

Automatic channel unmixing for high-throughput quantitative analysis of fluorescence images

Cris L. Luengo Hendriks¹, Soile V. E. Keränen², Mark D. Biggin² and David W. Knowles^{1*}

1) Life Science Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mailstop 84R0171, Berkeley, CA 94720, USA.

2) Genomics Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mailstop 84R0171, Berkeley, CA 94720, USA.

** Corresponding author: dwknowles@lbl.gov*

Abstract: Laser-scanning microscopy allows rapid acquisition of multi-channel data, paving the way for high-throughput, high-content analysis of large numbers of images. An inherent problem of using multiple fluorescent dyes is overlapping emission spectra, which results in channel cross-talk and reduces the ability to extract quantitative measurements. Traditional unmixing methods rely on measuring channel cross-talk and using fixed acquisition parameters, but these requirements are not suited to high-throughput processing. Here we present a simple automatic method to correct for channel cross-talk in multi-channel images using image data only. The method is independent of the acquisition parameters but requires some spatial separation between different dyes in the image. We evaluate the method by comparing the cross-talk levels it estimates to those measured directly from a standard fluorescent slide. The method is then applied to a high-throughput analysis pipeline that measures nuclear volumes and relative expression of gene products from three-dimensional, multi-channel fluorescence images of whole *Drosophila* embryos. Analysis of images before unmixing revealed an aberrant spatial correlation between measured nuclear volumes and the gene expression pattern in the shorter wavelength channel. Applying the unmixing algorithm before performing these analyses removed this correlation.

© 2007 Optical Society of America

OCIS codes: (180.2520) Fluorescence microscopy; (100.2000) Digital image processing; (100.2960) Image analysis.

References and links

1. K. Carlsson, and K. Mossberg, "Reduction of cross-talk between fluorescent labels in scanning laser microscopy," *J. Microsc.* **167**, 23–37 (1992).
2. J. Vassy, J. P. Rigaut, A. M. Hill, and J. Fouquier, "Analysis by confocal scanning laser microscopy imaging of the spatial distribution of intermediate filaments in foetal and adult rat liver cells," *J. Microsc.* **157**, 91–104 (1990).
3. K. Mossberg, and M. Ericsson, "Detection of doubly stained fluorescent specimens using confocal microscopy," *J. Microsc.* **158**, 215–224 (1990).
4. M. E. Dickinson, G. Bearman, S. Tille, R. Lansford, and S. E. Fraser, "Multi-spectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy," *BioTechniques* **31**, 1272–1278 (2001).
5. H. Shirakawa, and S. Miyazaki, "Blind spectral decomposition of single-cell fluorescence by parallel factor analysis," *Biophys. J.* **86**, 17391752 (2004).

6. J. R. Mansfield, K. W. Gossage, C. C. Hoyt, and R. M. Levenson, "Autofluorescence removal, multiplexing, and automated analysis methods for in-vivo fluorescence imaging," *J. Biomed. Opt.* **10**, 41207 (2005).
 7. H. Ahammer, T. T. J. DeVaney, M. Hartbauer, and H. A. Tritthart, "Cross-talk reduction in confocal images of dual fluorescence labelled cell spheroids," *Micron* **30**, 309–317 (1999).
 8. D. Chorvat Jr, J. Kirchnerova, M. Cagalinec, J. Smolka, A. Mateasik, and A. Chorvatova, "Spectral unmixing of flavin autofluorescence components in cardiac myocytes," *Biophys. J.* **89**, L55–57 (2005).
 9. C. L. Luengo Hendriks, S. V. E. Keränen, C. C. Fowlkes, L. Simirenko, G. H. Weber, A. H. DePace, C. Henriquez, D. W. Kaszuba, B. Hamann, M. B. Eisen, J. Malik, D. Sudar, M. D. Biggin, and D. W. Knowles, "3D morphology and gene expression in the *Drosophila* blastoderm at cellular resolution I: data acquisition pipeline," *Genome Biology* **7**, R123 (2006).
 10. S. V. E. Keränen, C. C. Fowlkes, C. L. Luengo Hendriks, D. Sudar, D. W. Knowles, J. Malik, and M. D. Biggin, "3D morphology and gene expression in the *Drosophila* blastoderm at cellular resolution II: dynamics," *Genome Biology* **7**, R124 (2006).
 11. K. R. Castleman, *Digital image processing* (Prentice Hall, Englewood Cliffs, New Jersey, 1996).
 12. S. V. Costes, D. Daelemans, E. H. Cho, Z. Dobbin, G. Pavlakis, and S. Lockett, "Automatic and quantitative measurement of protein-protein colocalization in live cells," *Biophys. J.* **86**, 3993–4003 (2004).
 13. D. Demandolx, and J. Davoust, "Multicolor analysis and local image correlation in confocal microscopy," *J. Microsc.* **185**, 21–36 (1997).
 14. E. M. M. Manders, F. J. Verbeek, and J. A. Aten, "Measurement of co-localization of objects in dual-colour confocal images," *J. Microsc.* **169**, 375–382 (1993).
-

1. Introduction

The advent of fast laser-scanning fluorescence microscopy allows large, three-dimensional images to be acquired in rapid succession. These data sets are providing unparalleled information about spatiotemporal macromolecular dynamics within organelles, cells, tissues and animals. They are also leading to the creation of multidisciplinary efforts in data management, visualization and quantitative image analysis so that biologically relevant information can be extracted and analyzed efficiently.

