribution. For example, this approach helped to demonstrate that NORs (nucleolar organizers) of all NOR-bearing chromosomes tend to cluster (Teller *et al.* 2007). NORs actually cluster around the nucleolus, and the discussed approach compensated for the variation in the position of the nucleolus. Strictly speaking, this approach requires that signals have similar radial distributions, but in practice differences between them, if they exist, may also be demonstrated using this approach. When the number of signals is small, absolute radial positions will tend to be smaller than the results based on the geometrical center of the nucleus. Practical observations showed, however, that when the distribution of signals is reasonably symmetrical, even six signals are usually enough for this approach to estimate realistically radial positions and differences between them. In particular, using a sample of 44 nuclei we determined radial positions of chromosome territories 4, 7, and 21 in HeLa cells (chromosome 7 was triploid, so that there were 7 chromosomes per cell). Estimations based on this approach and the relative radius approach differed by less than 5% (data not shown). Using the center of gravity as a reference, one can also demonstrate a trend to asymmetrical positioning, as was done for centromeres in nuclei of early mouse embryos (Merico *et al.* 2007).

#### *Choice of the approach to study radial distribution*

To validate the different approaches to analyzing a radial signal distribution we performed a special simulation study (see Supplementary Material S2). Importantly, our model allowed variation in the shape of modeled nuclei. Our goal was to assess the discriminating power of different approaches with regard to (1) presence of pattern: deviations from a random uniform distribution (i.e., the one where probability for signal occurrence is equal in all points of the nuclear volume); (2) type of pattern: first of all, preferred signal location at the center or periphery of the nucleus; (3) strength of pattern: i.e. the degree of attraction to certain structures used as a parameter in these models; and (4) sensitivity to the shape of nuclei. Our simulation studies confirmed that the distribution of MRPs in samples of nuclei is more suitable for discrimination of the various

patterns than the distributions of all observed distances, regardless of the specific measure employed. The simulation studies and observations on real material may be summarized into the following recommendations:

- Adequate results of the analysis depend in the first place on the adequate choice of model (e.g., 3D or 2D) that corresponds best to the organization of the studied nuclei and the targeted problem.
- Use of the nuclear center or nuclear border for reference is equally effective with regard to detection of the deviation from a random distribution, of the preferential localization toward the center or periphery, and of different degrees of such a preference. These measures cannot detect clustering as a deviation from random distribution.
- Use of the nuclear border as reference is preferable when absolute physical distances from the nuclear border or differences in such distances are biologically important.
- Use as the reference of the center of gravity of signals is efficient when major asymmetrical structural differences are observed within the nucleus.
- When observed differences in radial positioning are considered, it should be taken into account that all measures (with partial exception for the adaptive center approach) are sensitive to the shape of the nuclei: that is, similar differences will result from a non-random distribution of signal in a population of spherical nuclei and a random uniform distribution in a population of non-spherical nuclei.

Variation in shape and size of nuclei is a very important issue for studies of radial positioning. The longest absolute distances from a given reference point are represented only in the fraction of largest nuclei. Size normalization does not solve this problem because radial distributions depend very strongly on the shape of the nuclei (see below). As normalization transforms data in a linear fashion (proportionally), one can reliably compare only nuclei that have the same or similar shape. Calculations based on absolute values yield reliable information about differences between signals, which is usually sufficient for biological interpretation of results. In particular, the main advantage of using the nuclear border for reference, measuring real physical



distances from the real nuclear envelope to loci, is lost after normalization. Real spatial information is distorted especially strongly when data for each nucleus of a sample are normalized independently by the size of this nucleus. It is therefore strongly advisable to characterize samples by median (or mean, if appropriate) MRP values calculated from absolute MRP values for individual nuclei and respective median (or mean) size of nuclei. These two parameters may then be used for comparison of samples. The relative radius approach may be preferable when the shapes of compared nuclei are not too different from spherical (or are discoidal) and has proved to be useful for general, architectonic comparisons.

### Distances between signals and clustering

To determine the centers of the objects of interest and to measure the distances between them is a relatively simple task. Distances indeed provide important information about various biological processes. For instance, a change in the distance to a flanking control locus may indicate relocation of a transcriptionally activated locus (Morey *et al.* 2007), while distances between chromosomes and loci are useful for understanding their non-random spatial arrangement which plays an important role in chromosome translocations (Neves *et al.* 1999, Roix *et al.* 2003, Meaburn *et al.* 2007, Soutoglou *et al.* 2007).

When several objects, e.g., FISH signals of several genes, are present, one can measure distances between them. The median of these distances calculated for each cell may then be used as an observation in a sample of nuclei and samples of different nuclei may be compared on this basis using standard statistical tests. However, this approach is less efficient than comparing radial positions of signals. The set of distances from each signal to each other has intrinsically high variation—larger numbers of nuclei are needed. Another possibility is to use angles. First, one can measure angles between three objects (A, B, C). Such angles, however, have physical meaning only if the objects are indeed successive parts of a higher-order object (e.g., neighboring subregions of the same chromosome). Second, one can measure the central angle between two objects, the angle AOB, where O is the center of

the nucleus. However, as the nuclear center is not physically real, angles for objects situated close to the center are very imprecise. In short, though central angles were successfully used in several studies (Kozubek *et al.* 2002, Bolzer *et al.* 2005), the same information in most cases may be obtained more effectively from the analysis of distances.

Distances between objects provide information on clustering. A number of recent publications indicated transient ('kissing') or permanent spatial proximity of genomic loci with other loci or various nuclear structures. This proximity is ascribed very important roles in transcriptional regulation, the formation of transcription factories and X chromosome inactivation (for reviews see Fraser & Bickmore 2007, Misteli 2007, Sexton *et al.* 2007, Soutoglou & Misteli 2007). Especially in the case of transient contacts, the observer of fixed cell nuclei sees the objects not in direct proximity but at different distances from one other, with a certain excess of smaller distances: they are clustered. The opposite trend to clustering is called exclusion. If excluding objects are numerous, exclusion leads to an ordered uniform distribution of objects: all distances between them tend to be maximal and hence equal. If excluding objects are sparse, an excess of large distances between them is observed. An excess of large or small distances is a typical statistical problem, often solved by comparison of the observed distribution of distances with a theoretical one that would occur in the case of a random uniform distribution of objects. In practice, methods of this kind have been applied only under the assumption of a spherical shape of the nucleus (Kozubek *et al.* 2002). We will here consider only methods that do not need assumptions about a theoretical distribution. They are based on comparison of different signals stained in the same nucleus.

#### *A medium to large number of objects forming small clusters*

Consider a rather general case of soft clusters: both the number of objects per cluster and distance between objects in the cluster are allowed to vary between nuclei. A medium number means here 8–10 or more. Although the formulation of this problem may seem vague, its solution is straightforward. The analysis is based on the distance from a signal to its nearest neighbor (NN distance) which is an estab-

lished measure to describe patterns of clustering or exclusion. The distributions of the NN distances discriminate well between a random uniform distribution and different degrees of clustering or exclusion. They are also reasonably robust against the variation of the shape of the observed nuclei (the smaller are the NN distances in comparison to nuclear size, the less the shape of the nucleus affects their distribution). Nevertheless, it is preferable to compare two or more types of objects in the same nuclei, to avoid the influence of nuclear size. As an example we discuss 10 signals forming clusters of 2–4 signals.

One determines the NN distance ( $d$ ) for each signal in a nucleus ( $d_1$  to  $d_{10}$  in our case) and characterizes a nucleus by their median values ( $M_1$  to  $M_n$  in the case of  $n$  nuclei). The degree of clustering for signals of different type may then be compared simply by Mann–Whitney tests (comparison across groups of nuclei) or Wilcoxon signed-rank tests (comparison of distributions of two kinds of signals from the same nucleus) on the medians  $M$ . If the analyzed signals have different radial distribution, but form similar clusters, a difference in radial distribution will not affect the result. On the other hand, it should be taken into account that a difference in radial distribution will affect NN distances. The method discussed does not describe the number of signals per cluster or distance between clusters, but these parameters may be estimated independently. With a larger number of signals per nucleus it is possible to use spatial point pattern methodology (see, e.g., Beil *et al.* 2005, Buser *et al.* 2007). This approach allows a more detailed analysis of the observed patterns because it does not reduce the information to median NN distances in individual nuclei.

#### *Comparison between cells (independent samples)*

Another simple case for analysis and important for nuclear biology is clustering of two homologous loci when the distance between them is measured in different cell samples (e.g., different stages of cell cycle). For instance, it has been demonstrated that a temporary contact between the two X chromosome territories in nuclei of female embryos is necessary for X-chromosome inactivation (Bacher *et al.* 2006, Xu *et al.* 2006). To prove this contact, Xu and co-authors measured distances between pairs of Xic loci at the time point when these loci were supposed to

cluster, as well as before and after it. The difference between the distributions of such distances was shown to be significant by a Kolmogorov–Smirnov (KS) test. It was also emphasized that the observed maximal distances between distributions accumulated in the range of small distances, that is, the difference was due to an excess of clustered (lying close to one other) Xic loci. Additionally, lack of clustering was demonstrated for another locus situated on the X chromosome not far from the Xist region.

