

Open Microscopy Environment and FindSpots: integrating image informatics with quantitative multidimensional image analysis

David A. Schiffmann¹, Dina Dikovskaya¹, Paul L. Appleton¹, Ian P. Newton¹, Douglas A. Creager², Chris Allan¹, Inke S. Näthke¹, and Ilya G. Goldberg³

BioTechniques 41:199-208 (August 2006)
doi 10.2144/000112224

Biomedical research and drug development increasingly involve the extraction of quantitative data from digital microscope images, such as those obtained using fluorescence microscopy. Here, we describe a novel approach for both managing and analyzing such images. The Open Microscopy Environment (OME) is a sophisticated open-source scientific image management database that coordinates the organization, storage, and analysis of the large volumes of image data typically generated by modern imaging methods. We describe FindSpots, a powerful image-analysis package integrated in OME that will be of use to those who wish to identify and measure objects within microscope images or time-lapse movies. The algorithm used in FindSpots is in fact only one of many possible segmentation (object detection) algorithms, and the underlying data model used by OME to capture and store its results can also be used to store results from other segmentation algorithms. In this report, we illustrate how image segmentation can be achieved in OME using one such implementation of a segmentation algorithm, and how this output subsequently can be displayed graphically or processed numerically using a spreadsheet.

INTRODUCTION

Quantitative analysis of microscope images forms a growing part of modern cell biology. It is all too easy for such studies to generate vast amounts of data, which can become difficult to keep track of (1–5). The Open Microscopy Environment (OME) software system (1,3–6) was developed to provide a flexible, powerful, and freely available platform for storing, viewing, and analyzing such images, and has a clear audit trail to link all images with the analyses performed on them. It allows automated mining and analysis of image sets independent of the particular system originally used to capture the images. It facilitates analysis of these image sets using built-in or third-party image-analysis programs. Because images can be arranged in groups as required, and analyses can be performed on entire groups, automated image analysis of such groups of images is possible to

an extent not readily achievable by methods that require manual analysis of each image separately.

A particular strength of OME is that it allows for a highly structured description (data model) of the information that it stores (6) and encourages the reuse of these data models in different algorithms and “viewers”—software that can display the images themselves or that can take the results of such algorithms and display quantitative information about them (e.g., a spreadsheet application). The benefit of using common data models is the transparent exchange of structured information between independently developed software.

Common analysis tasks in imaging are finding spots or “blobs” in images based on various criteria and obtaining structured information to describe them (positions of the blobs, their sizes, shapes, intensities, etc.). Although various spot-finding tools have been described, none can fully take

advantage of the unique informatics abilities of OME. Here, we describe and validate FindSpots, an analysis package that is fully integrated into OME, version 2.4.

FindSpots can automatically detect and quantify spot-like objects within microscope images, whether they be two-dimensional (2-D) flat images or three-dimensional (3-D) image stacks. It can also be applied to images collected using more than one wavelength and can analyze time-lapse movies, which allows for subsequent tracking of the objects in 3-D. Thus, the algorithm is capable of operating in all five dimensions accessible on a modern digital microscope (4,7). The algorithm underlying FindSpots itself has been used in earlier work (8,9), but the current work represents the first integrated use of FindSpots and OME and the first example in the context of OME of its use to perform complex biological image analysis.

¹University of Dundee, Dundee, Scotland, UK, ²Massachusetts Institute of Technology, Cambridge, MA, and ³National Institutes of Health, Baltimore, MD, USA

THEORETICAL BASIS

Approaches to Object Detection/Tracking

Object finding (and therefore object tracking) in FindSpots and in many other image-analysis programs, including the commercial package Volocity® (Improvision®, Coventry, UK; www.improvision.com/products/Volocity), is based on a procedure known as global thresholding, whereby in any one image, all pixels above a certain minimum value are potentially considered to be part of an object (subject to a minimum spot-size requirement). An unavoidable problem with defining objects based on a threshold—especially a global one—is that whatever value it is set to, some subthreshold pixels that ideally should form part of the objects can be lost. However, lowering the threshold to include these pixels will cause too many artifactual pixels to be included (10). Where spot contrast or cellular background varies widely within an individual image, such unwanted omission or inclusion of pixels is more likely to be a problem, and it may be necessary to use adaptive thresholding algorithms, which allow the threshold to be varied within any one image (11) or employ advanced background correction methods. Indeed, methods other than thresholding (e.g., edge detection) can sometimes be helpful for image segmentation (i.e., the identification of objects within an image) (11). However, with any segmentation method, it is inevitable that some pixels that should form part of spots will be omitted and vice versa; no segmentation algorithm is perfect and any such quantitative analyses will inevitably be approximations (10,12). Moreover, in many cases, a low level of misclassification will not matter in practice.

Additionally, although we present here a simple but effective method for background correction, more advanced variations on this, as well as other specialized algorithms, also exist and are worth noting (www.moleculardevices.com/technotes_d1/MDC_D1_D50022_MorphologyFilters.pdf; References 10, 13, and 14), but a full discussion of these procedures

is beyond the scope of this article. In general, it should be noted that sophisticated segmentation algorithms are no substitute for high-quality image data when performing quantitative image analysis.

Finally, it should be noted that a number of different methods for setting a global threshold exist in the literature (for examples, see References 11, 15, and 16 and references therein), and FindSpots implements a number of these: an absolute intensity threshold, a variety of statistical intensity thresholds, maximum entropy, Kittler, Moment Preservation, and Otsu's. However, in practice only an absolute or statistical intensity threshold are suitable for typical fluorescence microscopy images, as the others will yield inappropriate thresholds for such images.