Many applications require multi-channel imaging to allow data for several different fluorescent dyes to be compared in the same context. Despite its wide application, multi-channel imaging using traditional organic dyes suffers from the inherent problem of overlapping emission spectra, leading to light from more than one dye being collected by each acquisition channel when the dyes are simultaneously excited (Fig. 1). This effect is called *cross-talk*. To computationally reduce cross-talk, various *channel unmixing* methods have been reported. The traditional unmixing scheme [1, 2] relies on the cross-talk levels being measured [1, 3] for a particular set of acquisition parameters, which are then used for all subsequent imaging. More recently, spectral imaging [4, 5, 6], which records the full emission spectrum per pixel, has been used to remove cross-talk. Other unmixing techniques have also been proposed which require knowledge of the object shape [7] or the use of principal component analysis [8]. However, neither the traditional nor these other techniques are suited to high-throughput imaging. For the traditional method, the cross-talk needs to be measured anew when the sample or acquisition parameters are changed. For spectral imaging, measuring the emission spectra substantially increases the acquisition time and the image size. Methods requiring a priori knowledge of the object shape are not robust, and principal component analysis has been shown to be unsuitable for channel unmixing [5].

Here we report an automated, fast method of channel unmixing that removes cross-talk from multichannel images using only the image data. The method does not require prior measurements of cross-talk levels or emission spectra and is independent of the relative gains between acquisition channels. It does require some spatial separation between different dyes in the image, but this separation is present in most applications. Our method also assumes that the cross-talk is unidirectional, where emission from the shorter wavelength dye is recorded in the longer wavelength channel. This, however, is a reasonable assumption if the emission peaks of the dyes are far enough apart, because emission spectra of traditional organic dyes are usually asymmet-

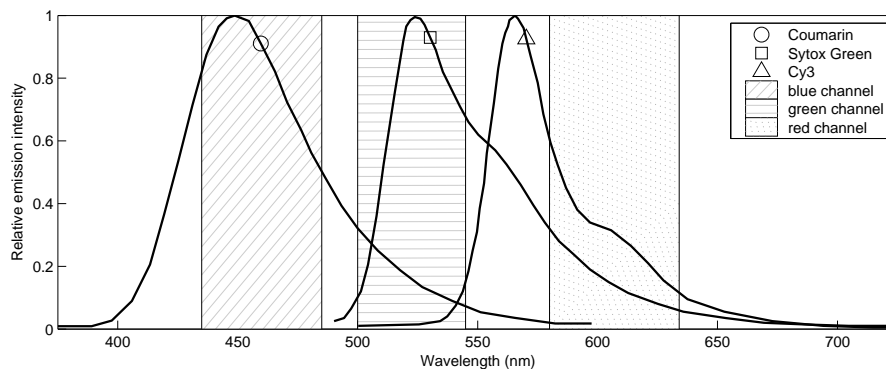


Fig. 1. Overlapping emission spectra of three fluorescent dyes. The shaded areas indicate the wavelength intervals that are acquired in each channel. Note how the channel recording the Sytox Green signal also records the tail of the Coumarin signal, but the Sytox signal is minimal within the Coumarin acquisition window. In the same way, the Sytox signal bleeds through to the Cy3 channel, but not the other way around.

ric, having long red-shifted tails (Fig. 1). The method also performs well in the presence of autofluorescence, even though it does not remove autofluorescence from the image.

We first demonstrate the utility of our method using images collected from a standard fluorescent microscopy slide. Single photon excitation was used, which allowed a comparison of our method with the traditional method. We then apply our method to a high-throughput analysis pipeline that quantifies gene expression and morphology at cellular resolution from images of whole *Drosophila* blastoderm embryos [9, 10]. The pipeline consists of acquisition of three-dimensional, multi-channel fluorescence images of whole embryos using two-photon excitation; automated segmentation of nuclei within the embryo and quantification of gene product in and around each nucleus. Using these images, we show that our automated channel unmixing method removed an aberrant spatial correlation between measured nuclear volumes and the fluorescence intensity of the expression pattern in the shorter wavelength channel.

2. Results

2.1. Computed versus measured cross-talk values

To evaluate our automated unmixing algorithm, we imaged one field of view from a commercial test slide and compared the results of our algorithm to that obtained with the traditional cross-talk measurement method. The test slide was chosen because it allowed us to demonstrate the ability of our method in cases where the different dyes have considerable spatial overlap. The field of view was imaged with three different acquisition configurations to record the *measured green image* (Fig. 2(a)), the *measured red image* (Fig. 2(b)), the *measured green-to-red cross-talk* (Fig. 2(c)), and the *measured pure red image* (Fig. 2(d)). We then used our algorithm to derive the *estimated green-to-red cross-talk* (Fig. 2(e)), and the *estimated pure (unmixed) red image* (Fig. 2(f)). Since the red dye excitation had no influence on the green channel (data not shown), we used the measured green image as if it were the *measured pure green image*.

