© 2018. Published by The Company of Biologists Ltd | Journal of Cell Science (2018) 131, jcs211847. doi:10.1242/jcs.211847

REVIEW

Image co-localization – co-occurrence versus correlation

Jesse S. Aaron, Aaron B. Taylor and Teng-Leong Chew*

ABSTRACT

Fluorescence image co-localization analysis is widely utilized to suggest biomolecular interaction. However, there exists some confusion as to its correct implementation and interpretation. In reality, co-localization analysis consists of at least two distinct sets of methods, termed co-occurrence and correlation. Each approach has inherent and often contrasting strengths and weaknesses. Yet, neither one can be considered to always be preferable for any given application. Rather, each method is most appropriate for answering different types of biological question. This Review discusses the main factors affecting multicolor image co-occurrence and correlation analysis, while giving insight into the types of biological behavior that are better suited to one approach or the other. Further, the limits of pixel-based co-localization analysis are discussed in the context of increasingly popular super-resolution imaging techniques.

KEY WORDS: Image analysis, Manders, Pearson, Co-localization, Fluorescence microscopy

Introduction

A common task in cell biology consists of assessing to what extent two biomolecules or structures are associated with each other within a cell. Optical microscopy offers a sensitive and specific means to infer such relationships. While various distinct approaches that attempt to characterize biomolecular associations from multi-color fluorescence images have been developed, they are all collectively referred to as co-localization analysis. Previous reviews on this subject have generally given a survey of various co-localization methods, with recommendations for use based on image characteristics (Dunn et al., 2011; Bolte and Cordelières, 2006; Zinchuk and Grossenbacher-Zinchuk, 2001; Costes et al., 2004; Zinchuk et al., 2007), or advocated for the superiority of one approach over another (Adler and Parmryd, 2010). However, the various methods discussed here and previously measure two distinct phenomena: co-occurrence and correlation. The former describes the extent of spatial overlap between two fluorophores. The latter determines the degree to which the abundance of two spatially overlapping fluorophores are related to each other. Determining the most appropriate metric, therefore, should be guided by the nature of the question being posed. Co-occurrence measurements are often best utilized to determine what proportion of a molecule is present within particular area, compartment or organelle. It does not give insight into any concentration relationship between two molecules. Correlation analysis is most applicable when assessing a functional or stoichiometric relationship between two overlapping species. It does not, however, measure the extent of their spatial co-occurrence. Both measures can supply complementary information about a biological system. This Review explores how co-localization can be more strategically employed. We outline strengths and pitfalls of correlation and co-occurrence, both in terms of what phenomena they measure, as well as common sources of inaccuracies. Finally, we relate these methods to emerging super-resolution imaging techniques and introduce new approaches based on spatial point statistics.

Advanced Imaging Center, Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Dr., Ashburn, VA USA.

*Author for correspondence (chewt@janelia.hhmi.org)

T.-L.C., 0000-0002-3139-7560

Prerequisites for analysis

It is important to note that neither image correlation nor co-occurrence are direct measures of molecular interaction. The resolving power of a microscope is, conventionally, limited to approximately half the wavelength of emitting light (Abbe, 1873), while typical interaction distances between bio-molecules are <10 nm. Even with the advent of super-resolution imaging techniques, intramolecular interactions cannot be unambiguously observed. Only nearfield biophysical techniques, such as Förster resonant energy transfer (FRET), can be used to directly measure molecular interactions (Truong and Ikura, 2001; Piston and Kremers, 2007). Nevertheless, correlation and co-occurrence offer a means to infer such relationships.