The tracking method in OME is implemented as a separate routine called TrackSpots. TrackSpots is a simple algorithm that tracks objects found by FindSpots or other segmentation algorithms that use OME's data model to describe spots. For each spot in a time point, TrackSpots searches the subsequent time point for a nearest neighbor and assigns it to the given spot's trajectory. An essentially identical algorithm is used for particle tracking in many other software programs. More advanced tracking methods also exist, although it is beyond the scope of this work to discuss these in detail. These include SpotTracker (17), which is optimized for tracking single particles in very noisy 2-D microscope images, and the TIKAL image-analysis platform (18), which allows 2-D or 3-D tracking of objects within cells while correcting for cell movements or deformations.

Object Finding in FindSpots

A description of the FindSpots algorithm has been published (9). In brief, for each image, FindSpots calculates a threshold intensity level; any pixels below this value are considered to be background signal and are ignored, and any pixels above the threshold may form part of a spot as long as the spot thus formed is above a preset minimum volume. FindSpots defines a spot as a set of adjoining pixels above

the threshold. The algorithm works recursively: starting with an initial seed pixel, it checks all the adjoining pixels once to see if they are above the threshold (i.e., if they are bright enough). Any such pixel is added to the spot being formed, and the process starts again from that pixel, checking all its adjoining pixels for those above the threshold. Whenever an adjoining pixel is found to be below threshold, the algorithm excludes that pixel from the spot and does not search beyond it. This process is mathematically guaranteed to reach every contiguous pixel above the threshold and therefore will work with spots of any shape.

Object Tracking in OME Using TrackSpots

FindSpots can, in combination with the TrackSpots routine, analyze time-lapse movies, and these results can be used to track spots between time frames, assigning them to trajectories.

The initial implementation of TrackSpots (8,9) predated the OME system and was a command-line program that could operate on a single five-dimensional (5-D) file format produced using a DeltaVision® microscope (Applied Precision®, Issaquah, WA, USA). This program performed both finding and tracking functions and produced a table of outputs that could be further processed in Microsoft® Excel®. It could not be exported easily to other file formats and was difficult to work with due to the many files and spreadsheets to keep track of when processing just tens of images, not to say hundreds or thousands. In its present implementation, the algorithm was altered so that it can readily handle any image format OME can import. Notably, the tracking function was removed and implemented in a separate algorithm in order to prototype different tracking techniques more easily and to demonstrate modular analysis in OME by defining common data models.

When using FindSpots with TrackSpots in its current implementation for object tracking, it is recommended to apply it only to images with well-spaced objects in order to avoid the introduction of errors in tracking. The tracking algorithm currently used

is a simple nearest-neighbor search for the spot in the next time point, but this algorithm can be easily replaced and will continue to work with the original FindSpots, or with a different type of segmentation algorithm that can use the same data model for the 5-D objects it finds. The result of using the tracking feature is a set of trajectories composed of the spot objects originally created by FindSpots, as well as additional measurements for each spot such as the X, Y, Z displacement and distance to the next spot.

Threshold Calculation

Although a single and constant threshold can be used for all images, it is preferable to determine the threshold specifically for each image to allow for variation in background levels. This can also be important in time-course data where photobleaching may occur in which setting a fixed threshold based on early time points could lead the routine to miss spots at later times. A number of algorithms to determine threshold are available in FindSpots, but for typical fluorescence microscopy images, only *Absolute*, *RelativeToMean*, and *RelativeToGeometricMean* are appropriate. These are, respectively, an absolute invariant pixel value or values corresponding to a user-defined number of standard deviations above the arithmetic (μ_a) or geometric (μ_g) means:

$$\begin{aligned} & \textit{RelativeToMean} \\ \textit{Threshold} &= \mu_a + n\sigma_a \end{aligned}$$

$$\begin{aligned} & \textit{RelativeToGeometricMean} \\ \textit{Threshold} &= \mu_g + n\sigma_g \end{aligned}$$

where n is a constant for all images in any one analysis, σ_a is the arithmetic standard deviation, and σ_g is the geometric standard deviation. These statistics are calculated for each X, Y, Z stack by a separate OME module and then passed on to FindSpots by the OME analysis engine. We typically used values of $n = 3$ or $n = 4$ for the images discussed here. The other important parameter to set is the minimum spot volume, which should be large enough that random bright dots (noise) that are too small to be spots are not mistakenly identified as spots of interest.

Background Correction and Other Considerations

When performing quantitative image analysis, it is important to correct for differences in background intensity so that all images are on a level playing field—especially if looking for subtle effects. Where possible, such differences should be minimized by identical treatment of all samples and by optimization of the optical system (13), but it is almost inevitable that differences will remain. Depending on the intended analysis, these differences should be removed prior to running FindSpots or used to adjust the spot statistics generated by the program. Many simple approaches to background correction exist, and as long as all images are treated equally, it often will not matter which method is chosen. For fluorescence images, one such method is to measure manually the average pixel intensity of one or more small regions of each image that should, in principle, be constant between all images analyzed in the experiment (but in practice will vary, which is why it needs to be taken into account) and use the result to correct either the entire image prior to processing, or the reported spot intensities. The location of such regions will depend on the question being studied and the known biology of the system, but may include the cytoplasm or simply the black space between cells on a slide.

In cases where the spots occupy a relatively small total volume relative to the volume of the whole image, as is the case with the kinetochore data presented herein, the average pixel intensity over all pixels in the image will provide a simple but effective measure of the background, and this can be obtained directly from OME. However, as discussed above for thresholding, it is important to crop the image in order to avoid this background estimate being overly skewed by irrelevant parts of the image.

SPECIFICATIONS AND IMPLEMENTATION

We discuss here some of the practical considerations necessary for

successful application of the method to the analysis of microscope images. Detailed instructions for installing OME and using FindSpots are available online at www.openmicroscopy.org/howto.

Image Formats

Any file imported into an OME system can be analyzed using FindSpots. Indeed, this is one of the major advantages of an OME-associated analysis application such as FindSpots: any image that is imported in OME can be processed using FindSpots, and as more importers for proprietary formats are developed for OME, FindSpots or any other segmentation algorithm used in OME will be able to read them. A list of currently supported formats can be found online at www.openmicroscopy.org/getting-started/import.html.