An opposite example is the recent study of Angelman syndrome/Prader–Willi syndrome (AS/PWS) loci by Teller *et al.* (2007). Transient clustering of these loci at the late S-phase has earlier been suggested as a possible mechanism for maintenance of opposite imprinting, while Teller and co-authors argued against kissing of AS/PWS loci. Using ANOVA on ranks (Kruskal–Wallis test followed by Dunn’s pairwise comparisons) they analyzed distances between AS/PWS loci at different stages of the cell cycle. They found no difference between the stages of the cell cycle. Lack of a functionally meaningful association was supported by the fact that variation between nuclei at all stages was several times greater than the differences between the stages.

A methodologically related problem is the clustering of loci from the same chromosome that may play an important role in regulation of these loci. If the genetic distance between loci is large, the methods discussed above are wholly appropriate. In case of genetic distances of 3–5 Mb and less, it appears more productive to apply methods used to study chromatin folding (Yokota *et al.* 1995, Simonis *et al.* 2006), which are based on the analysis of the relation between genetic and nuclear distances between loci.

#### *Comparison within the same cells (dependent samples)*

Two pairs of genes represent the simplest situation in which heterologous genes from different chromosomes come to spatial proximity. Several examples have been reported (Spilianakis *et al.* 2005, Brown *et al.* 2006, Lomvardas *et al.* 2006). In particular, active alpha- and beta-globin genes, as well as active homologous alpha-globin loci, are frequently juxtaposed in human erythroblast nuclei. However, neither beta-globin nor globins in mouse erythroblasts tend to be found in spatial proximity (Brown *et al.* 2006), which emphasizes the necessity of rigorous quan-

titative analysis for each particular experimental system. The problem is difficult, however, because distances between loci depend strongly on their radial distributions, as was clearly illustrated for chromosome territories by Boei *et al.* (2006).

Hence, we consider genes *A* and *B* represented by visually indistinguishable homologous alleles  $A_1, A_2$  and  $B_1, B_2$ . An important property of heterologous clusters is that an *A* allele can cluster with any of the two *B* alleles, and vice versa. We here propose a simple test (Figure 10) related to the test suggested by Shiels *et al.* (2001). This approach had a power (sensitivity) of practically 1 for the detection of clustering of either homologous or heterologous pairs with a maximal distance up to 40% of the nuclear radius for simulated data and a significance level of 5% for as few as 5 replications. Performance of the test was not affected by the distribution of signals (random uniform or giving preference to a peripheral position). Most importantly, the test also proved to be very robust to differences in the radial positioning of genes *A* and *B*. For simulated data, performance of the test was practically unchanged for median relative radial positions of the two genes of 0.3 and 0.7, respectively. It should be noted, however, that owing to the complicated geometrical dependences between  $d_{\text{het}}$  and  $d_{\text{hom}}$  (see Figure 10), this test can be slightly anti-conservative (it tends to find differences where there are none), especially for large numbers of replications (i.e., large sample size, which, however, is rarely a problem for biologists collecting small samples of 3D data) and one-sided hypotheses.

With more than two genes, the number of possible combinations increases and formulation of the hypothesis to be tested becomes a very difficult task. If the chosen hypothesis is not too complicated, an appropriate test can be designed. Unfortunately, the power of such tests may prove low even for simple hypotheses. We illustrate this issue for the simplest possible case: three genes and the hypothesis that there is only one type of cluster, namely heterologous clusters composed of one allele of each gene (see Supplementary Material S3).

#### Clustering of genes to nuclear bodies

The close physical proximity of distant genomic loci in the nuclear space is currently often explained by the targeting of these genes to the same nuclear structures (Fraser & Bickmore 2007): transcription

factories (Ragoczy *et al.* 2006, Osborne *et al.* 2007), speckles (Shopland *et al.* 2003, Moen *et al.* 2004, Brown *et al.* 2006), and PML bodies (Shiels *et al.* 2001, Wang *et al.* 2004). *In vivo* observations by Dundr *et al.* (2007) showed that upon activation transgene arrays of small nuclear RNA genes (U2) moved to relatively immobile Cajal bodies. This introduces one more important problem: the analysis of the degree of association of genes with these nuclear structures. One can simply try to estimate the ratio of signals that stay in direct contact with nuclear bodies, but defining such contact, e.g., by overlap or position within some threshold distance from the nuclear body, is not easy and is mostly arbitrary. Analysis of this issue as a clustering problem is therefore very promising. A serious (though universal) difficulty in application of this method is that owing to variation in intensity of staining and visual fusion of nuclear structures themselves the segmentation of, e.g., speckles and Pol II foci is difficult.

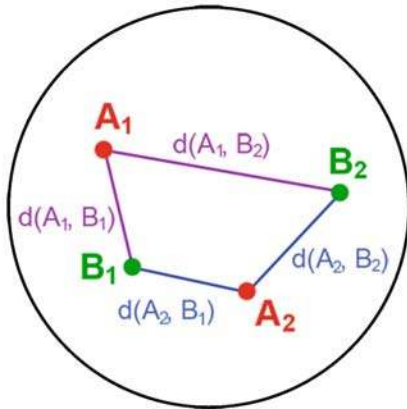
The statistical problem can be formulated as follows: we have several genes (G-objects) that may cluster to a similar or higher number of, e.g., nuclear bodies (B-objects). When the number of signals analyzed is large, there are established methods to determine clustering or exclusion between signals of two different types. They are implemented, for instance, in the GeoStoch Software (Mayer *et al.* 2004) and in the 'spatstat' package for R (Baddeley & Turner 2005). However, very few attempts have been made to apply such methods to cell biology data (Beil *et al.* 2005, 2007, Jinno *et al.* 2007).

When the numbers of nuclei studied and/or the number of signals observed are small, the problem becomes even more difficult because not only the number and the radial distributions of objects but also the characteristic sizes of the objects play a role. To our knowledge, only one method of analysis which is applicable in such a situation has been published (Shiels *et al.* 2001). For each nucleus, they record the mean of the cross nearest-neighbor distances (cNN distances) for each nucleus. cNN distances are the shortest distances from the center of each object of type G to the center of an object of type B ( $dl_i$  for nucleus  $i$ ). In a similar way the authors determine the mean of NN distances between two B-objects OR (*sic!*) a B- and a G-object ( $d0_i$  for nucleus  $i$ ). In order to test whether a particular gene tends to be close to at least one nuclear body, a paired *t*-test with pairs  $d0_i, dl_i$  is

## Test for clustering (or exclusion) of heterologous genes

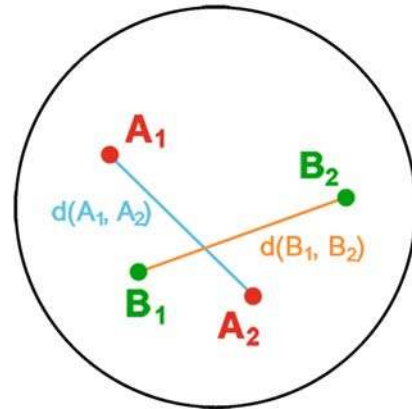
Let  $d(A,B)$  denote the distance between the alleles of genes A and B. If genes are distributed identically and are not clustered or excluded, the distances between homologous and heterologous genes should be equal.

$$d_{\text{het}} = 0.5 \min [d(A_1, B_1), d(A_1, B_2)] + 0.5 \min [d(A_2, B_1), d(A_2, B_2)]$$



The mean minimum distance between heterologous pairs  $d_{\text{het}}$  is used as a measure of the proximity of heterologous pairs of genes

$$d_{\text{hom}} = \min [d(A_1, A_2), d(B_1, B_2)]$$



The minimum distance between homologous genes  $d_{\text{hom}}$  is used as a measure of the proximity of homologous pairs of genes

**In the absence of clustering or exclusion,  $d_{\text{het}}$  should be roughly the same as  $d_{\text{hom}}$ . If a sufficient number of replicates is available, a Wilcoxon-Signed-Rank-Test can be used to decide whether  $d_{\text{het}}$  and  $d_{\text{hom}}$  are significantly different.**

Figure 10. Analysis of the positioning of heterologous genes.

used. Large negative differences ( $d0_i - d1_i$ ) indicate that genomic loci tend to be closer to nuclear bodies than nuclear bodies are to one another, that is, they indicate clustering of genes to nuclear bodies. This test is appealing because of its conceptual simplicity and easy implementation using standard statistical software. It is also conservative in the

sense that it is biased against the support of the hypothesis of clustering and tends to overestimate exclusion (the proof is omitted). For example, in the case of two genes and six nuclear bodies located in the nucleus completely randomly (i.e., randomly uniformly), the hypothesis of random distribution was incorrectly rejected at the rate of 15% in the

two-sided *t*-test at a significance level of 5% (sample size 30 nuclei).

A relatively simple permutation test may be used as an alternative method to determine whether G-objects (genes) tend to cluster with B-objects (nuclear bodies). As an example we consider a simple model situation in which six genes are probabilistically associated with 15 bodies. The maximal distance between the centers of objects of two types (e.g. the maximal radius of a nuclear body) is 10% of the diameter of the circular nucleus. We use cNN distances to measure the degree of clustering between objects of different types. To determine whether these distances differ significantly from what

we would expect if the positioning of B-objects and G-objects was independent of each other, we developed a simple permutation test (Figure 11).