In this context, optimized sample preparation (Wysocki and Lavis, 2011; Grimm et al., 2015; Shaner et al., 2005; Chudakov et al., 2010) and proper image acquisition settings are both essential for accurate analysis (North, 2006; Waters, 2009; Stelzer, 1998; Nakamura, 2005). In addition, post-acquisition corrections should be made for inhomogeneous illumination (Sternberg, 1983; Dickinson et al., 2002; Leong et al., 2003). Co-registration of the component images either due to chromatic aberration, focal plane drift or multi camera acquisitions (Lange et al., 2008), may also be necessary (Zitová and Flusser, 2003). Accurate measurements also depend on digital removal of unwanted, non-biologically relevant signal (Wu et al., 2010; Young et al., 2004; Russ and Neal, 2016). Finally, it is critical to isolate the pixels that contain signal, while ignoring those pixels containing predominantly noise, which can be achieved via thresholding (Mehmet Sezgin and Bülent Sankur, 2004; Nakagawa and Rosenfeld, 1979; Glasbey, 1993; Otsu, 1975; Pun, 1980; Russ and Russ, 1987).

Area and object analysis

At its simplest, co-localization analysis can consist of measuring the area of overlap between the signals of interest in two images. Fig. 1 illustrates a hypothetical example of the measurement of overlap between areas or objects. Fig. 1A shows a two-color image of multiple objects that are significantly larger than the diffraction limit, with spatial overlap between the Color 1 (Fig. 1B) and Color 2 (Fig. 1C) images, shown separately for clarity. An appropriate signal threshold intensity was calculated with the commonly used Otsu's method (Otsu, 1975) and all pixels below this value were assigned a zero value. The remaining pixels in the Color 1 and Color 2 images are shown in white in Fig. 1D and Fig. 1E, respectively. The area of overlap is then found by determining which pixel locations contain non-zero values in both images. Referred to as the

Journal of Cell Science (2018) 131, jcs211847. doi:10.1242/jcs.211847

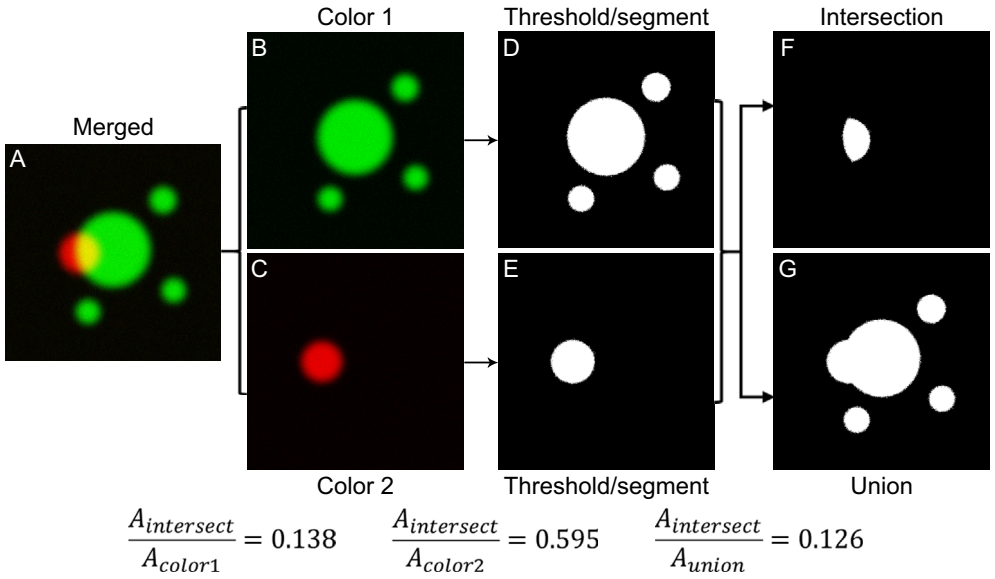


Fig. 1. Area overlap analysis is a simple means to assess biological associations. (A) A simulated two-color image indicates the presence of some overlap between red and green pixels. (B,C) Individual color channels are subject to threshold application to isolate the signals of interest. (D,E) Pixels containing signal are shown in white, whereas the remaining area is given a value of zero intensity. (F,G) Pixel locations containing signal in both images (intersection, F), whereas pixel locations containing signal in either image (union, G). Shown below, the fractional area in F, relative to the areas in D, E and G, were found to be 0.138, 0.595 and 0.126, respectively.