OME/FindSpots is similar to the ImageJ software tool (4,19) in many ways, as the latter is also an open-source program with support for many formats and several segmentation algorithms. What distinguishes OME/FindSpots from ImageJ (and many other commercial imaging applications) is a 5-D model of microscopy images and a mechanism for expressing, storing, and managing structured numerical data.

Image Capture and Preparation

Images containing spot-like objects to be analyzed using FindSpots can be collected on essentially any microscopy platform. Images can consist of single planes or multiple optical sections. They can contain data collected from any number of emission wavelengths using any number of time points in a time series.

Depending on the contents of the image, and especially if a set of images will be analyzed using FindSpots in an automated way, optimally each image should be captured or cropped (in all three axes) so as to contain primarily the region or cell/tissue of interest, with minimal black background or neighboring cells/tissue. This can usually most easily be accomplished using the same software suite originally used to capture the images. Doing this prevents the subsequent automatic thresholding

or background determination (discussed above) from being skewed by unwanted signals from other parts of the image.

Indeed, successful quantitative analysis of microscope images requires careful attention to the initial collection of images at the microscope. Factors that may introduce unwanted variation must be minimized or eliminated. For example, all images from a single experiment that are to be compared directly against each other should always be collected using the same microscope and lens, the same illumination settings (e.g., laser power), and with identical exposure times and binning settings (where this is an option) for any given wavelength channel.

FindSpots Output

Prior to further analysis, it is important to verify that the results of applying FindSpots to a set of images are sensible in order to ensure that the parameters have been set correctly and that the image files used are in fact suitable for analysis. Inspection of the results of a few representative images is normally sufficient for this. Simple checks include whether the total number of spots found is reasonable, and whether their positions match those detectable by eye.

The results of running FindSpots can be viewed through a web browser. The location of spots identified by the software can be verified by viewing an overlay of FindSpots' output with the original image, as shown in Figures 1B and 2B. Each blue circle represents the "centroid" (center of mass

weighted by pixel intensity) of a spot found by FindSpots, either limited to the corresponding optical section being displayed or across all optical sections, according to user settings. Note that in some instances, what appears to be two or more separate spots may be identified by the algorithm as a single spot; this can occur if at any point in the image stack, there are pixels bridging the gap between the two spots. Since the algorithm works in 3-D, these two spots are actually identified as a single spot, so only one centroid would be reported for such a cluster. Depending on the goals of the analysis, this is not necessarily a problem, and indeed this phenomenon can also occur with object finding algorithms other than global thresholding. In some cases, it may be possible to resolve this by raising the threshold slightly.

The raw numerical output can be transferred to a spreadsheet for further mathematical processing, such as averaging, summation, or plotting of the spot data. Indeed, the power of OME stems in large part from its ability to analyze many images quantitatively and automatically. To process the results of such an analysis, specially written macros have been developed to import the results directly into Microsoft Excel, where the spot statistics from many images can be collated into a single file and analyzed. For example, the total spot intensity for each emission channel can be determined for each image and plotted for all images.

Using the Software

FindSpots is part of the standard installation of OME, which can be downloaded at www.openmicroscopy.org. Currently, installation is simplest on a Macintosh running OS X, because a graphical installer is available for this system, but not currently for other