These two estimates were computed from the joint histogram of the measured red versus measured green images (Fig. 3(a)), using the theory described in the Methods. This skewed histogram shows the data along the green channel axis to be shifted towards the red axis which indicated a significant amount of cross-talk from the green channel into the red, but not from the red into the green. To determine the cross-talk level, the algorithm calculates the slope of

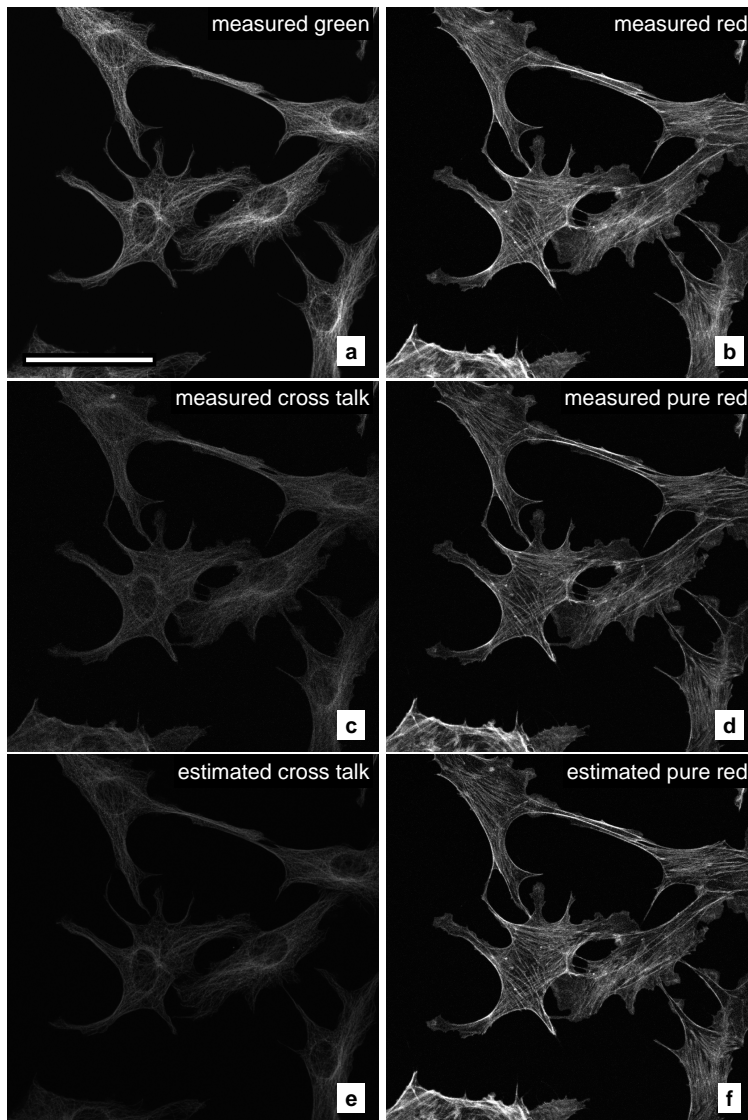


Fig. 2. Comparison of the automated unmixing method with the measured cross-talk using a standard test slide of bovine pulmonary artery endothelial cells. (a) The *measured green image* (F-actin, BODIPY) and (b) the *measured red image* (microtubules, Texas Red) were recorded using simultaneous excitation at 488 nm and 543 nm. The gain and offset of each channel were independently set to fill the 12 bit dynamic range of the images. (c) The image of the *measured cross-talk* from the green to the red channel was then recorded in the red channel using only green dye excitation at 488 nm, and using the same gain and offset as that for the measured red image. (d) The *measured pure red image* was recorded in the red channel using only red dye excitation at 543 nm, again using the same detector settings. (e) The *estimated cross-talk image* and (f) the *estimated pure red image* calculated from the measured green and measured red images shown in (a) and (b). The bar in (a) is 50 μm . The same, small gamma change has been performed on all images to enhance the contrast in the dark areas and thus make the cross-talk better visible.