For simulated data, this permutation test performed markedly better than the paired *t*-test described above. Its rate of false positives corresponds to the nominal significance level if positioning of G- and B-objects is independent. In almost all other situations that we considered (varying numbers of B- and G-objects with different average sizes of B-objects and different strengths of clustering between B- and G-objects), the rate of true positives is larger than that of the paired *t*-test. Especially as the number of B-objects increases or the degree of clustering

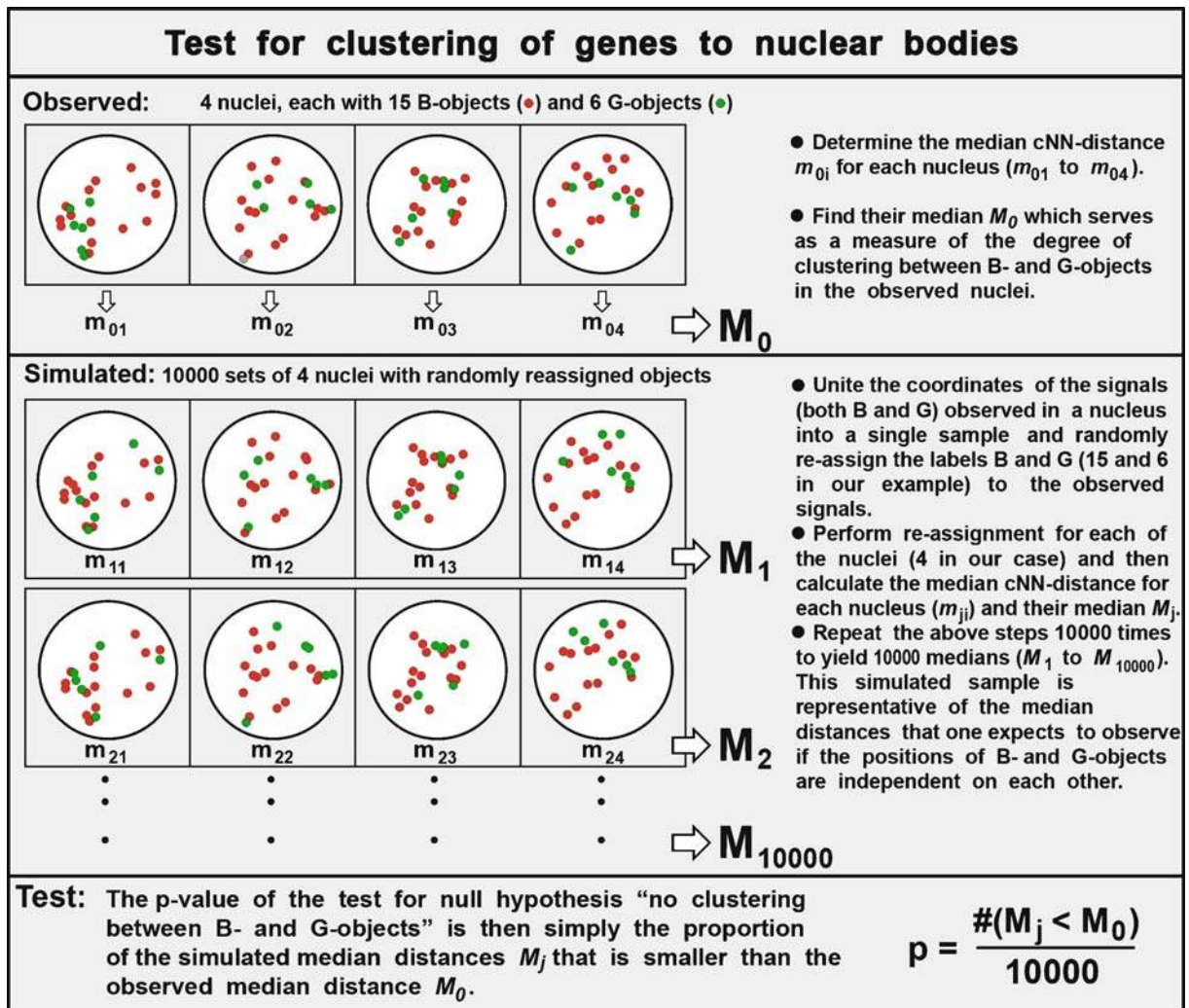


Figure 11. Analysis of the positioning of genes and nuclear bodies.

between G- and B-objects becomes weaker, the paired *t*-test often shows misleading significant *positive* values of the difference ( $d0_i - d1_i$ ), indicating exclusion between G- and B-objects, even if in reality they are clustered. However, the power of both the permutation test and the paired *t*-test for identifying weak clustering is low (data not shown).

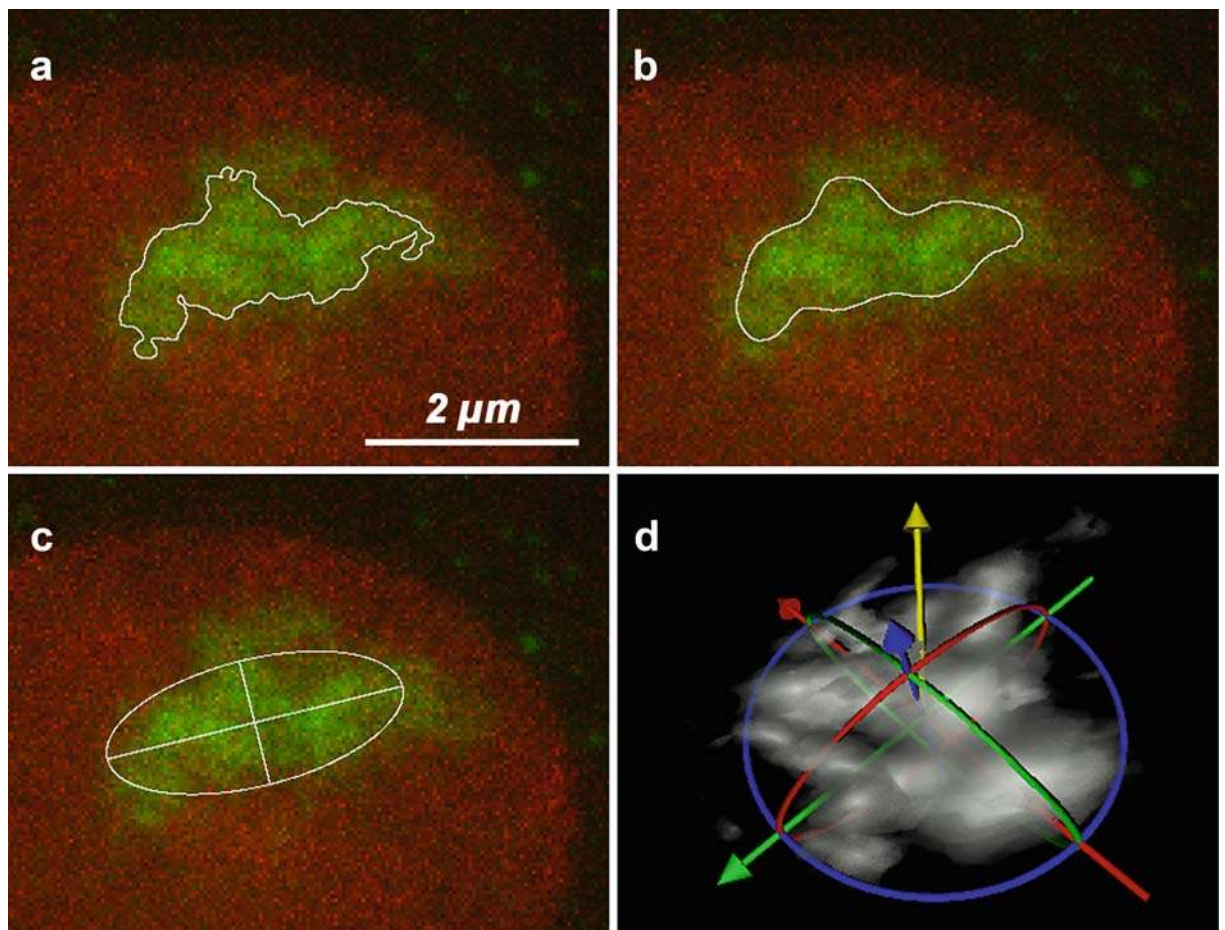
### Shape and orientation of objects

In addition to non-random positioning, nuclear structures may have specific shapes and orientations

related to their functions. A well-known example of this kind is difference in shape between chromosome territories of active and inactive X chromosomes (Eils et al. 1996).