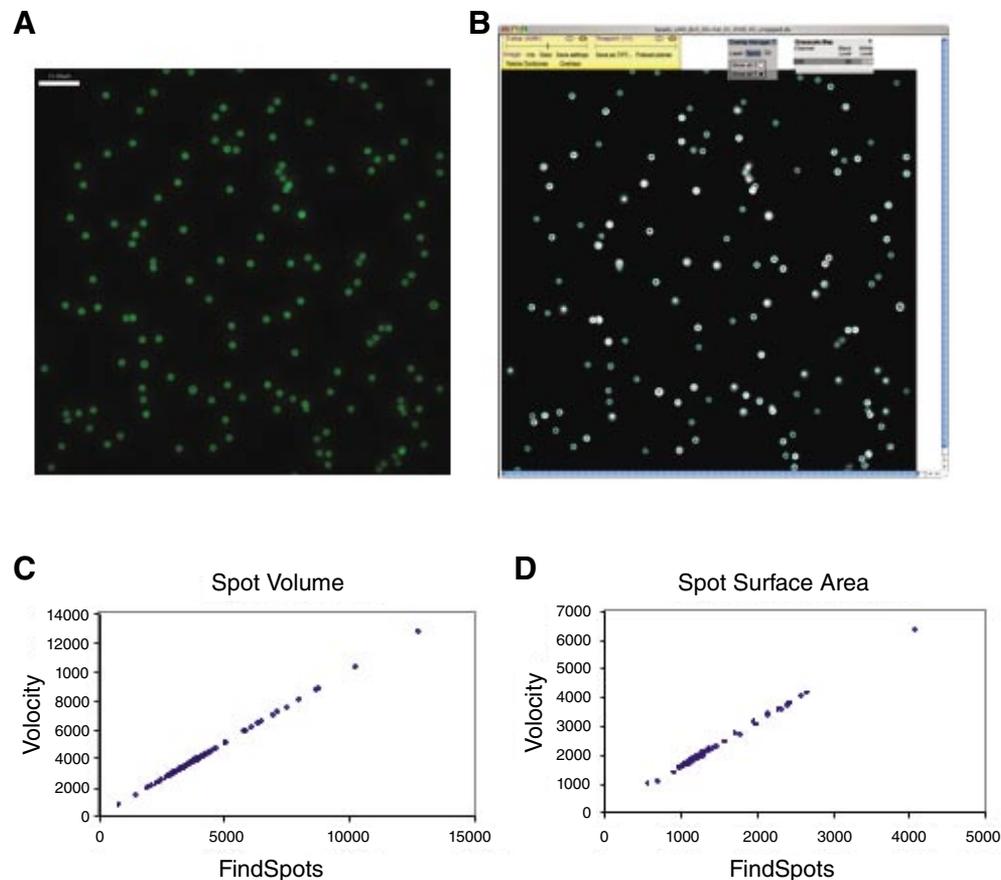


Figure 1. Validation of FindSpots by analysis of fluorescent beads. A set of deconvolved optical sections of fluorescein isothiocyanate (FITC)-conjugated beads (2.5 μm diameter) was analyzed using the object-finding features of FindSpots and Volocity. (A) A maximum-projection view of all the optical sections (generated in Volocity), showing the beads analyzed by both methods. (B) A screen-shot from FindSpots: a single optical section is shown (hence not all beads are equally visible), overlaid with all spots identified in this and all other optical sections (blue circles). Note that the image is shown in gray scale solely for improved contrast of the blue circles. Scattergraphs showing the correlations between (C) the spot volume and (D) the surface area measurements obtained using Volocity and FindSpots. Note the weaker correlation for surface area.

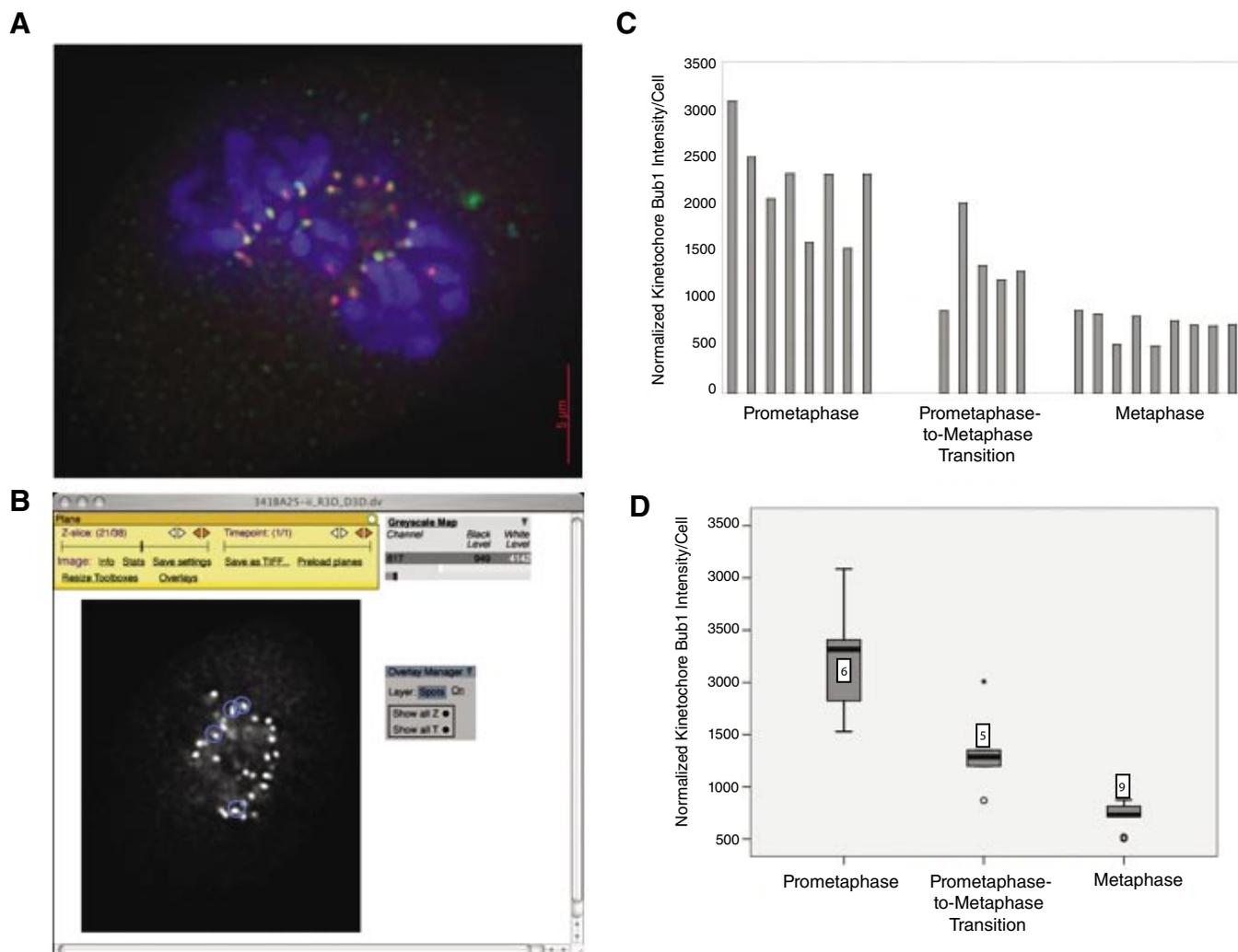


Figure 2. Using FindSpots to establish protein dynamics of Bub1 at kinetochores in fibroblasts. (A) Optical section of a wild-type cell, labeled with the DNA marker 4',6-diamidino-2-phenylindole (DAPI) (blue); CREST, a kinetochore-specific antibody (red); and anti-Bub1 antibody (green). Image generated using softWoRx Explorer. (B) FindSpots output (gray scale view of kinetochore marker channel only) for optical section shown in A. Blue circles indicate spots identified by FindSpots—only those whose centroids are located in the current optical section are displayed. The image data used to generate this section is identical to that in panel A, but the images appear different as each software package uses different standard rules for displaying images: if panel B is simply reflected horizontally and rotated clockwise through 90°, it will match panel A. (C) Histogram summary of 22 cells analyzed via FindSpots. (D) Boxplot (35) of same data, showing more clearly the reduction of Bub1 at kinetochores from prometaphase to metaphase. The plot was generated by SPSS version 13.0; numbers indicate the sample count in each case. The upper and lower boundaries of the box represent, respectively, the lower (LQ) and upper (UQ) quartiles of the data, and the thick line within the box shows the median. The vertical lines (“whiskers”) show the upper and lower ranges of the data, excluding outliers. Mild outliers are defined in SPSS as those points more than $1.5 \times (UQ-LQ)$ from the median and are shown as an “o”, whereas extreme outliers are those more than $3 \times (UQ-LQ)$ from the median, and are represented by “*”.

operating systems. The OME server needs to run on a UNIX®-based operating system (e.g., OS X, Linux, or BSD). However, OME’s user interfaces can be run on any platform. Images to be analyzed are first imported into OME and grouped into one or more data sets before FindSpots can be run on the desired data set. After setting user-defined analysis parameters, the software will automatically analyze each image within the data set in turn. Results can be viewed online in a web

browser or imported into Microsoft Excel for further analysis.