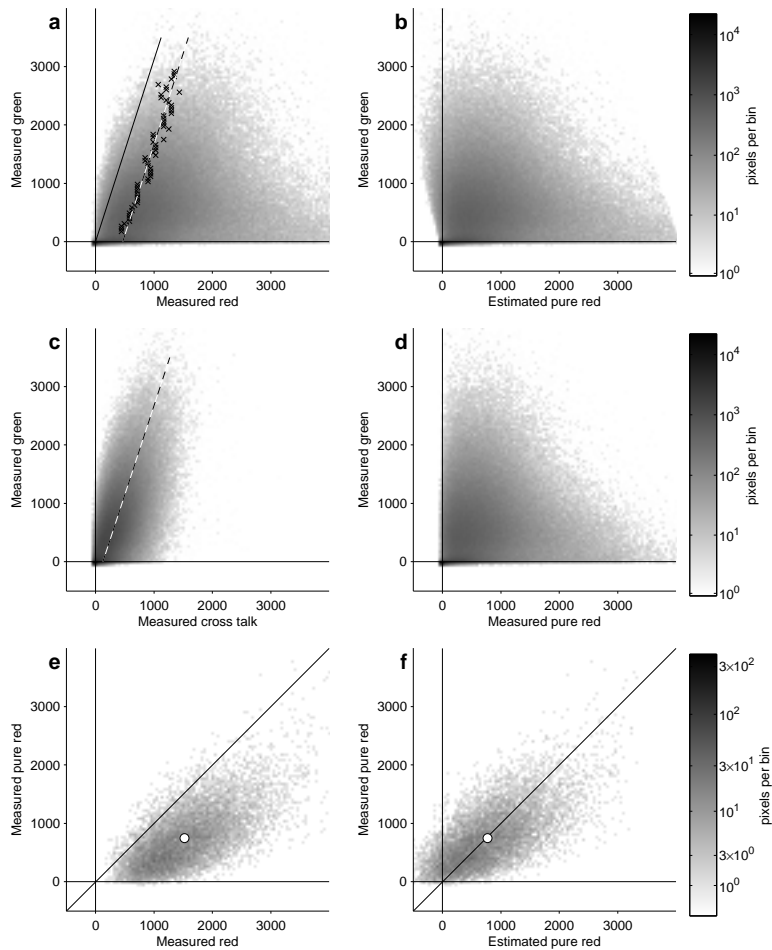


Fig. 3. Joint histograms are used to estimate cross-talk levels. (a) Red-green joint histogram before and (b) after unmixing with the proposed algorithm. The crosses are the points detected by the algorithm on each horizontal line through the histogram. The dashed line is the fit through these points. The offset along the x -axis (red channel) of this fitted line was a result of the autofluorescence, which had a different strength in the two channels, and was ignored. The continuous line, going through the origin, is of the same slope as this fitted line, and corresponds to the vertical axis in the corrected histogram (b). (c) Joint histogram of the measured green image versus the green-to-red cross-talk. Notice that the red channel axis offset is less than in (a). This is because the 488 nm laser excites less autofluorescence than the 543 nm laser. (d) Red-green joint histogram using the pure red image, which shows the expected shape for the histogram plotted in (b). (e) Joint histogram of the measured red image versus the measured pure red image, using only pixels with green intensity of 2000 or more. The cross-talk shifts the plot off the diagonal towards the right. (f) Histogram as in (e) but using the estimated pure (unmixed) red image rather than the measured red image. The algorithm was able to shift the plot back towards the diagonal. The white dots in (e) and (f) are the center of mass of the pixel data.

the best fit line (Fig. 3(a), dashed line) through maxima in the histogram along lines of constant green channel intensity (Fig. 3(a), crosses). The slope of this line (0.32 ± 0.03) gives the fraction of the green channel that had bled through to the red channel. The x -axis (red channel) offset of this line was a result of the autofluorescence and was ignored. The pure red image (Fig. 2(e)) and the estimated cross-talk image (Fig. 2(f)) were then simply calculated from this value. The joint histogram of the measured green channel versus the estimated pure red channel (Fig. 3(b)) shows the skew has been corrected when cross-talk has been removed.

Next we determined the cross-talk level using the traditional unmixing method, by calculating the linear least squares slope of the histogram of the measured green image versus the measured cross-talk image (Fig. 3(c)). The result (0.328 ± 0.001) was consistent with the cross-talk level estimated by our algorithm. This was confirmed by the joint histogram of the measured green image versus the measured pure red image (Fig. 3(d)). The similarity of this histogram to that of Fig. 3(b), indicates our automated method is as effective as the traditional unmixing method in reducing cross-talk. Finally we plotted the joint histograms of the measured pure red image versus the measured pure red image (Fig. 3(e)) and the measured pure red image versus the estimated (unmixed) pure red image (Fig. 3(f)). For these histograms only pixels with a large green channel component (> 2000) are shown, since these are most affected by cross-talk. Notice that the considerable offset from the diagonal in Fig. 3(e), due to cross-talk, was fully corrected once the cross-talk has been removed (Fig. 3(f)). Dividing the horizontal offset of the dot to the diagonal in Fig. 3(f) (780 ± 9) into the average green intensity for these pixels (2423 ± 8) yielded another estimate of the cross-talk level (0.323 ± 0.004) which was consistent with the values computed by our method and the traditional approach.

2.2. Automated channel unmixing removes aberrant correlation between measured nuclear volumes and gene expression intensity

To demonstrate the utility of the unmixing algorithm, we present results of its application to a high-throughput image analysis project (<http://BDTNP.lbl.gov>). Figure 4(a) shows an optical section through the middle of a *Drosophila* embryo at stage 5. Total DNA was labeled with Sytox Green (green), the mRNA product for gene *fushi tarazu* (*ftz*) was labeled with Coumarin (blue), and the mRNA product for *even-skipped* (*eve*) was labeled with Cy3 (red). Figure 4(b)-(d) show the blue, green and red channels, respectively, of the portion of the embryo in Fig. 4(a) indicated by the white box. Unmixing the three channels resulted in the unmixed green and red images (Fig. 4(e),(f)). The effect of the unmixing is readily visible, and the degree of cross-talk is apparent from the skew of the corresponding joint histograms (Fig. 4(g),(h)), each of which show cross-talk from the shorter wavelength channel to the longer wavelength channel. Figure 4(i),(j) show the joint histograms after unmixing.