### Roundness factors

Roundness factors (in a general sense) are shape parameters comparing the analyzed object with a sphere. They are ratios of surface areas and volumes measured and calculated (e.g., ratio of the measured volume of an object and the volume of a sphere with the same surface area as that measured for the



**Figure 12.** Description of the shape and orientation of objects. (a, b) Description of shape with roundness factors. The measured perimeter of an object in 2D (correspondingly its surface area in 3D) depends strongly on the selected landmark density: the perimeter of the profile of a chromosome territory is 16.5  $\mu\text{m}$  with fine landmark density (a) and only 11.6  $\mu\text{m}$  with a coarser landmark density (b). The area is less affected: 5.66  $\mu\text{m}^2$  and 5.55  $\mu\text{m}^2$ , respectively. The resulting roundness factors are very different: 0.26 and 0.52. (c, d) Description of shape and orientation by PCA. The shape of an object can be approximated with an ellipse in 2D (c) and with an ellipsoid in 3D (d). The radii of the ellipse are the square root of the eigenvalues and the orientation of the axes is defined by the eigenvectors. The yellow arrow shows the orientation of the optical axis.

object). Hence, they directly use surface areas and volumes. Although ratios can reduce the effects of arbitrary threshold setting, the compensation will be satisfactory only if a certain linearity condition is satisfied. Moreover, calculating surface areas or circumferences from images is not straightforward: for highly structured (fractal) surfaces (like those of chromosome territories), a single *true* surface area simply does not exist! The area obtained always increases if the density of landmarks used to calculate surfaces is increased. If smoothing is applied to the raw data or to the extracted surface, the area obtained will decrease. Volume measurements are usually not sensitive to the landmark density or smoothing: as a result, the landmark density greatly affects the roundness factor value (Figure 12a,b). Furthermore, roundness factors cannot differentiate between a star and an ellipse/ellipsoid.

#### *Principal component analysis*

A more robust alternative for shape description is principal component analysis (PCA). In this case the spatial distribution of intensities within the object is analyzed. In other words, one thinks of an object as a cloud built of voxels with different intensities. PCA is equivalent to fitting a 2D or 3D Gaussian distribution to the object, and can be depicted by fitting an ellipse (2D) or an ellipsoid (3D) to the object. PCA finds the direction of the largest variance (standard deviation), the direction of the second largest variance, perpendicular to the first, and so on. The results are eigenvectors (pointing in the found directions) and eigenvalues (the variances in these directions). They can be depicted by drawing an ellipse (2D) or an ellipsoid (3D) whose radii are proportional to the square roots of the eigenvalues (Figure 12c,d). Although describing a cloud of irregular form (as chromosome territories often look like) by an ellipsoid might seem oversimplified, it should be understood that a more detailed shape description would also demand more parameters. Ellipses and ellipsoids are the best description of a shape that may be given for 2D and 3D shapes with only two or three parameters, respectively.

Principal components are threshold-robust in the same manner as the centers of objects (see Robustness of measurements and Figure 5) and are computed in a way similar to calculation of centers. Of course,

PCA does not differentiate between a symmetrical star and a circle. Furthermore, PCA is based on variances and therefore has the same ‘shortcomings’ as variances in 1D, namely a high sensitivity to outliers (individual voxels situated far from the considered object will have a disproportionate effect on the shape of the estimated ellipse/ellipsoid). Nevertheless, PCA assures a reasonable and robust description of overall shapes. Aspect ratios (ratios of eigenvalues) allow one to differentiate between overall shapes: sphere-like, spindle-like, disk-like. Directions of axes may be compared to a specific direction or plane. For example, one can compare orientation of the nucleus in relation to some surface (e.g., of the slide on which cells were grown, to give the simplest example) (Figure 12d) or analyze the orientation of a structure with regard to the inward/outward direction in the nucleus.

Cells in some tissues are polarized, symmetrically (e.g., smooth-muscle cells) or in a unipolar fashion (e.g., epithelial cells). Correspondingly, the question rises about the positions of chromosome territories and chromosomal subregions relative to the axes of such cells. Preliminary data (T. Cremer’s laboratory, unpublished) suggest that in such cases centromeres indeed may have differential non-random distribution along the nuclear axis. PCA is useful here to assess the orientation of the nucleus. Some of the problems of this kind may be studied using the dedicated software, FILO, that is now under development (P.R. Gudla and S.J. Lockett, personal communication).

#### **Co-localization**

The term co-localization is often used to describe the positioning of two objects close to one other, which we discussed in the previous paragraph. Here we consider co-localization in the narrow sense: the trend for signals from different fluorochrome tags to be present in the same or in different pixels. If the centres of objects labeled by different fluorochromes tend to be situated at a distance notably smaller than characteristic sizes of these objects, or the size of the larger of the two objects, then fluorochromes co-localize or one of them localizes to the other, respectively. One can also draw a line through an object and consider the distributions of the two fluorochrome intensities along this line (Bolte *et al.* 2004). Provided that the signal-to-noise ratio in the

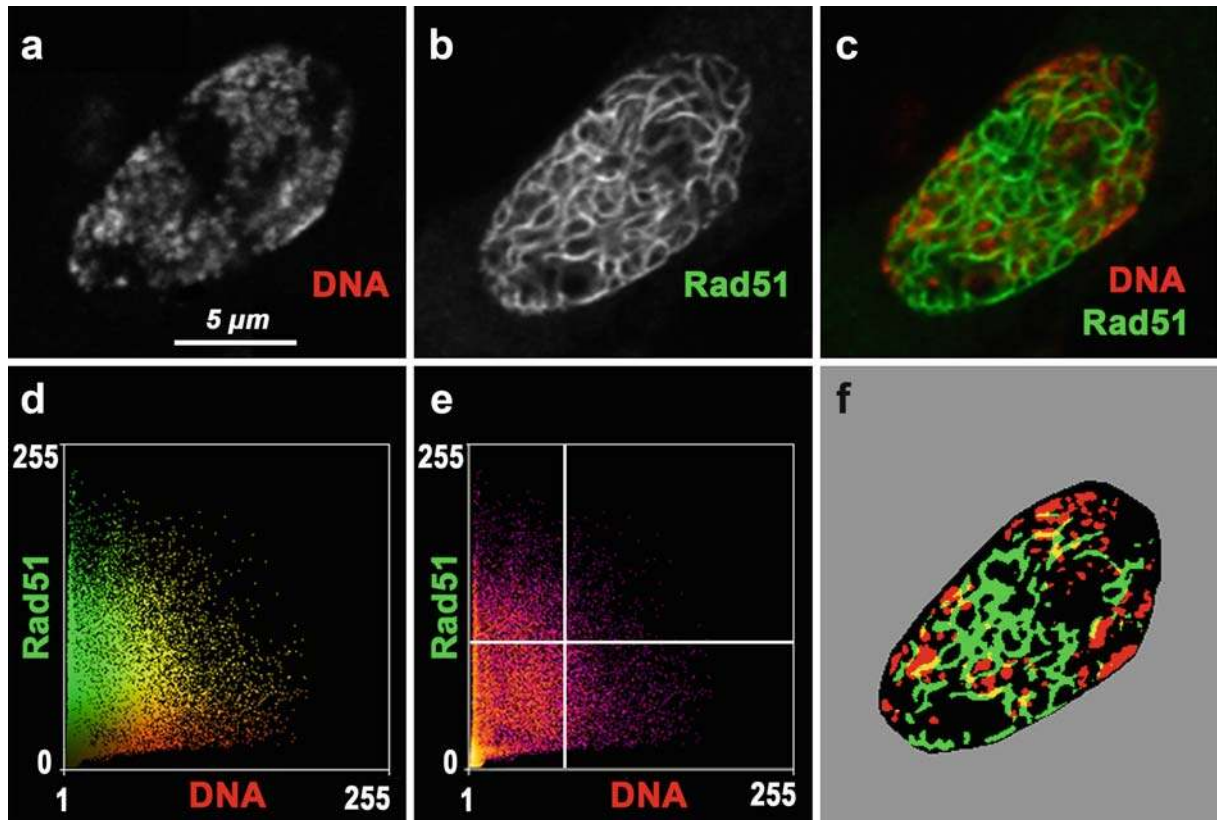
object is sufficiently high, the degree to which FWHM values (full width at half maximum: the length of the line drawn through an intensity peak at the middle of its height) of two channels overlap suggests a measure for co-localization, especially appropriate when the sizes of the objects are known. However, this method is time consuming (Bolte & Cordelieres 2006) and correction for chromatic shift should be taken into account for small objects. Correction for chromatic shift is of course important for any co-localization study (see Chromatic shift: measuring and correcting for it).