intersection of the two images, the result is shown in Fig. 1F. The overlap area, however, is most meaningful as a relative measure. For instance, the intersection area (Fig. 1F) as a fraction of the segmented Color 1 image area (Fig. 1D) indicates the fractional overlap of Color 1 with Color 2. Likewise, the intersection area as a fraction of the segmented Color 2 image area (Fig. 1E) specifies the fractional overlap of Color 2 with Color 1. Finally, the overlapping area can be expressed as a fraction of the total area, or the union of the Color 1 and Color 2 images (Fig. 1G).

The area overlap analysis can be extended further. As suggested by Fig. 1D and E, contiguous areas of above-threshold pixels can be grouped into discrete objects, with associated properties such as size, location and shape (Wu et al., 2010). This allows a statistical analysis of the relative distribution of objects across image color channels. For example, the distribution of distances from each object's center-of-mass in one color channel to its nearest neighbor in the other channel can yield meaningful insight when compared to negative controls or over time (Bolte and Cordelières, 2006). To assess the significance of any changes in object interaction, Helmuth et al., proposed a rigorous significance testing framework (Helmuth et al., 2010).

However, the limits of diffraction must always be considered. Assuming a diffraction-limited resolution of ~300 nm, objects – such as whole cells, nuclei and larger organelles – can lend themselves well to this type of analysis. However, smaller structures, such as actin filaments, microtubules, small vesicles, or single molecules or molecular clusters, can cause misinterpretation since their apparent sizes are determined by diffraction.

Co-occurrence: Manders' coefficients

Aside from the threshold calculations used in the example in Fig. 1, there is no consideration of the individual intensity values in the areas that contain both signals of interest. Manders introduced a method that determines the overlap of two images while taking into account pixel intensity, which we term co-occurrence (Manders et al., 1993). In other words, it accounts for the total amount (or abundance) of fluorophores that overlap with each other. This results in two coefficients, such that

$$M_1 = \frac{\sum_{i=1}^{n} x_{i,coloc}}{\sum_{i=1}^{n} x_i}\,, \tag{1}$$

where

$$x_{i,coloc} = \left\{ \begin{matrix} x_i \; if \; y_i > 0 \\ 0 \; if \; y_i = 0 \end{matrix} \right\} \tag{2}$$

and:

$$M_2 = \frac{\sum_{i=1}^{n} y_{i,coloc}}{\sum_{i=1}^{n} y_i}\,, \tag{3}$$

where

$$y_{i,coloc} = \left\{ \begin{matrix} y_i \; if \; x_i > 0 \\ 0 \; if \; x_i = 0 \end{matrix} \right\}. \tag{4}$$

Here, x_i and y_i refer to the i^{th} above-threshold pixel value in the color 1 and color 2 images, respectively, with n total pixels in each image being analyzed. As noted, $x_{i,coloc}$ and $y_{i,coloc}$ may have non-zero values only when the corresponding y_i and x_i, respectively, are also above threshold. Thus, M_1 can be defined as the co-occurrence fraction of color 1 with color 2. Likewise, M_2 is the co-occurrence fraction of color 2 with color 1. Furthermore, Manders proposed an overall overlap coefficient (MOC), such that

$$MOC = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}}\,, \tag{5}$$

where x_i, y_i and n are defined as previously. Manders' coefficients provide an important distinction over the simpler area overlap calculation (Fig. 1) because they give greater importance to brighter pixels and less weight to values near zero or the threshold. Using Manders' approach will ameliorate (but not ignore) the effects of inadvertently including dim unwanted signals, such as autofluorescence or other near-threshold signals. However, images that contain bright unwanted signal (such as that from high non-specific labeling, large camera offset, or out-of-focus light) or large amounts of image saturation due to poor acquisition parameters will inflate the MOC value. Such artefacts can give the impression of greater co-occurrence than is actually present.