Analysis performed without channel unmixing revealed a spatial correlation between the nuclear volumes, measured from the green channel, and the expression pattern imaged in the blue channel. The correlation was clearly aberrant and likely caused by cross-talk. To test this hypothesis, two cohorts of images were selected having either *ftz* or *eve* expression in the blue channel. Both these gene expression patterns consist of seven stripes as illustrated by Fig. 5(a),(b). Figure 5(c) shows a plot of the average relative levels of gene expression, as a function of embryo egg length, for *ftz* (purple line, 24 embryos) and *eve* (yellow line, 33 embryos), measured from lateral strips along both sides of each embryo. In Fig. 5(d),(e), the average nuclear volumes are plotted, for the same lateral strips, before and after channel unmixing, respectively. Clearly, the anomalous correlation of measured nuclear volumes with expression pattern in Fig. 5(d) has been removed by channel unmixing, as shown by Fig. 5(e).

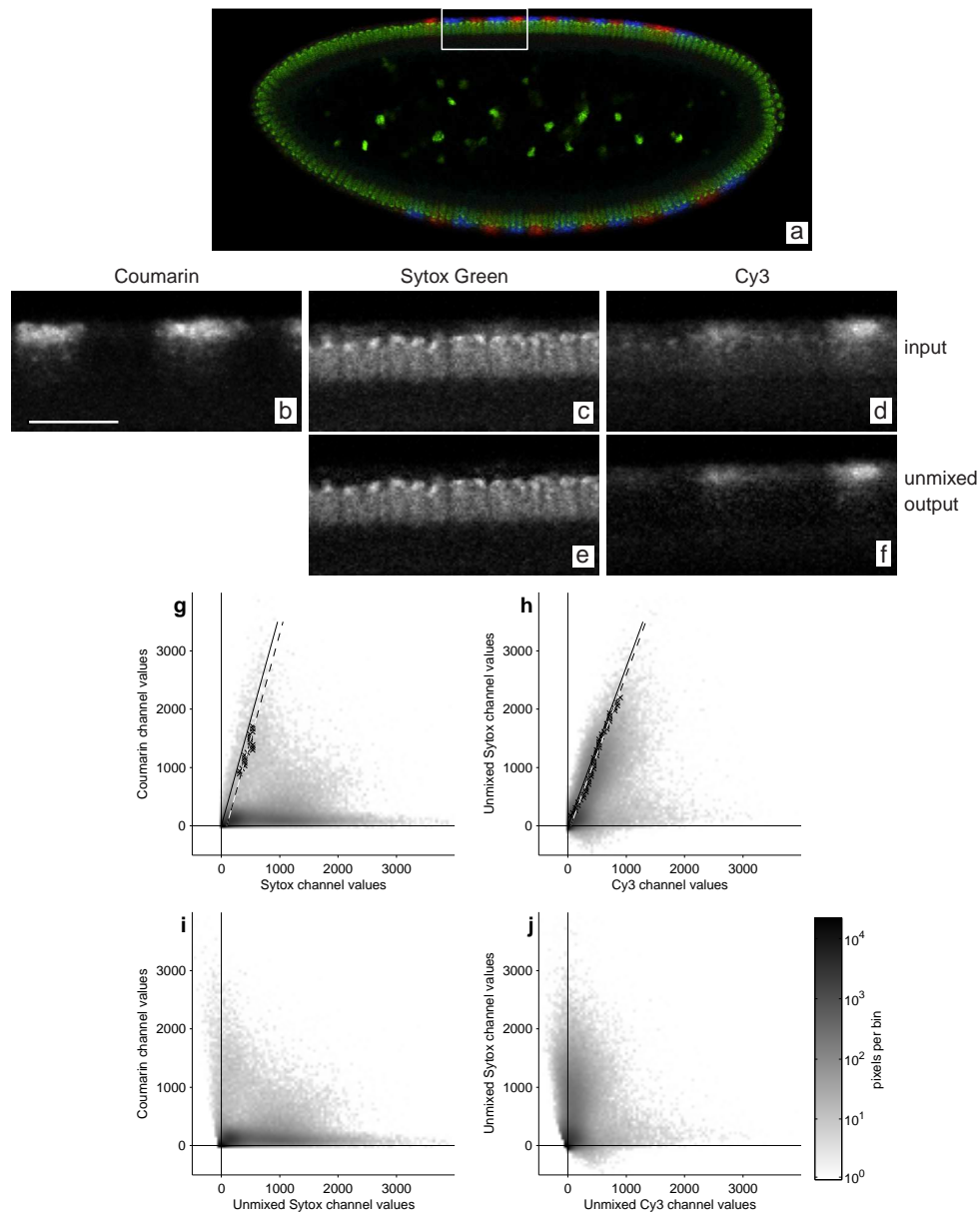


Fig. 4. Application of the method to a high-throughput image analysis pipeline. (a) An optical slice through the middle of a 3D confocal image of a fruit fly embryo, stained for DNA with Sytox Green, for *ftz* mRNA with Coumarin (blue) and for *eve* mRNA with Cy3 (red). The white rectangle indicates the region of interest used for (b)-(f). (b) Region of interest from the blue channel, (c) the green channel and (d) the red channel as measured. (e) The image from the green and (f) red channel after unmixing using the proposed algorithm. (g) Joint histogram of the blue channel versus the green channel. (h) Joint histogram of the green channel versus the red channel. (g) and (f) show the points detected by the algorithm (*crosses*) and the linear fit through these points (*dashed line*). (i) Joint histogram of the blue channel versus the green channel after automatic unmixing. (j) Joint histogram of the green channel and the red channel after automatic unmixing. Comparison of (b) with (e) and (c) with (f) shows how images in the green and red channels are improved after unmixing. The bar in (b) is 20 μm . The same, small gamma change has been performed on (b)-(f) to enhance the contrast in the dark areas and thus make the cross-talk better visible.