VALIDATION

Validation Using Fluorescent Beads

To validate the FindSpots analysis algorithm, we used an image of fluorescent beads as a simple test sample. We analyzed this using FindSpots and an alternative package,

Velocity version 3.6, and performed a pairwise comparison of the results for each spot identified. Velocity is an established commercial image-analysis software package designed for visualizing 3-D microscope images and for detecting and tracking objects within these images. Although it is a very powerful package, Velocity, unlike FindSpots, is not integrated as part of an image informatics suite and cannot automatically perform such analyses

on a collection of images; rather, the user is required to repeat the analysis manually for each image.

InSpeck™ Green microspheres (Molecular Probes™; Invitrogen, Paisley, UK) with a diameter of 2.5 μm were mounted in Hydromount™ (National Diagnostics, East Riding of Yorkshire, UK) and imaged using a DeltaVision wide-field deconvolution microscopy system (Applied Precision). The deconvolved image used for analysis (Figure 1A) was cropped to reduce distortions at the boundary, yielding an 81-section, 929 × 919 pixel, 16-bit image.

The exact values employed for threshold level and minimum spot volume were not critical in this case, but it was important to use the same value for each method to permit a meaningful software comparison. Hence, an identical absolute threshold value and equivalent minimum object volumes (accounting for different units used) were set manually for each program, which led to the detection of all beads in the image (although groups of beads which were very close together were detected as single spots) as shown in Figure 1B. To compare the results of the two software routines, spot statistics were compared for all spots found. The mean absolute percentage difference was calculated to provide a comparison of how the absolute values reported by each package compared, whereas the Pearson correlation coefficient, *r*, showed the extent of the linear correlation between the statistics generated by the two packages. Both packages found the same number (*n* = 131) of

spots in the image. To ensure that we compared statistics from corresponding beads in each program's output, output tables were sorted by the X, Y, Z coordinates of the spot centroids. Visual comparison of the locations of a random subset of 15 spots showed this to be a generally reliable way of matching up spots between the two packages—manual correction was needed for just two spots.

The comparison between FindSpots and Volocity, summarized in Table 1 and Figure 1, showed that the two packages gave results that were highly correlated (*r* > 0.999) for all seven parameters common to both packages. However, spot surface areas showed a distinctly lower, albeit still very high, correlation. A more striking difference was seen when comparing absolute values: both packages yielded essentially identical values for all spot parameters, except for surface areas, which differed by an average of 57%. This turned out to reflect a difference in the methods used to calculate the surface areas of objects. Nonetheless, these results show that the spot statistics produced by the two software packages were highly correlated, confirming that FindSpots can be used to make internally consistent comparisons between images.

In summary, this analysis confirmed that the spot-finding algorithm used by FindSpots works as expected. It also highlights the need to be careful if analyzing data from a single experiment with more than one software package—because of software-specific differences in certain calculations, it is recommended to use the same package

for all quantitative measurements to be compared against one another where possible.

Biological Images: Application and Validation of FindSpots

A particular strength of using FindSpots in OME is that it can automatically process large numbers of images, the results of which can be analyzed en masse using Microsoft Excel. As part of a study investigating the regulation of mitosis, we measured the amount of two mitotic checkpoint proteins, Bub1 and BubR1, at kinetochores during different stages of mitosis using images of mitotic cells. Comparing the results to those expected from the literature (measured by other means) provided a method to validate FindSpots on real, complex biological images, and also illustrated how OME/FindSpots can be applied to answer biological questions.

Many of the components of the spindle checkpoint machinery, including Bub1 and BubR1, associate dynamically with kinetochores throughout mitosis (20,21). We used FindSpots to quantify the amount of kinetochore-associated spindle checkpoint proteins, Bub1 and BubR1, in cells at different stages of mitosis. The mitotic spindle checkpoint functions to ensure proper chromosome attachment to mitotic spindles before chromosomes segregate into daughter cells (22). Defects in this checkpoint result in mitotic defects, including chromo-

Table 1. Comparison of FindSpots with Another Object-Detection Package

Spot Parameter Spot Statistic	FindSpots Versus Volocity (n = 131)	
	Pearson Correlation Coefficient	Mean % Absolute Difference ^a
Total intensity	1.0000	0.0738 ± 0.045
Mean intensity	1.0000	0.0358 ± 0.017
Volume	1.0000	0.110 ± 0.060
Surface area	0.99908	57.6 ± 2.7
Centroid X	1.00000	0.0171 ± 0.027
Centroid Y	1.00000	0.051 ± 0.23
Centroid Z	0.99987	0.229 ± 0.29

The image in Figure 1A was analyzed using equivalent parameters in FindSpots and Volocity.

^aDefined for each type of spot statistic, as the mean, over all spots, of the absolute percentage difference (APD) between software packages, where APD = |(FindSpots value - Volocity value)/FindSpots value| × 100; ± the standard deviation of the APD. Shown to 3 significant figures ± 2 significant figures.

somal instability, an important characteristic of tumor cells (23).