However, the MOC is relatively insensitive to differences in signal-to-noise ratio (SNR) (Manders et al., 1993). Since noise will cause a random deviation in the intensity of each pixel, the effect of increasing noise relative to signal is largely averted

Journal of Cell Science (2018) 131, jcs211847. doi:10.1242/jcs.211847

when summing over a large number of pixels during calculation of the M_1, M_2 or MOC coefficients. Fig. 2A illustrates this phenomenon. As can be seen, a reduction of >50% in SNR (from 19 to 8) results in only a relatively small change in the MOC value of 3%.

While being robust against fluctuating SNR, an accurate MOC value is contingent on effective removal of unwanted signal. Because undesirable signal sources will increase pixel intensity (even in areas that also contain the signal of interest), they should always be carefully subtracted from the image. Fig. 2B illustrates this effect. On the left, a hypothetical image is shown that contains only the signal of interest and accompanying noise, which exhibits a MOC value of 0.691. A relatively dim out-of-focus light contribution (middle image), shown at twice actual brightness, was added to create the image shown on the right. As a result, the MOC value increases by 20% (0.691 vs 0.842) compared to the original image, thereby giving a false impression of higher co-occurrence. For this reason, strategies, such as unsharp masking (Polesel et al., 2000) can help to avoid overestimation of the MOC value by removing unwanted signal, such as that arising from camera offset or out-of-focus light. Higher intensity or higher spatial frequency sources of signal contamination, such as bright non-specific labeling or autofluorescence, may be more difficult to remove. Thus, while relatively unaffected by fluctuations in SNR, the MOC depends sensitively on accurate removal of non-biologically relevant signal sources.

A second, but frequently overlooked, aspect of the Manders' coefficients is the need to determine whether or not the co-occurrence observed between two images is primarily due to random chance. Costes et al. proposed an approach to perform a statistical significance test of any image co-occurrence or correlation measure (Costes et al., 2004). To do this, one image is subjected to a spatial 'scrambling', whereby pixel values are re-assigned to random locations within the image. Importantly, this scrambling should occur with 'blocks' of neighboring pixels to account for the fact nearby pixels are typically correlated with each other. An appropriate estimate of the pixel block size can be determined by performing an image autocorrelation to assess texture size within the image. Once scrambled, this image can then be compared to the other channel image (which remains unaltered) and an MOC value is calculated. The process of scrambling and recalculating MOC values is then repeated for many iterations. The resulting collection of randomized MOC values is then compared to the 'true' MOC value between the original images. Typically, if the true MOC exceeds 95% of the scrambled values, it can be deemed statistically significant and not primarily due to random chance. While robust, this method can also be computationally intensive. An alternative method proposed by Dunn, et al. simply rotates one channel image 90° to act as a negative control (Dunn et al., 2011), which is compared with the original images. While considerably simpler, this method does not provide a statistically robust means to validate a co-occurrence or correlation measure.