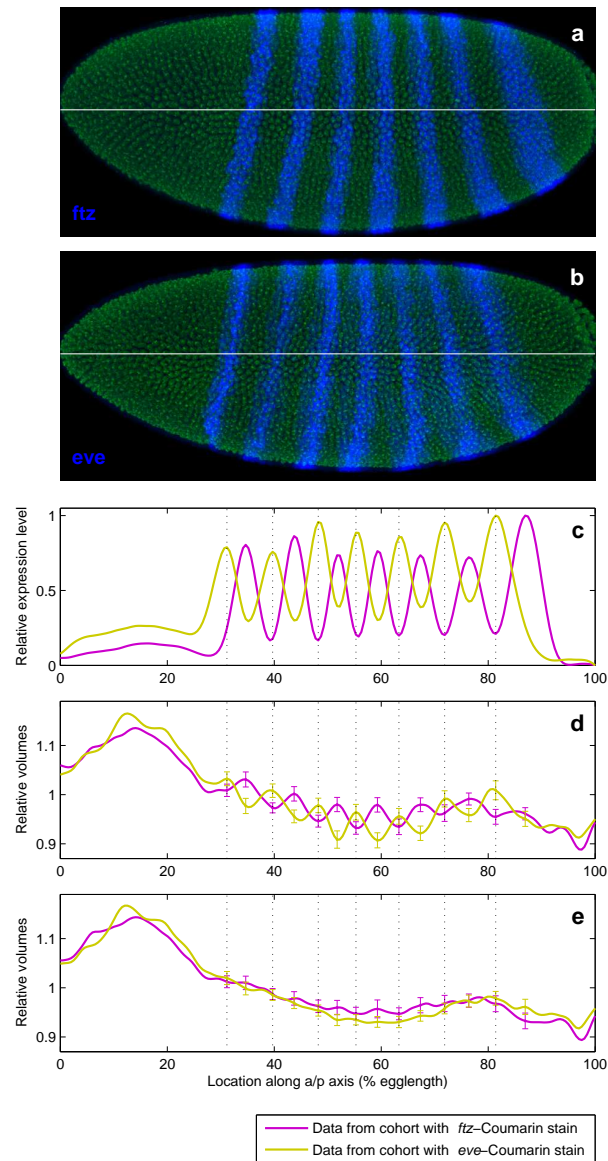


Fig. 5. Channel unmixing removes the effect of cross-talk and increases the accuracy of analysis results. (a) A maximum intensity projection along the optical axis from a single two-photon fluorescence image of an embryo stained for *eve* with Coumarin. DNA is shown in green. (b) An image of an embryo stained for *ftz* with Coumarin. (c) Plot of averaged *eve* (yellow line) and *ftz* (purple line) mRNA expression levels along the anterior/posterior axis on the embryo. (d) Measured nuclear volumes when either *eve* or *ftz* is stained for in the Coumarin channel. (e) Measured nuclear volumes from the same images after fully automatic channel unmixing using the algorithm presented here. Plots in (c)-(e) are the averages of the data from a cohort of embryos. The vertical dotted lines in these plots, at the location of maximum *eve* intensity, indicate the correlation between expression level and measured volumes. (d) and (e) also show the corresponding 95% confidence limits.

3. Discussion

For high-throughput, quantitative, image-based analysis, existing cross-talk reduction methods are not applicable. Here, we have presented an algorithm that estimates the cross-talk level between two channels, assuming unidirectional cross-talk between different dyes that have some spatial separation in the image. Both these assumptions are practical as we have shown from joint histograms between image channels. Our method corrects skew in the histogram, caused by cross-talk, by detecting image pixels which have recorded only the shorter wavelength dye. Only a small fraction of the total image pixels are required for this, and thus most multi-channel images will fulfill both assumptions. The only case where our method is unable to unmix channels is when there are no pixels that have recorded the shorter wavelength dye only. This condition exists when there is complete spatial overlap between the signals in both channels, and while this is unlikely, our method simply detects this condition from the histogram, reports it, and does not attempt to unmix the images. For some studies, such as colocalization, images are recorded to detect specific overlapping events. While our algorithm does not attempt to detect colocalized pixels, and as long as the channel overlap is partial, the unmixing algorithm is completely compatible with images recorded for colocalization studies. The algorithm also is sensitive to, but does not remove, autofluorescence, and is able to correctly unmix the two channels even in its presence. All of these properties make the algorithm robust enough for unsupervised use. Although the algorithm was developed for unidirectional cross-talk, if cross-talk in the other direction cannot be ignored, it can still be estimated simply by applying our algorithm to both axes. Unmixing bi-directional cross-talk requires solving a system of linear equations rather than performing a simple subtraction.