Other methods of co-localization analysis do not consider objects and do not take into account positions of pixels: the values will not change if pixels are shuffled—the same way for both channels, of course. These methods consider pixel intensities. All these methods can be implemented using a collection of plugins for ImageJ. Three extreme cases are possible: (1) co-localization, the more Red, the more Green; (2) exclusion (or contra-localization), the more Red, the less Green; (3) no co-localization, intensities of Red and Green are independent. The usual statistical measure in such cases is the Pearson correlation coefficient, which is indeed often used to study co-localization (Manders *et al.* 1993, Bolte & Cordelieres 2006). Unfortunately, this solution is often not appropriate. A good qualitative idea about the relation between intensities of two fluorochromes is given by plots showing relative frequency of pixels with given intensities in the Red and Green channels (Figure 13d,e). Nearly always such plots show that the majority of pixels in the image have low signals in both channels (Figure 13e). This is not necessarily caused by background: many labeled macromolecules do indeed show high concentrations in small nuclear domains, while in large parts of the nucleus the genuine concentrations of fluorochrome are very low, close to the background level. A good example of such a distribution is given by many histone modifications: the same modifications are present essentially everywhere, though they may be notably enriched in chromatin regions with certain transcriptional status (Vakoc *et al.* 2005, Kouzarides 2007). The high proportion of pixels with low intensities dramatically affects the Pearson correlation coefficient; in particular, strong negative correlation is observed very rarely for unprocessed images. Manders *et al.* (1993) suggested the overlap coefficient, which is similar to the Pearson correla-

tion but not sensitive to coincident small intensities in both channels. However, it still depends strongly on the proportion of pixels with small intensities because of pixels where the intensity in the other channel is high. For instance, data shown in Figure 13 yield Pearson correlation  $R = -0.03$  and Manders overlap coefficient  $R_m = 0.69$ . Both these values are not too different from 0 and 0.5, representing independence of channels. Exclusion from the analysis of the pixels with small intensities by applying the thresholds shown in Figure 13e results in  $R = -0.79$  and  $R_m = 0.13$ , which indicate mutual exclusion. Still, because pixels in the left bottom corner of the used range persist, the observed values remain notably different from those showing complete contra-localization ( $-1$  and  $0$ , for  $R$  and  $R_m$ , respectively). Useful qualitative information may also be obtained by mapping fluorochrome intensities to the image: in the simple map showing the distribution of intensities above and below the chosen thresholds one immediately sees that at the sites where fluorochromes are present in high concentration, they avoid each other (Figure 13f). Maps may also be based on ratios of fluorochrome intensities (setting a reasonable threshold is important also for these maps, because ratios of low intensities are not informative).

Manders and co-authors (1993) also suggested several measures that characterize co-localization with regard to each channel: in particular, the proportion of the total intensity in one channel coincident with non-zero intensity in the other channel (this parameter describes how much of, e.g., the Red is coincident with Green). Later Costes *et al.* (2004) and van Steensel *et al.* (1996) suggested approaches that estimate Pearson correlation after transformation of one of the channels. It is noteworthy that if two fluorochromes co-localize, the intensities in a pixel will tend to be either both above the mean level, or both below it. On this basis Li *et al.* (2004) developed ICA (intensity correlation analysis), which makes use of PDM (product of the differences from the mean) values for pixels.  $PDM = (A - a)(B - b)$ , where the upper-case letters denote pixel intensities in two channels and lower-case letters are respective means for the whole image. ICA plots show qualitatively the trend to co-localization, contra-localization, or independent distribution in a very clear way ('symmetrical' for co- and contra-localization) and allow PDM mapping



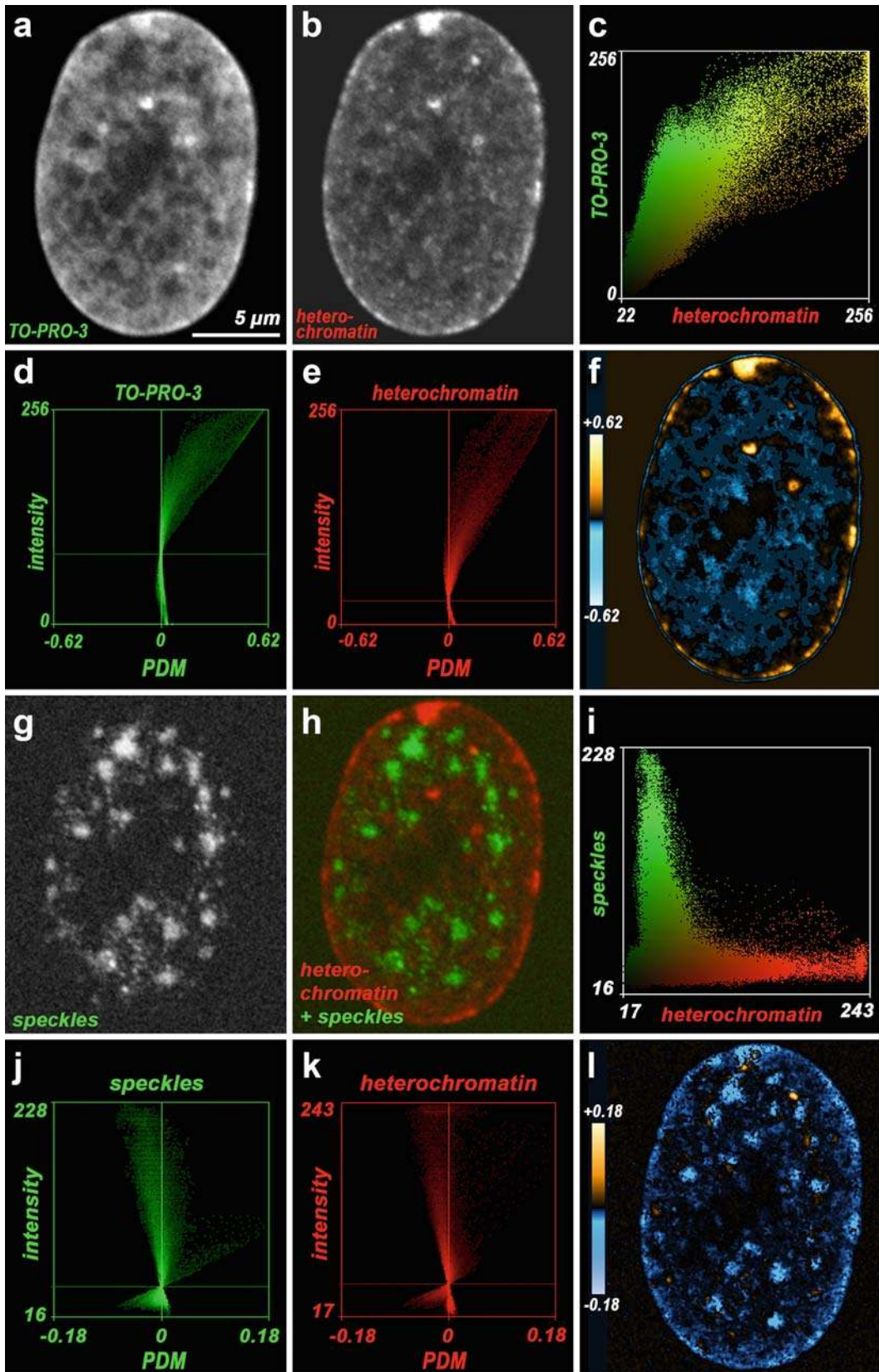


**Figure 13.** An example of co-localization analysis: DNA and over-expressed Rad51-GFP in nuclei of transgenic human fibroblasts. (a) Replicationally labeled DNA. (b) Over-expressed Rad51-GFP forming 'fibers'. (c) Overlay. Images (a–c) show an optical mid-section of the nucleus. (d) Color scatterplot shows which combinations of intensities of the two channels in a pixel are present in the image and which color they have. (e) Frequency scatterplot shows how common are the pixels with certain combinations of intensities. Frequency is coded by color: from dark magenta marking single pixels through red to yellow marking the most common pixels with low intensity in both channels. Setting intensity thresholds corresponds to drawing lines in a plot. (f) Distribution of pixel intensities in the nucleus. Red, intensity of DNA signal above the threshold; green, intensities of Rad51 is above the threshold; yellow, both above the thresholds; black, both below the thresholds.

of the image (Figure 14). For the quantitative estimation of co-localization, the intensity correlation quotient (ICQ) is used (Li *et al.* 2004, Khanna *et al.* 2007). ICQ is the proportion of pixels with PDM > 0 reduced by 0.5 to vary from  $-0.5$  to  $0.5$  (complete contra- and co-localization, respectively). As PDM depends on mean intensities, ICQ is also threshold dependent.

None of the measures discussed above may be recommended for all cases. Our experience suggests that different measures may be more appropriate for different samples. If the distribution of fluorochromes is asymmetrical (e.g., small spots of the Red on large patches of Green), asymmetrical measures (e.g., the proportion of Red pixels coincident with Green ones, weighted or not weighted by

the intensities of Red) correspond better to the nature of data. The sensitivity of all quantitative measures of co-localization to intensity thresholds (and the combination of thresholds for the two channels) raises two serious problems. The first problem is uniform handling of nuclei within a sample. Here one currently has to rely on arbitrary setting of thresholds by the researcher (the results of automatic thresholding or using a certain centile of integrated intensity rarely look satisfactory to biologists; see also Intensity normalization). Secondly—and even more important—is the range of intensities to be analyzed. On one hand, when fluorescence intensities are low, a threshold is difficult to set; on the other hand, genuinely low intensities cannot be ignored (Fay *et al.* 1997, Tashiro *et al.* 2000). Deconvolution



has been reported to improve the results of co-localization analysis (Landmann & Marbet 2004; Sedarat *et al.* 2004), however, problems with the choice of the deconvolution parameters and reproducibility of the deconvolution results cannot be ignored (see Deconvolution), especially with regard to diffuse signals with low intensity. To work around this problem, Tashiro *et al.* (2000) used the solution discussed by us in general in the earlier section Estimating the error associated with arbitrary settings. They considered co-localization at a low threshold level (all signal included, even though some noise too) and a high level (nothing but signal is included, albeit not all signal). Furthermore, the co-localization trend may be different for high and low fluorochrome intensities (ICA graphs show this point especially clearly). Let us consider a simple example. Medics and patients co-localize pretty well in hospitals where their concentrations are high, do not co-localize in general, and may even contra-localize when richer and poorer quarters of a city are considered: all three facts are informative. For the analysis of co-localization it is strongly advisable—irrespective of the measure of co-localization—to use a range of thresholds (to assure robustness) and/or analyze co-localization differentially for appropriate intensity ranges.