Unsynchronized mouse fibroblasts were fixed and double-stained with antibodies against the kinetochore marker, CREST, and antibodies against either Bub1 or BubR1 (we present data only on Bub1 here). Cells were also stained with 4',6-diamidino-2-phenylindole (DAPI) to visualize chromosomes. 3-D images of mitotic cells were cropped in the X/Y plane so that each image contained only one cell and minimum nearby space. Images were also cropped in the Z plane so that only regions of the cells containing kinetochores were included. Kinetochore spots were identified within each 3-D image stack using FindSpots by thresholding on the CREST channel. The total fluorescence intensity at these spots was then automatically computed for the other wavelength(s) corresponding to Bub1 and/or BubR1 staining. This allowed kinetochore-associated signals to be accurately defined even in images where the Bub1 or BubR1 signal-to-noise ratio was not sufficiently high to easily define kinetochores. For each image, the measured intensity value was normalized by dividing it by the total spot volume in that image (i.e., the total kinetochore volume). The data were then sorted according to the mitotic stage of each individual cell, namely prophase, prometaphase, prometaphase-to-metaphase transition (when most but not all chromosomes were aligned on the metaphase plate), metaphase, and anaphase.

Figure 2 shows an example of the images used (Figure 2A), the results of applying FindSpots to it (Figure 2B), and two alternate ways to represent the data from all images analyzed in this experiment (Figure 2, C and D).

The relative amount of kinetochore-associated Bub1 that we measured throughout mitosis using this method agrees well with previous studies of Bub1 dynamics at kinetochores during mitosis (20,24), which independently validates the use of our method to quantify proteins in biological images.

Applying FindSpots to Study the Effect of a Tumor-Suppressor Protein

We employed a similar approach (unpublished data) to gain insight into the mechanism of spindle checkpoint defects that we found in cells lacking the tumor suppressor, adenomatous polyposis coli (APC) (25,26). Colorectal cancer, one of the most common cancers in the Western world, is associated with mutations in APC; lack of and mutations in the APC protein can interfere with normal mitosis (27–29). APC binds to the kinetochore-associated mitotic checkpoint proteins, Bub1 and BubR1 (30).

We used OME/FindSpots to measure levels of kinetochore-associated Bub1 and BubR1 in APC-deficient mouse fibroblasts and compared these with levels in wild-type control cells, using the approach described above. To compare cells at uniform mitotic stages, we arrested them with monastrol. This Eg5 kinesin inhibitor prevents centrosome separation and arrests cells with their chromosomes arranged in a rosette around centrosomes, leaving their checkpoint active (31).

Results from wild-type cells and APC-deficient cells were compared using standard statistical tests. The data was first analyzed using an F-test to determine if there was a significant difference (at the 5% level) between the variances of the two groups. For both Bub1 and BubR1, this difference was not significant here, and so unpaired 2-tailed Student's t-tests assuming equal variances were performed to compare the means.

Analyzing the relative amount of Bub1 (Figure 3) and BubR1 at kinetochores using this method revealed a statistically very significant decrease in kinetochore-associated Bub1 in APC mutant cells ($P < 0.01$), but for BubR1 the difference was not significant at the 5% level (data not shown). This result further illustrates how OME and FindSpots can be used to extract statistically significant biological information from complex image data.

DISCUSSION

This study arose out of the need to answer a biological question through the quantitative analysis of a large number of images obtained from fluorescence microscopy. Alternative

methods investigated were either too complex to be of use or would have required a manual analysis of each image—given that potentially hundreds of images were to be analyzed, an automated method was desirable. FindSpots/OME was identified as an appropriate tool, was further developed to produce the tool described here, and is well suited to meet this task. One of the attractive features that led to its selection was that even though the suite of tools then available in OME was not fully adequate to complete this task at first, this software is open-source and thus could be readily adapted. OME features the ability to import various proprietary file formats and the easy automatic processing of large sets of images, without the need for user-written scripts. In addition, it allows the integrated management of large amounts of quantitative results so that the effects of parameter variation can be easily evaluated, both visually and quantitatively. This approach to image analysis is powerful due to the integration of analysis with a data management system that automatically tracks all analysis results, making it straightforward to optimize parameters and manage results. Although the data models used by OME to describe the results from FindSpots did not need alteration for this study and will not be discussed further here, they formed the basis by which a collection of disparate software (e.g., API's DeltaVision/softWoRx microscopy system, OME, FindSpots, and Excel) was used together to accomplish a complex image-analysis task.

Limitations and Solutions

The integration of FindSpots with OME is an important addition to the arsenal of tools available to those seeking to obtain quantitative data from microscope images. At the same time, it is important to be aware of potential pitfalls and limitations in its use.

For analysis of cellular images as used here, the need to crop images manually to yield more meaningful automatic thresholding or background correction, as discussed earlier, can limit the throughput of analyses. However, depending on the question

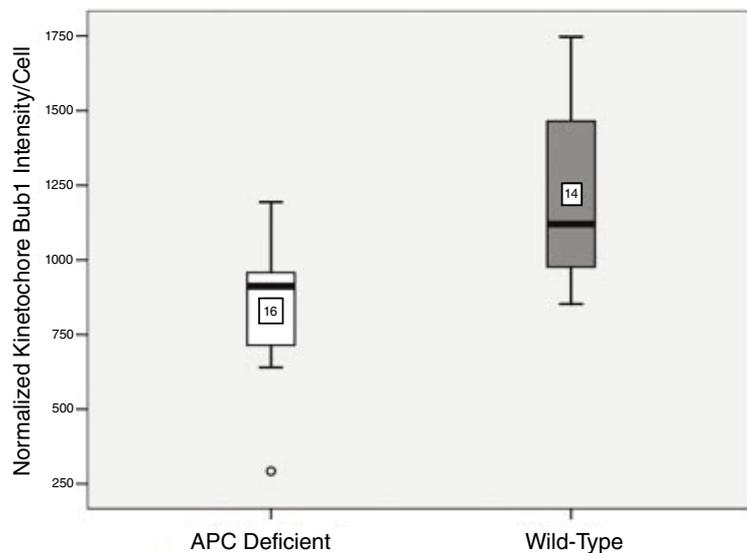


Figure 3. Application of FindSpots to the analysis of mitotically arrested cells. The kinetochore-associated serine-threonine kinase, Bub1, was studied in mouse fibroblast cells treated with monastrol, which arrests cells in mitosis with unseparated centrosomes. Cells with wild-type adenomatous polyposis coli (APC) levels ($n = 14$) were compared with those with reduced APC levels in mutant cells ($n = 15$, excluding one outlier). The box plot clearly shows the reduction in kinetochore-associated Bub1 intensity in APC-deficient cells. The significance of the symbols is as in Figure 2D.

being asked, and the distribution of the cells/tissue under study, this step may not be necessary or can be rendered unnecessary by modifying the original imaging protocol so as to minimize the inclusion of unwanted regions in the images.