The algorithm has been used successfully in a fully automated analysis pipeline that has measured gene expression and morphology in thousands of whole *Drosophila* blastoderm embryos [9, 10]. Results of this application demonstrated that channel unmixing is not only important for the direct quantification of dye signals but also for higher order analyses which, in this case, uses measured dye signals to quantify morphological features.

4. Methods

4.1. The linear mixing model

Our automated unmixing methods assumes both the fluorescence yield and the detector operation are in their linear domains and the fluorescent dyes in the sample have not been saturated [1]. These conditions are achieved in many systems by properly setting laser intensity, detector gain, and detector integration time. The method also assumes that photobleaching is negligible. For three dyes with concentrations f_n , the measured light intensity S_n in three corresponding detector channels [1] is given by

$$\begin{aligned} S_1 &= a_1 f_1 + c_{2,1} a_2 f_2 + c_{3,1} a_3 f_3 + g_1 \\ S_2 &= c_{1,2} a_1 f_1 + a_2 f_2 + c_{3,2} a_3 f_3 + g_2 \\ S_3 &= c_{1,3} a_1 f_1 + c_{2,3} a_2 f_2 + a_3 f_3 + g_3 \end{aligned}$$

In these equations the parameters a_n represent sample and imaging dependent parameters such as the excitation and emission spectra of the dye, the quantum efficiency of the detector, the laser wavelength and intensity, and optical filter band-pass characteristics. The parameters $c_{n,m}$ are the channel mixing constants, and g_n are the autofluorescence components in each channel. The pure “unmixed” fluorescence signals, $S'_n = a_n f_n$, are then obtained by solving this set of linear equations with $g_n = 0$. The autofluorescence components g_n are ignored to make the system of equations determined. Because we assume unidirectional cross-talk from S_1 into S_2

and from S_2 into S_3 , many of the constants $c_{n,m}$ can be ignored. This leaves only $c_{1,2}$ and $c_{2,3}$ as a significant contribution to cross-talk. Under this assumption, $S'_1 = S_1$. The second channel contains some cross-talk from the first channel, which can be directly subtracted after appropriate weighting: $S'_2 = S_2 - c_{1,2}S'_1$. This unmixed signal can then again be used to correct the third channel: $S'_3 = S_3 - c_{2,3}S'_2$.

4.2. Cross-talk determination using the joint histogram

To measure the cross-talk level $c_{1,2}$ we used the two-dimensional joint histogram which shows the intensity distribution of all image pixels between two image channels. Consider the two channels with signals S_1 and S_2 , which collect fluorescence from dyes 1 and 2. Dye 1 is the shorter wavelength dye, its emission is primarily recorded in S_1 . The emission from dye 2 is primarily recorded in S_2 . In the joint histogram of S_1 versus S_2 , each image pixel is accumulated into bins according to its intensity in the two channels [11]. The joint histogram gives insight into the correlation between S_1 and S_2 , including the amount of colocalization between the two dyes [12] and the channel cross-talk [13].

In the absence of cross-talk, all image pixels that register fluorescence from one dye only will accumulate along either of the two axis of the joint histogram. Specifically, we call *cluster 1* the pixels without contribution from dye 2, which lie along the S_1 (vertical) axis. Pixels that register fluorescence in both channels will be distributed throughout the histogram, away from the axes. However, in the presence of cross-talk, pixels that register fluorescence from dye 1 will also record a fraction $c_{1,2}$ of that fluorescence level in S_2 . S_2 will therefore have a linear dependence on S_1 , which is particularly apparent for the pixels in cluster 1. This cross-talk causes cluster 1 to be linearly skewed away from the S_1 axis of the joint histogram with a gradient of $1/c_{1,2}$ [13, 14]. Thus, one can measure the cross-talk level by measuring the gradient of the best fit line through this cluster. The cross-talk $c_{2,3}$ for the third channel can be determined in the same manner from a joint histogram of S'_2 versus S_3 , where S'_2 is the unmixed version of S_2 . Note that this method will only work if cluster 1 can be detected, and thus there must exist a small fraction of image pixels which contain dye 1 but not dye 2.

4.3. The algorithm

To estimate the cross-talk level $c_{1,2}$ the algorithm first generates the joint histogram, $H(S_1, S_2)$, for the two channels S_1 and S_2 . In our implementation we have divided each axis into 100 bins and removed any artifacts produced by the binning by convolving the histogram with a two dimensional Gaussian of $\sigma = 1$ bin. The signal S_1 from the shorter wavelength channel, which bleeds through to the longer wavelength channel, is on the vertical axis (y). Next, the algorithm locates the first local maximum $m(y)$ for each line $H_y(x) = \{H(x, y)|y\}$, which are horizontal lines in the joint histogram. Each of these lines $H_y(x)$ is the histogram of S_2 intensities for pixels with a fixed value y in S_1 . The locations of these local maxima form a set of points $(y, m(y))$ which represent cluster 1. A linear least squares fit to this set of points yields an offset and a slope. The slope is the inverse of the cross-talk level $c_{1,2}$ for S_1 into S_2 . The offset is caused by additional fluorescent components such as autofluorescence, and can be ignored.