Figure 14. Co-localization analysis using ICA: nuclear speckles, heterochromatin and nuclear counterstain in human fibroblasts. (a) Nuclear counterstain, TO-PRO-3. (b) Heterochromatin, immunostaining with an antibody against anti-H3K27me3. (c) Color scatter plot. (d, e) ICA plots for counterstain; crosshair shows mean intensities for TO-PRO-3 (d) and heterochromatin (e). In both plots, points corresponding to higher intensity levels are in the positive PDM range, showing that fluorochromes are co-localized. (f) PDM mapping of the nucleus. Color code for PDM values is shown in the scale on the left side of the figure. The regions where intensities strongly covariate are highlighted in yellow. In this case they are represented mainly by the layer of peripheral heterochromatin along the nuclear border. Brighter blue staining marks regions enriched in H3K27me3. (g) Nuclear speckles, staining with an antibody against the speckle marker protein SC35. (h) Overlay of the heterochromatin and speckle staining. (i) Color scatter plot. (j, k) ICA plots for speckles (j) and heterochromatin (k). In both plots points corresponding to higher intensity levels are in the negative PDM range, showing that fluorochromes are contra-localized. (l) PDM mapping of the nucleus. Mutual exclusion of fluorochromes (blue) is especially prominent in the layer of peripheral heterochromatin and in speckles. Images (a), (b), (g) and (h) show a confocal mid-section of the nucleus.

## Concluding remarks

The problems of quantitative image analyses of nuclear architecture and cell structure in general discussed in this article are complex. Perspectives for the progress in the next few years depend on ongoing developments in different research areas. Most importantly, the development of new light-microscopic techniques has overcome the classical Abbe limit of conventional light-microscopic approaches and allows them to reach a resolution previously restricted to electron microscopy. It can be predicted that structures with dimensions below some 100 nm and quite realistically down to some 10 nm can be resolved by new light-microscopic instrumentation within the next few years. Configurations of transcriptionally active and silent genes within their chromatin context will hopefully be directly visualized. However, we do not expect that light-microscopic setups with ultra-high resolution will outcompete electron-microscopic approaches. The whole range of microscopic approaches must be used in combination to extract the maximum possible information.

Higher resolution makes the demands for reliability of all steps involved in quantitative microscopic analysis even more demanding. In particular, the demands for structural preservation increase strongly in order to solve topographical problems at the ultrastructural level. Although in this article we have focused on the confocal laser scanning microscope, which is still the workhorse for 3D microscopic studies, the problems that need to be solved to assure reliability of quantitative 3D (space) and 4D (space

Table 2. Measures that facilitate reliability of image analysis

- Appropriate choice of the microscope type and good calibrations of the instrument.
- Correct choice of image acquisition settings.
- Planning of the approach to data analysis as a part of planning the biological experiments.
- Appropriate choice of image transformations performed at the preprocessing stage and as a part of image analysis; thorough consideration of the effect of all these transformations and possible artifacts.
- Avoidance of preprocessing and processing that is not clearly justified by the aims of the study and the nature of the raw data.
- Correct choice of appropriate statistical tests.
- Discussion of approaches to preprocessing, processing and statistical analysis with specialists if they are not well established for the intended task.

and time) image analysis (Table 2) are, by and large, the same with all types of microscopic setups. Progress in organic chemistry and molecular biology has already increased the spectrum of specific fluorescent labels for both proteins and nucleic acids for studies not only of fixed but also of living cells. This spectrum will certainly continue to increase and help to develop multicolor labeling schemes with fluorochrome combinations which cover the entire range of the visible spectrum and beyond much more densely than possible to date and with properties that allow the reliable discrimination of all fluorochromes from each other. Bleed-through of fluorescence from one channel into another has remained a major problem of image recording in multicolor formats. Fluorochromes with improved bleaching properties are also urgently needed, which allow tracing even of single fluorescent molecules over long periods of time.

The prospects of imaging techniques at the single-cell level should be weighed against new molecular approaches to study the 3D organization of the genome. For example, it has recently become possible to detect close proximity of genomic loci in the interphase nucleus using purely molecular 3C/4C/5C methods (Simonis *et al.* 2006, Zhao *et al.* 2006, Hagege *et al.* 2007) and a similar method (PLA, proximity ligation analysis) has been suggested for proteins (Soderberg *et al.* 2006). While these methods can pinpoint DNA and protein interactions *in cis* (loci on the same chromosome) and *in trans* (loci on different chromosomes), their power is not unlimited. For instance, discriminative features of the different, presently discussed models of nuclear architecture (see Cremer & Cremer 2006 for review) cannot be tested by 3C/4C/5C methods. Most importantly, 3C/4C/5C methods provide information for DNA interactions on a statistical basis and therefore require large samples of nuclei subjected to cross-linking of DNA and proteins. Nevertheless, new factual data obtained using this and other methods have attracted attention to the spatial parameters of nuclear architecture that only recently did not seem functionally important. Despite the simplicity of their formulation, many of the respective problems are very complex. Improved methods of imaging and image analysis will help to solve these problems. It is the combination of new molecular, microscopic and image analysis methods

that will determine progress of nuclear studies in the near future.

### Acknowledgements

The material used for illustration was kindly supplied by I. Solovei, D. Koehler, C. Casas-Delucchi, S. Brero (University of Munich), S. Tashiro (Hiroshima University), and M. Keuper (University of Freiburg). The authors are grateful to all colleagues who have discussed with us the problems considered in the article. The work was supported by DFG CR60/23–2 to C. Cremer and Cr59/22–3 to T. Cremer.

### References

- Akhtar A, Gasser SM (2007) The nuclear envelope and transcriptional control. *Nat Rev Genet* **8**: 507–517.
- Albiez H, Cremer M, Tiberi C *et al.* (2006) Chromatin domains and the interchromatin compartment form structurally defined and functionally interacting nuclear networks. *Chromosome Res* **14**: 707–733.
- Bacher CP, Guggiari M, Brors B *et al.* (2006) Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat Cell Biol* **8**: 293–299.
- Baddeley A, Turner R (2005) Spatstat: an R package for analyzing spatial point patterns. *J Stat Software* **12**: 1–42.
- Beil M, Fleischer F, Paschke S, Schmidt V (2005) Statistical analysis of the three-dimensional structure of centromeric heterochromatin in interphase nuclei. *J Microsc* **217**: 60–68.
- Betzig E, Patterson GH, Sougrat R *et al.* (2006) Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**: 1642–1645.
- Boei JJ, Fomina J, Darroudi F, Nagelkerke NJ, Mullenders LH (2006) Interphase chromosome positioning affects the spectrum of radiation-induced chromosomal aberrations. *Radiat Res* **166**: 319–326.
- Bolte S, Cordelieres FP (2006) A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* **224**: 213–232.
- Bolte S, Brown S, Satiat-Jeunemaitre B (2004) The N-myristoylated Rab-GTPase m-Rabmc is involved in post-Golgi trafficking events to the lytic vacuole in plant cells. *J Cell Sci* **117**: 943–954.
- Bolzer A, Kreth G, Solovei I *et al.* (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol* **3**: e157.
- Boutet de Monvel J, Scarfone E, Le Calvez S, Ulfendahl M (2003) Image-adaptive deconvolution for three-dimensional deep biological imaging. *Biophys J* **85**: 3991–4001.
- Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA (2001) The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet* **10**: 211–219.