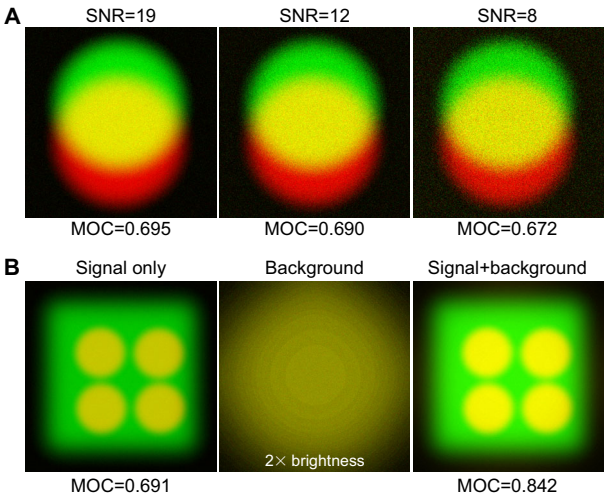


Fig. 2. Manders' overlap coefficient is insensitive to signal-to-noise ratio but sensitive to out-of-focus signal. (A) Illustrated here is the relatively small effect of decreasing signal-to-noise ratio (SNR) on the value of the Manders' overlap coefficient (MOC). In this example, a decrease in SNR of 60% results only in a 3% decrease in MOC for the images (left to right). The random fluctuation in pixel intensity owing to noise is largely balanced out when many pixels are summed up during calculation of the MOC value. (B) Illustration of the effect of out-of-focus signal. Here, an original image (left) displays an MOC value of 0.69. Upon adding a relatively low-intensity background signal with low spatial frequency (middle image), shown for clarity at twice the actual brightness, the resulting image (right) exhibits a >20% increase in MOC value, even after the threshold value has been correctly adjusted.

Fig. 3 shows two contrasting examples for significance testing of a MOC calculation. In Fig. 3A, a two-color image with largely overlapping structures is shown on the left; the MOC value for this image pair is correspondingly high with a value of 0.91. The green signal was then randomly scrambled 1000 times using pixel blocks that match the size of the features in the image (~60×60 pixels), and the MOC value was calculated for each case. An example of such a scrambled image is shown in Fig. 3A, middle. A histogram of the resulting MOC values calculated for 1000 scrambled green versus the unscrambled red image is shown in Fig. 3A, right. The red arrow indicates the true (unscrambled) MOC value. In this case, the true MOC value exceeds all of those calculated for the randomly scrambled images. Thus, there is very high likelihood that the co-occurrence calculated for these images is not due to chance. An opposite example is shown in Fig. 3B. The left panel shows a different two-color image with a similar MOC value than that in Fig. 3A. Nevertheless, when Color 1 (green) is scrambled 1000 times (a single example is shown in Fig. 3B, middle), the resulting randomized MOC value histogram (right), although narrowly distributed, tends to fall near the original value as indicated by the red arrow. Therefore, in this example, it cannot be discounted that the MOC value is due to random chance and any inferences drawn from it are not statistically supportable.

Taken together, co-occurrence analyses offer advantages and disadvantages that depend on the biological context and should, thus, be carefully considered. Co-occurrence analysis offers a clear intuitive means to assess a relationship between two images by calculating the concentration-weighted overlap between them. This approach is largely insensitive to noise and small errors in thresholding, which is particularly useful for inherently dim samples. However, the presence of non-biologically relevant signal can give a false overestimate of the overlap between two images. Thus, careful correction for illumination heterogeneity (Leong et al., 2003), appropriate sample preparation and acquisition parameters (North, 2006; Waters, 2009), and proper subtraction of unwanted signal is essential to achieve accurate results (Wu et al., 2010; Young et al., 2004; Russ and Neal, 2016). Furthermore, it is imperative to measure the statistical significance of any co-occurrence or correlation value in order to assess the degree to which similar results can be obtained by chance and – by inference – would carry little biological meaning.