Some simple tests have been implemented to ensure robustness of the algorithm and to catch input images of poor quality. Firstly, to ensure accurate determination of the maxima $(y, m(y))$, we require there to be at least 100 pixels in the histogram along each line $H_y(x)$ from which a maximum is determined. This test simply ensures the presence of a minimum amount of data to accurately determine the maximum. Second, because we are looking for cluster 1, the set of pixels that lies closest to the y -axis, we require that the number of pixels on each line $H_y(x)$ to the left of the maximum $m(y)$ is less than the number to the right, $\sum_{x=0}^{m(y)} H_y(x) < \sum_{x=m(y)}^{\infty} H_y(x)$. Thirdly, because the fractional cross-talk is independent of pixel intensity, the

detected points $(y, m(y))$ must lie along a straight line. To ensure this we require that at least 8 maxima survive the first two conditions and that they produce a linear correlation coefficient that exceeds 0.7. Note, these three tests are lenient and are defined simply to catch aberrant input image data. Certainly, more stringent tests could be devised to ensure the unmixing algorithm worked correctly in specific cases. However, this was unnecessary for the data presented in this work, which were of sufficient quality to easily pass these tests and in most cases produced correlation coefficients well above 0.9.

4.4. Test slide

To evaluate the unmixing algorithm we used a standard fluorescent slide (FluoCells #2 Molecular Probes, Carlsbad, California) of bovine pulmonary artery endothelial cells stained with Texas Red phalloidin, which binds F-actin (red dye), and anti- α -tubulin antibody conjugated to a BODIPY labeled antibody, which binds microtubules (green dye). The slide was imaged on Zeiss 510 confocal microscope (Carl Zeiss MicroImaging, Inc., Thornwood, NY) with a 63x oil immersion objective (1.4 NA). Single photon excitation was used at 488 nm and 543 nm to excite the green and red dyes respectively. The fluorescence emission was collected by independent photomultiplier detectors at wavelengths between 500 nm and 560 nm (green channel) and wavelengths greater than 560 nm (red channel).

4.5. Embryo image acquisition and analysis

Whole *Drosophila* embryos were imaged, for the Berkeley *Drosophila* Transcription Network Project [9, 10], on a Zeiss 510 laser scanning confocal microscope (Carl Zeiss MicroImaging, Inc., Thornwood, NY) using a Plan-Apochromat 20x, 0.75 numerical aperture objective lens. Two selected mRNA gene products were hybridized with probes and labeled with Coumarin and Cy3 (Perkin-Elmer, Wellesley, MA), respectively, and nuclear DNA was stained with Sytox Green (Molecular Probes, Carlsbad, CA). The three dyes were excited simultaneously using two-photon excitation at 750 nm, provided by a Chameleon ultra-fast laser (Coherent, Inc., Santa Clara, CA). Images of up to 1024 by 1024 by 140 pixels were recorded by three independent channels (blue, green and red, as shown in Fig. 1). The images were processed in a fully automated image analysis pipeline [9]. For each embryo, this pipeline produced a table with the location and extent of all the nuclei, measured from the green channel, and the relative gene expression per nucleus, measured from the blue and red channels.

For the analysis results shown in Fig. 4, a single image was used, taken from an embryo which had mRNA for the gene *fushi tarazu* (*ftz*) labeled with Coumarin, and mRNA for *even skipped* (*eve*) labeled with Cy3.

For the analysis results shown in Fig. 5, where we were interested in the cross-talk from the blue channel to the green channel, two cohorts of embryo images were used. One cohort contained images of 24 embryos which had *ftz* mRNA expression labeled with Coumarin. The other cohort contained images of 33 embryos that had *eve* mRNA expression labeled with Coumarin. In both cohorts, embryos were from a tight (20 minute) temporal window during the 14th mitotic interphase, right before gastrulation, where membrane invagination along the ventral surface was between 50% and 100% [9]. The cohorts were selected from embryos which had their dorsal/ventral axis at $(90^\circ \pm 22.5^\circ)$ to the optical axis of the microscope. This presented the sides of the embryo perpendicular to the optical axis, so that the image analysis was most accurate in these regions [9]. The gene expression and nuclear volumes were then extracted from the table produced by the image analysis pipeline. Two strips running along the anterior/posterior (a/p) axis of the embryo were selected. Each strip was centered on one of the embryo's lateral midlines and covered one-sixth of its surface. The measured nuclear volumes and gene expression levels within each strip were then normalized and projected onto the a/p

axis. The volume normalization was such that the average nuclear size within the strip was 1. The expression level normalization was such that the maximum and minimum levels were 1 and 0, respectively.

Acknowledgments

Funded by a National Institutes of Health grant 1R01 GM70444-01A1 and conducted at Lawrence Berkeley National Laboratory under contract DE-AC03-76SF00098 from the U.S. Department of Energy. The images shown in Fig. 4 and 5 have been acquired for the Berkeley *Drosophila* Transcription Network Project (<http://bdtntp.lbl.gov/>).