As with any image-analysis package used to find and report parameters about the objects within an image, certain parameters are more robust than others. In order of decreasing robustness (increasing difficulty of obtaining accurate numbers), these parameters are count, position, shape, and total intensity. For example, slight variations in threshold can give very different results for an object's total signal without necessarily affecting its position. In practice, however, in images where these objects are much brighter than the surrounding regions [as is often the case in labeled or green fluorescent protein (GFP)-expressing cells or tissues], small changes in threshold will in fact not significantly affect the reported volume or total intensity for such spots.

As discussed earlier, global thresholding, as used in FindSpots, is not suitable for all image types; in cases with sharp variations in contrast across the image, a local thresholding algorithm, or indeed segmentation

by methods other than thresholding, may be needed. This may help, for example, to minimize the merging of neighboring spots into single objects, without requiring such a high threshold that much of each spot is lost. Global thresholding is a robust method that can be applied in many cases of biological imaging, as evidenced by its use in many image-analysis packages, including Volocity. Nonetheless, as mentioned above, no segmentation method is perfect, and for some applications, a customized object-detection algorithm may need to be developed. The power of OME is that such algorithms can be developed in a choice of common programming environments and then be integrated seamlessly into OME, where they can be used to analyze images automatically within such data sets, as demonstrated here for the FindSpots algorithm. If the algorithms use the same data models to describe their inputs and outputs, they can seamlessly interact, allowing one to swap out easily the current FindSpots or TrackSpots algorithms for entirely different algorithms that perform similar functions.

Users should also note that no matter how good a segmentation method is, it is not a substitute for common sense and a visual check of a subset

of the results to ensure that objects are detected broadly as expected.

Conclusions

The use of FindSpots with OME in the biological examples given here made it far quicker to obtain quantitative information than would have been feasible by manual analysis of individual images. Indeed, a similar approach can in principle be applied to much larger data sets, which should prove useful for other applications in basic cell biology and in drug screening, as large-scale imaging studies are key elements of drug discovery (32–34).

Although the present work focuses on microscope images, OME and its associated software are not limited to just those images, and they may also find use for the storage and management of images from other devices that have the potential to generate large quantities of digital image data, such as X-ray crystallography equipment, infrared Western blot scanning equipment (such as the Odyssey® Imager; LI-COR Biotechnology, Lincoln, Nebraska, USA), or medical devices such as endoscopes. OME already has support for DICOM—a format commonly used in medical radiology.

In summary, we have shown that FindSpots, when combined with OME, is a powerful tool for image analysis. As well as validating the FindSpots algorithm, the results discussed here represent the first detailed account of segmentation in OME and the first use of FindSpots/OME to identify and quantify mitotic kinetochores. They also demonstrate more generally its potential for quantitative image analysis of large data sets. Using a choice of common programming environments, such as MATLAB® (The MathWorks, Natick, MA, USA), Perl, Java™, or C, users can generate customized analysis algorithms and apply these to entire collections of images in a similar fashion to FindSpots. This illustrates how OME allows a plug-in approach and common data models representing image information to implement a workflow for image informatics and analysis.

ACKNOWLEDGMENTS

I.S.N. and I.G.G. both have senior authorship of this paper. We thank Dr. Steve Ogg for initial help in analyzing kinetochore images, and all members of the N athke laboratory and the Dundee OME team for their help. This work is supported by grants from Cancer Research UK, The Wellcome Trust, and the National Institutes of Health.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

NOTE ADDED IN PROOF

While this paper was in the proof stage, it was announced that a new version of Volocity software will soon be released which will, unlike the current version discussed here, be able to perform tasks on multiple datasets in an automated way.