Journal of Cell Science (2018) 131, jcs211847. doi:10.1242/jcs.211847

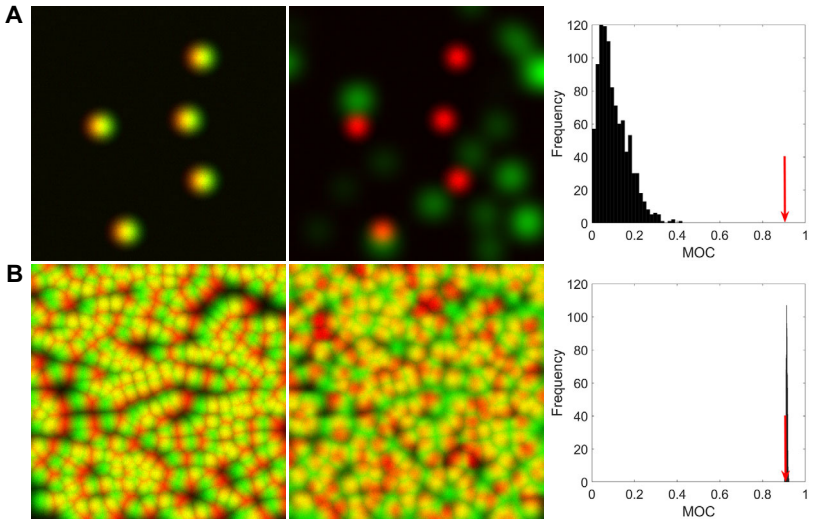


Fig. 3. Image randomization to test the statistical significance of the Manders' overlap coefficient. (A) A true multicolor image (left) that exhibits a Manders' overlap coefficient (MOC) value of 0.91. By using Costes' randomization method, the green-color image was scrambled 1000 times in blocks whose size was determined by autocorrelation. A single example of a randomized image (right). MOC was calculated for each of the scrambled green images and the original red-color image. The histogram shows the resulting values, in which the true (unscrambled) MOC value (red arrow) exceeds the randomized MOC values in all cases, indicating high statistical confidence in the MOC. (B) A different multicolor image (left), with an almost identical MOC value to A (MOC=0.90). Costes' randomization method was applied to the green-channel image as before. An example of a scrambled green image, overlaid with the original red image is shown (right). The resulting MOC histogram shows that the true MOC value is not consistently greater than the randomized values, suggesting the co-occurrence is mainly due to chance and, thus, has poor statistical significance.

In general, Manders' coefficients are useful for assessing to what extent a structure or molecule can be found in a particular location or organelle. For example, it has been used to quantify the co-occurrence of a molecule of interest with mitochondria (Bravo et al., 2011; Seibler et al., 2011), the plasma membrane (Yeung et al., 2008; Spira et al., 2012) or the endoplasmic reticulum (Horner et al., 2011; Arruda et al., 2014). Importantly, however, Manders' approach does not signify that any predictable relationship exists between the intensities in one image and the corresponding values in the other. The MOC only expresses the degree to which two structures spatially overlap, in an intensity-weighted manner. It cannot positively indicate, for example, that areas of high intensity in one image correspond to low-intensity areas in the other and vice versa, which may be a sign of an important biological phenomenon.

Correlation: Pearson's coefficient

While the Manders' overlap coefficients express the extent of co-occurrence between two images, correlation-based co-localization analyses are based on a different interpretation of co-localization. In this case, the guiding assumption states that, if two imaging targets are functionally related, then their abundances will also be predictably related to each other wherever they exist in the same region. This assumption is particularly relevant when probing two molecules that are thought to bind to (or repulse) each other within the cell – even if such interactions may be rare. In other words, correlation methods measure the relationship between the signal intensities in one image and the corresponding values in another, not the degree to which the signals co-occur.

Pearson's correlation coefficient (PCC) is a common metric to measure the predictability of this relationship (Pearson, 1896; Manders et al., 1993). In more mathematical terms, the PCC can be thought of as the covariance between the two images, normalized by the product of their standard deviations:

$$PCC = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \quad (6)$$

where x_i and $\bar{x}$ represent the i^{th} pixel intensity and the average pixel intensity (ignoring sub-threshold pixels), respectively, in the segmented color 1 image. Likewise, y_i and $\bar{y}$ are the corresponding i^{th} and mean values for the color 2 image. The value of n represents the total number of segmented pixels in both images. Recall that the value of n must be the same for both images since correlation analysis should be performed on the intersection of the images in question.