- Brezger A, Fahmeir L, Hennerfeind A (2007) Adaptive Gaussian Markov random fields with applications in human brain mapping. *J Roy Statist Soc, Ser C (Applied Statistics)* **56**: 327–345.
- Bridger JM, Boyle S, Kill IR, Bickmore WA (2000) Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. *Curr Biol* **10**: 149–152.
- Brown JM, Leach J, Reittie JE *et al.* (2006) Coregulated human globin genes are frequently in spatial proximity when active. *J Cell Biol* **172**: 177–187.
- Buser C, Fleischer F, Mertens T, Michel D, Schmidt V, Walther P (2007) Quantitative investigation of murine cytomegalovirus nucleocapsid interaction. *J Microsc* **228**: 78–87.
- Cannell MB, McMorland A, Soeller C (2006) Image enhancement by deconvolution. In: Pawley JB, ed. *Handbook of Biological Confocal Microscopy*. New York: Springer, pp. 488–500.
- Carlton PM (2008) Three-dimensional structured illumination microscopy and its application to chromosome structure. *Chromosome Res* **16**(3): 351–365.
- Chambeyron S, Bickmore WA (2004) Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* **18**: 1119–1130.
- Chuang CH, Belmont AS (2007) Moving chromatin within the interphase nucleus-controlled transitions? *Semin Cell Dev Biol* **18**: 698–706.
- Chuang CH, Carpenter AE, Fuchsova B, Johnson T, de Lanerolle P, Belmont AS (2006) Long-range directional movement of an interphase chromosome site. *Curr Biol* **16**: 825–831.
- Cobb BS, Morales-Alcelay S, Kleiger G, Brown KE, Fisher AG, Smale ST (2000) Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev* **14**: 2146–2160.
- Conchello JA, Lichtman JW (2005) Optical sectioning microscopy. *Nat Methods* **2**: 920–931.
- Costes SV, Daelemans D, Cho EH, Dobbin Z, Pavlakis G, Lockett S (2004) Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys J* **86**: 3993–4003.
- Cremer T, Cremer C (2006) Rise, fall and resurrection of chromosome territories: a historical perspective. Part II. Fall and resurrection of chromosome territories during the 1950s to 1980s. Part III. Chromosome territories and the functional nuclear architecture: experiments and models from the 1990s to the present. *Eur J Histochem* **50**: 223–272.
- Cremer M, von Hase J, Volm T *et al.* (2001) Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res* **9**: 541–567.
- Cremer M, Kupper K, Wagler B *et al.* (2003) Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. *J Cell Biol* **162**: 809–820.
- Cremer T, Cremer M, Dietzel S, Muller S, Solovei I, Fakan S (2006) Chromosome territories—a functional nuclear landscape. *Curr Opin Cell Biol* **18**: 307–316.
- Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA (1999) Differences in the localization and morphology of chromosomes in the human nucleus. *J Cell Biol* **145**: 1119–1131.
- Denk W, Strickler JH, Webb WW (1990) Two-photon laser scanning fluorescence microscopy. *Science* **248**: 73–76.
- Dietzel S, Zolghadr K, Heppner C, Belmont AS (2004) Differential large-scale chromatin compaction and intranuclear positioning of transcribed versus non-transcribed transgene arrays containing beta-globin regulatory sequences. *J Cell Sci* **117**: 4603–4614.
- Dundr M, Ospina JK, Sung MH *et al.* (2007) Actin-dependent intranuclear repositioning of an active gene locus in vivo. *J Cell Biol* **179**: 1095–1103.
- Dyba M, Hell SW (2002) Focal spots of size  $\lambda/23$  open up far-field fluorescence microscopy at 33 nm axial resolution. *Phys Rev Lett* **88**: 163901.
- Eils R, Dietzel S, Bertin E *et al.* (1996) Three-dimensional reconstruction of painted human interphase chromosomes: active and inactive X chromosome territories have similar volumes but differ in shape and surface structure. *J Cell Biol* **135**: 1427–1440.
- Fay FS, Taneja KL, Shenoy S, Lifshitz L, Singer RH (1997) Quantitative digital analysis of diffuse and concentrated nuclear distributions of nascent transcripts, SC35 and poly(A). *Exp Cell Res* **231**: 27–37.
- Fehr J, Ronneberger O, Kurz H, Burkhardt H (2005) Self-learning segmentation and classification of cell nuclei in 3D volumetric data using voxelwise gray scale invariants. In: Kropatsch W, Sablating R, eds. *Pattern Recognition. Proceedings of the 27th DAGM*. Berlin: Springer, pp. 377–384.
- Fölling JV, Belov R, Kunetsky R *et al.* (2007) Photochromic Rhodamines provide Nanoscopy with Optical Sectioning. *Angew Chem Int Ed* **46**: 6266–6270.
- Foster HA, Bridger JM (2005) The genome and the nucleus: a marriage made by evolution. Genome organisation and nuclear architecture. *Chromosoma* **114**: 212–229.
- Fraser P, Bickmore W (2007) Nuclear organization of the genome and the potential for gene regulation. *Nature* **447**: 413–417.
- Geisler C, Schönle A, von Middendorf C *et al.* (2007) Resolution of Lambda /10 in fluorescence microscopy using fast single molecule photo-switching. *Appl Phys, A* **88**: 223–226.
- Genovesio A, Liedl T, Emiliani V, Parak WJ, Copepy-Moisan M, Olivo-Marin J-C (2006) Multiple particle tracking in 3-D+t microscopy: method and application to the tracking of endocytosed Quantum Dots. *IEEE Trans Image Processing* **15**: 1062–1070.
- Gerlich D, Beaudouin J, Kalbfuss B, Daigle N, Eils R, Ellenberg J (2003) Global chromosome positions are transmitted through mitosis in mammalian cells. *Cell* **112**: 751–764.
- Geman D, Reynolds G (1992) Constrained restoration and the recovery of discontinuities. *IEEE Trans Pattern Anal Mach Intell* **14**: 367–383.
- Gudla PR, Nandy K, Collins J, Meaburn KJ, Misteli T, Lockett SJ (2008) A high-throughput system for segmenting nuclei using multiscale techniques. *Cytometry A* (in press).
- Gustafsson MG (2000) Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J Microsc* **198**: 82–87.
- Habermann FA, Cremer M, Walter J *et al.* (2001) Arrangements of macro- and microchromosomes in chicken cells. *Chromosome Res* **9**: 569–584.
- Hagege H, Klous P, Braem C *et al.* (2007) Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat Protoc* **2**: 1722–1733.
- Hell SW, Stelzer EHK (1992) Fundamental improvement of resolution with a 4Pi-confocal fluorescence microscope using two-photon excitation. *Opt Commun* **93**: 277–282.
- Hell SW, Lindek S, Cremer C, Stelzer EHK (1994) Measurement of 4pi-confocal point spread function proves 75 nm axial resolution. *Appl Phys Lett* **64**: 1335–1337.

- Hess ST, Girirajan TP, Mason MD (2006) Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys J* **91**: 4258–4272.
- Hess ST, Gould TJ, Gudheti MV, Maas SA, Mills KD, Zimmerberg J (2007) Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. *Proc Natl Acad Sci U S A* **104**: 17370–17375.
- Holmes TJ, Biggs D, Abu-Tarif A (2006) Blind deconvolution. In: Pawley JB, ed. *Handbook of Biological Confocal Microscopy*. New York: Springer, pp. 468–487.
- Huang B, Wang W, Bates M, Zhuang X (2008) Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* **319**: 810–813.
- Jinno S, Fleischer F, Eckel S, Schmidt V, Kosaka T (2007) Spatial arrangement of microglia in the mouse hippocampus: a stereological study in comparison with astrocytes. *Glia* **55**: 1334–1347.
- Khanna R, Li Q, Bewersdorf J, Stanley EF (2007) The presynaptic CaV2.2 channel-transmitter release site core complex. *Eur J Neurosci* **26**: 547–559.
- Klar TA, Jakobs S, Dyba M, Egnér A, Hell SW (2000) Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc Natl Acad Sci U S A* **97**: 8206–8210.
- Koberna K, Ligasová A, Malinský J et al. (2005) Replication in 3-D: Evidence for Similar-Sized Replication Foci Throughout S-Phase. *J Cell Biochem* **94**: 126–138.
- Kosak ST, Groudine M (2004) Form follows function: the genomic organization of cellular differentiation. *Genes Dev* **18**: 1371–1384.
- Kosak ST, Skok JA, Medina KL et al. (2002) Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* **296**: 158–162.
- Kosak ST, Scalzo D, Alworth SV et al. (2007) Coordinate gene regulation during hematopoiesis is related to genomic organization. *PLoS Biol* **5**: e309.
- Kouzarides T (2007) SnapShot: histone-modifying enzymes. *Cell* **131**: 822.
- Kozubek S, Lukasova E, Amrichova J, Kozubek M, Liskova A, Slotova J (2000) Influence of cell fixation on chromatin topography. *Anal Biochem* **282**: 29–38.
- Kozubek S, Lukasova E, Jirsova P et al. (2002) 3D Structure of the human genome: order in randomness. *Chromosoma* **111**: 321–331.
- Kupper K, Kolbl A, Biener D et al. (2007) Radial chromatin positioning is shaped by local gene density, not by gene expression. *Chromosoma* **116**: 285–306.
- Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* **8**: 104–115.
- Landmann L, Marbet P (2004) Colocalization analysis yields superior results after image restoration. *Microsc Res Techn* **64**: 103–112.
- Li Q, Lau A, Morris TJ, Guo L, Fordyce CB, Stanley EF (2004) A syntaxin 1, Galpha(o), and N-type calcium channel complex at a presynaptic nerve terminal: analysis by quantitative immunocolocalization. *J Neurosci* **24**: 4070–4081.
- Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R (2006) Interchromosomal interactions and olfactory receptor choice. *Cell* **126**: 403–413.
- Mahy NL, Perry PE, Bickmore WA (2002) Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J Cell Biol* **159**: 753–763.
- Manders EMM, Verbeek FJ, Aten JA (1993) Measurements of colocalization of objects in dual-color confocal images. *J Microsc* **169**: 375–382.
- Manley S, Gillette JM, Patterson GH et al. (2008) High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nat Methods* **5**: 155–157.
- Markham J, Conchello JA (2001) Artefacts in restored images due to intensity loss in three-dimensional fluorescence microscopy. *J Microsc* **204**: 93–98.
- Martini N, Bewersdorf J, Hell SW (2002) A new high-aperture glycerol immersion objective lens and its application to 3D-fluorescence microscopy. *J Microsc* **206**: 146–151.
- Mayer R, Brero A, von Hase J, Schroeder T, Cremer T, Dietzel S (2005) Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol* **6**: 44.
- Mayer V, Schmidt V, Schweiggert F (2004) A unified simulation framework for spatial stochastic models. *Simulation Modelling Practice and Theory* **12**: 307–326.
- McNally JG, Karpova T, Cooper J, Conchello JA (1999) Three-dimensional imaging by deconvolution microscopy. *Methods* **19**: 373–385.
- Meaburn KJ, Misteli T, Soutoglou E (2007) Spatial genome organization in the formation of chromosomal translocations. *Semin Cancer Biol* **17**: 80–90.
- Merico V, Barbieri J, Zuccotti M et al. (2007) Epigenomic differentiation in mouse preimplantation nuclei of biparental, parthenote and cloned embryos. *Chromosome Res* **15**: 341–360.
- Misteli T (2007) Beyond the sequence: cellular organization of genome function. *Cell* **128**: 787–800.
- Moen PT Jr, Johnson CV, Byron M et al. (2004) Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. *Mol Biol Cell* **15**: 197–206.
- Morey C, Da Silva NR, Perry P, Bickmore WA (2007) Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development* **134**: 909–919.
- Murray JM, Appleton PL, Swedlow JR, Waters JC (2007) Evaluating performance in three-dimensional fluorescence microscopy. *J Microsc* **228**: 390–405.
- Neusser M, Schubel V, Koch A, Cremer T, Müller S (2007) Evolutionarily conserved, cell type and species-specific higher order chromatin arrangements in interphase nuclei of primates. *Chromosoma* **116**: 307–320.
- Neves H, Ramos C, da Silva MG, Parreira A, Parreira L (1999) The nuclear topography of ABL, BCR, PML, and RARalpha genes: evidence for gene proximity in specific phases of the cell cycle and stages of hematopoietic differentiation. *Blood* **93**: 1197–1207.
- Olivio-Marin J-C (2002) Extraction of spots in biological images using multiscale products. *Pattern Recognition* **35**: 1989–1996.
- Ono A, Kono K, Ikebe D et al. (2007) Nuclear positioning of the BACH2 gene in BCR-ABL positive leukemic cells. *Genes Chromosomes Cancer* **46**: 67–74.
- Osborne CS, Chakalova L, Mitchell JA et al. (2007) Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol* **5**: e192.