REFERENCES

- Harding, A. 2006. Seeing is believing. *Scientist* 20:46-51.
- Biel, S., R. Wepf, and P. Schwarb. 2005. Multimodal imaging in life science requires intelligent digital data management. *G.I.T. Imaging and Microscopy* 4/05:45-47.
- Hogan, H. 2005. Where's that picture? *Biophotonics International July*:32-36.
- Eliceiri, K.W. and C. Rueden. 2005. Tools for visualizing multidimensional images from living specimens. *Photochem. Photobiol.* 81:1116-1122.
- Swedlow, J.R., I. Goldberg, E. Brauner, and P.K. Sorger. 2003. Informatics and quantitative analysis in biological imaging. *Science* 300:100-102.
- Goldberg, I.G., C. Allan, J.M. Burel, D. Creager, A. Falconi, H. Hochheiser, J. Johnston, J. Mellen, et al. 2005. The Open microscopy environment (OME) data model and XML file: open tools for informatics and quantitative analysis in biological imaging. *Genome Biol.* 6:R47.
- Andrews, P.D., I.S. Harper, and J.R. Swedlow. 2002. To 5-D and beyond: quantitative fluorescence microscopy in the postgenomic era. *Traffic* 3:29-36.
- Platani, M., I. Goldberg, A.I. Lamond, and J.R. Swedlow. 2002. Cajal body dynamics and association with chromatin are ATP-dependent. *Nat. Cell Biol.* 4:502-508.
- Platani, M., I. Goldberg, J.R. Swedlow, and A.I. Lamond. 2000. In vivo analysis of Cajal body movement, separation, and joining in live human cells. *J. Cell Biol.* 151:1561-1574.
- Jesch, S.A., A.J. Mehta, M. Velliste, R.F. Murphy, and A.D. Linstedt. 2001. Mitotic Golgi is in a dynamic equilibrium between clustered and free vesicles independent of the ER. *Traffic* 2:873-884.
- Castleman, K.R. 1996. *Digital Image Processing*. Prentice Hall, Upper Saddle River, NJ.
- Young, I.T., J.J. Gerbrands, and L.J. van Vliet. 1998. Fundamentals of image processing. In V.K. Madisetti and D.B. Williams (Eds.), *The Digital Signal Processing Handbook*. CRC Press, Boca Raton.
- Russ, J.C. 2002. *The Image Processing Handbook*, 4th ed. CRC Press, Boca Raton.
- Zhou, J., D. Panda, J.W. Landen, L. Wilson, and H.C. Joshi. 2002. Minor alteration of microtubule dynamics causes loss of tension across kinetochore pairs and activates the spindle checkpoint. *J. Biol. Chem.* 277:17200-17208.
- Seul, M., L. O'Gorman, and M.J. Sammon. 2000. *Practical Algorithms for Image Analysis with CD-ROM: Description, Examples and Code*. Cambridge University Press, Cambridge, UK.
- Petrou, M. and P. Bosdogianni. 1999. *Image Processing: The Fundamentals*. John Wiley & Sons, Chichester, England.
- Sage, D., F.R. Neumann, F. Hediger, S.M. Gasser, and M. Unser. 2005. Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics. *IEEE Trans Image Process.* 14:1372-1382.
- Bacher, C.P., M. Reichenzeller, C. Athale, H. Herrmann, and R. Eils. 2004. 4-D single particle tracking of synthetic and proteinaceous microspheres reveals preferential movement of nuclear particles along chromatin-poor tracks. *BMC Cell Biol.* 5:45.
- Girish, V. and A. Vijayalakshmi. 2004. Affordable image analysis using NIH Image/ImageJ. *Indian J. Cancer* 41:47.
- Howell, B.J., B. Moree, E.M. Farrar, S. Stewart, G. Fang, and E.D. Salmon. 2004. Spindle checkpoint protein dynamics at kinetochores in living cells. *Curr. Biol.* 14:953-964.
- Jablonski, S.A., G.K. Chan, C.A. Cooke, W.C. Earnshaw, and T.J. Yen. 1998. The hBUB1 and hBUBR1 kinases sequentially assemble onto kinetochores during prophase with hBUBR1 concentrating at the kinetochore plates in mitosis. *Chromosoma* 107:386-396.
- Zhou, J., J. Yao, and H.C. Joshi. 2002. Attachment and tension in the spindle assembly checkpoint. *J. Cell Sci.* 115:3547-3555.
- Masuda, A. and T. Takahashi. 2002. Chromosome instability in human lung cancers: possible underlying mechanisms and potential consequences in the pathogenesis. *Oncogene* 21:6884-6897.
- Shah, J.V., E. Botvinick, Z. Bonday, F. Furnari, M. Berns, and D.W. Cleveland. 2004. Dynamics of centromere and kinetochore proteins; implications for checkpoint signaling and silencing. *Curr. Biol.* 14:942-952.
- Fearnhead, N.S., M.P. Britton, and W.F. Bodmer. 2001. The ABC of APC. *Hum. Mol. Genet.* 10:721-733.
- Nathke, I.S. 2004. The adenomatous polyposis coli protein: the Achilles heel of the gut epithelium. *Annu. Rev. Cell Dev. Biol.* 20:337-366.
- Green, R.A. and K.B. Kaplan. 2003. Chromosome instability in colorectal tumor cells is associated with defects in microtubule plus-end attachments caused by a dominant mutation in APC. *J. Cell Biol.* 163:949-961.
- Dikovskaya, D., I.P. Newton, and I.S. Nathke. 2004. The adenomatous polyposis coli protein is required for the formation of robust spindles formed in CSF *Xenopus* extracts. *Mol. Biol. Cell* 15:2978-2991.
- Tighe, A., V.L. Johnson, and S.S. Taylor. 2004. Truncating APC mutations have dominant effects on proliferation, spindle checkpoint control, survival and chromosome stability. *J. Cell Sci.* 117:6339-6353.
- Kaplan, K.B., A.A. Burds, J.R. Swedlow, S.S. Bekir, P.K. Sorger, and I.S. Nathke. 2001. A role for the adenomatous polyposis coli protein in chromosome segregation. *Nat. Cell Biol.* 3:429-432.
- Kapoor, T.M., T.U. Mayer, M.L. Coughlin, and T.J. Mitchison. 2000. Probing spindle assembly mechanisms with monastrol, a small molecule inhibitor of the mitotic kinesin, Eg5. *J. Cell Biol.* 150:975-988.
- Milligan, G. 2003. High-content assays for ligand regulation of G-protein-coupled receptors. *Drug Discov. Today* 8:579-585.
- Mitchison, T.J. 2005. Small-molecule screening and profiling by using automated microscopy. *ChemBioChem* 6:33-39.
- Monk, A.J. 2005. Faster, surer prediction. Harnessing automated microscopy for drug discovery. *Biochemist October*:25-28.
- Tukey, J.W. 1977. *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.

Received 15 March 2006; accepted 1 June 2006.

Address correspondence to Inke S. N athke, Division of Cell & Developmental Biology, School of Life Sciences, Wellcome Trust Biocentre, University of Dundee, Dundee DD1 5EH, Scotland, UK e-mail: inke@lifesci.dundee.ac.uk and Ilya G. Goldberg, Image Informatics and Computational Biology Unit, Laboratory of Genetics, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA. e-mail: igg@nih.gov

To purchase reprints of this article, contact: Reprints@BioTechniques.com