- Parada LA, McQueen PG, Munson PJ, Misteli T (2002) Conservation of relative chromosome positioning in normal and cancer cells. *Curr Biol* **12**: 1692–1697.
- Park PC, De Boni U (1998) A specific conformation of the territory of chromosome 17 locates ERBB-2 sequences to a DNase-hypersensitive domain at the nuclear periphery. *Chromosoma* **107**: 87–95.
- Pawley JB, ed. (2006) *Handbook Of Biological Confocal Microscopy*, 3rd edn. New York: Springer.
- Ragoczy T, Bender MA, Telling A, Byron R, Groudine M (2006) The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev* **20**: 1447–1457.
- Reisert M, Ronneberger O, Burkhardt H (2007) Holomorphic filters for object detection. In: Hamprecht FA, Schnörr C, Jähne B, eds. *Pattern Recognition. Proceedings of the 29th DAGM*. Berlin: Springer, pp. 415–424.
- Roix JJ, McQueen PG, Munson PJ, Parada LA, Misteli T (2003) Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat Genet* **34**: 287–291.
- Ronneberger O, Fehr J, Burkhardt H (2005) Voxelwise gray scale invariants for simultaneous segmentation and classification pattern recognition. In: Kropatsch W, Sablating R, eds. *Pattern Recognition. Proceedings of the 27th DAGM*. Berlin: Springer, pp. 85–92.
- Rust MJ, Bates M, Zhuang X (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* **3**: 793–795.
- Sedarat F, Lin E, Moore EDW, Tibbits GF (2004) Deconvolution of confocal images of dihydropyridine and ryanodine receptors in developing cardiomyocytes. *J Appl Physiol* **97**: 1098–1103.
- Sexton T, Schober H, Fraser P, Gasser SM (2007) Gene regulation through nuclear organization. *Nat Struct Mol Biol* **14**: 1049–1055.
- Shiels C, Islam SA, Vatcheva R *et al.* (2001) PML bodies associate specifically with the MHC gene cluster in interphase nuclei. *J Cell Sci* **114**: 3705–3716.
- Shiels C, Adams NM, Islam SA, Stephens DA, Freemont PS (2007) Quantitative analysis of cell nucleus organisation. *PLoS Comput Biol* **3**: e138.
- Shopland LS, Johnson CV, Byron M, McNeil J, Lawrence JB (2003) Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. *J Cell Biol* **162**: 981–990.
- Simonis M, Klous P, Splinter E *et al.* (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet* **38**: 1348–1354.
- Simonis M, Kooren J, de Laat W (2007) An evaluation of 3C-based methods to capture DNA interactions. *Nat Methods* **4**: 895–901.
- Soderberg O, Gullberg M, Jarvius M *et al.* (2006) Direct observation of individual endogenous protein complexes *in situ* by proximity ligation. *Nat Methods* **3**: 995–1000.
- Soutoglou E, Misteli T (2007) Mobility and immobility of chromatin in transcription and genome stability. *Curr Opin Genet Dev* **17**: 435–442.
- Soutoglou E, Dorn JF, Sengupta K *et al.* (2007) Positional stability of single double-strand breaks in mammalian cells. *Nat Cell Biol* **9**: 675–682.
- Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA (2005) Interchromosomal associations between alternatively expressed loci. *Nature* **435**: 637–645.
- Sun HB, Shen J, Yokota H (2000) Size-dependent positioning of human chromosomes in interphase nuclei. *Biophys J* **79**: 184–190.
- Swedlow JR (2007) Quantitative fluorescence microscopy and image deconvolution. *Methods Cell Biol* **81**: 447–465.
- Taddei A, Van Houwe G, Hediger F *et al.* (2006) Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature* **441**: 774–778.
- Tanabe H, Muller S, Neusser M *et al.* (2002) Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proc Natl Acad Sci U S A* **99**: 4424–4429.
- Tashiro S, Walter J, Shinohara A, Kamada N, Cremer T (2000) Rad51 accumulation at sites of DNA damage and in postreplicative chromatin. *J Cell Biol* **150**: 283–291.
- Teller K, Solovei I, Buiting K, Horsthemke B, Cremer T (2007) Maintenance of imprinting and nuclear architecture in cycling cells. *Proc Natl Acad Sci U S A* **104**: 14970–14975.
- Thomson I, Gilchrist S, Bickmore WA, Chubb JR (2004) The radial positioning of chromatin is not inherited through mitosis but is established *de novo* in early G1. *Curr Biol* **14**: 166–172.
- Vakoc CR, Mandat SA, Olenchok BA, Blobel GA (2005) Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell* **19**: 381–391.
- van Steensel B, van Binnendijk EP, Hornsby CD *et al.* (1996) Partial colocalization of glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons. *J Cell Sci* **109**: 787–792.
- Verveer PJ, Swoger J, Pampaloni F, Greger K, Marcello M, Stelzer EH (2007) High-resolution three-dimensional imaging of large specimens with light sheet-based microscopy. *Nat Methods* **4**: 311–313.
- von Tiedemann M, Fridberger A, Ulfendahl M, Tomo I, Boutet de Monvel J (2006) Image adaptive point-spread function estimation and deconvolution for *in vivo* confocal microscopy. *Microsc Res Tech* **69**: 10–20.
- Wallace W, Schaefer LH, Swedlow JR (2001) A workingperson's guide to deconvolution in light microscopy. *Biotechniques* **31**: 1076–1082.
- Walter J, Schermelleh L, Cremer M, Tashiro S, Cremer T (2003) Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. *J Cell Biol* **160**: 685–697.
- Walter J, Joffe B, Bolzer A *et al.* (2006) Towards many colors in FISH on 3D-preserved interphase nuclei. *Cytogenet Genome Res* **114**: 367–378.
- Wang J, Shiels C, Sasieni P *et al.* (2004) Promyelocytic leukemia nuclear bodies associate with transcriptionally active genomic regions. *J Cell Biol* **164**: 515–26.
- Westphal V, Hell SW (2005) Nanoscale resolution in the focal plane of an optical microscope. *Phys Rev Lett* **94**: 143903.
- Willig KI, Harke B, Medda R, Hell SW (2007) STED microscopy with continuous wave beams. *Nat Methods* **4**: 915–918.
- Willig KI, Kellner RR, Medda R, Hein B, Jakobs S, Hell SW (2006) Nanoscale resolution in GFP-based microscopy. *Nat Methods* **3**: 721–723.

- Wolf DE, Samarasekera C, Swedlow JR (2007) Quantitative analysis of digital microscope images. *Methods Cell Biol* **81**: 365–96.
- Xu N, Tsai CL, Lee JT (2006) Transient homologous chromosome pairing marks the onset of X inactivation. *Science* **311**: 1149–52.
- Yokota H, van den Engh G, Hearst JE, Sachs RK, Trask BJ (1995) Evidence for the organization of chromatin in megabase pair-sized loops arranged along a random walk path in the human G0/G1 interphase nucleus. *J Cell Biol* **130**: 1239–1249.
- Zhao Z, Tavossidana G, Sjolinder M *et al.* (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet* **38**: 1341–1347.