In more conceptual terms, the PCC determines to what extent the signal intensity variation in one image can be explained by the corresponding variation in the other, assuming a linear relationship. Since the calculation of the PCC involves the difference of pixel intensity from the population mean, it can be either positive or negative and can range from −1 to 1. The magnitude of the PCC gives a measure of the predictability of the relationship. Values close to 1 or −1 indicate a near-perfect ability to infer a color 2 intensity, given the corresponding color 1 intensity and vice versa. Values near zero indicate that there is little predictive value between the images and that the two species being imaged do not have a clear correlation. The sign of the PCC indicates the 'direction' of the relationship between color 1 and color 2, with a positive sign indicating that, when color 1 intensity increases, color 2 tends to increase proportionally, pointing to a molecular attraction. A negative sign implies the reverse, suggesting molecular repulsion. The ability to distinguish positive and negative correlation, therefore, represents a significant strength of the PCC coefficient over the Manders' methods, owing to its ability to quantify both positive and negative associations.

Critical to the understanding of correlation-based image similarity measurements is the use of the scatterplot. Scatterplots are constructed by plotting the intensity value of each pixel in one image along the x-axis and on the y-axis the intensity value of the same pixel location in the second image, thereby forming a bivariate histogram that describes the relationship between the corresponding intensity values. It is important to note that a given region within a scatterplot does not necessarily correspond to specific areas within the images, as this representation contains only intensity information.

Most importantly, a multi-color image may exhibit relatively high co-occurrence, while – at the same time – it can be poorly correlated, and vice versa. For example, analysis of the multicolor image shown in Fig. 4A results in relatively large Manders' coefficients, with M_1, M_2 and MOC values of 0.81, 0.89 and 0.89, respectively. However, upon inspection of the corresponding scatterplot (Fig. 4B), a predictive relationship between the intensities of corresponding green and red pixels is not apparent, and calculation of the PCC confirms this with a relatively low value of 0.11. Fig. 4C and D illustrate a contrasting example. In this case, the M_1, M_2 and MOC values are found to be 0.13, 0.15 and 0.14, respectively, indicating little co-occurrence. Nevertheless, the corresponding PCC value of

Journal of Cell Science (2018) 131, jcs211847. doi:10.1242/jcs.211847

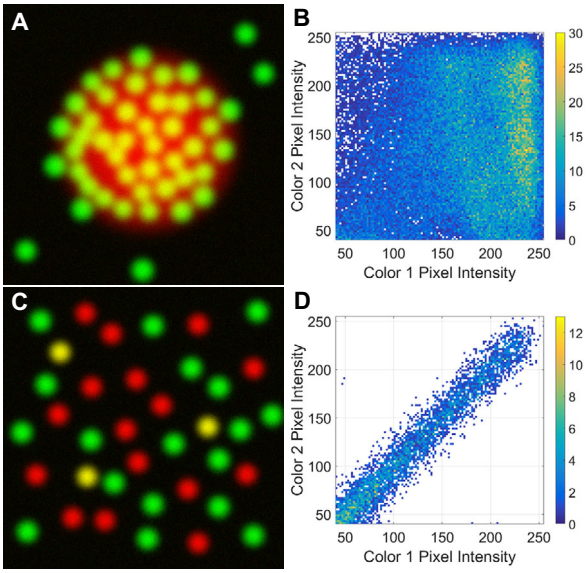


Fig. 4. Co-occurrence and correlation can occur independently of each other. (A,B) The image shows an example of high co-occurrence (MOC=0.89) but with low correlation (PCC=0.11) (A). The low correlation can be clearly seen in the corresponding intensity scatterplot (B), as there is no clear statistical dependence between the corresponding pixels of each color image. (C) By contrast, the image shown here displays a relatively low co-occurrence between the red and green channels (MOC=0.14). (D) Nevertheless, when only the intersection of both color images is considered, the corresponding pixel intensities have a clear linear relationship, with a PCC value over the intersection of the red and green images almost equaling one. This serves to illustrate that correlation and co-occurrence are independent phenomena, as they measure distinct behavior in an image pair. While co-occurrence gauges the overlap between signals in two images, correlation measures the relationship between those signals in overlapping areas.

0.97 indicates an almost perfect correlation among the image intersection. Therefore, although the data in Fig. 4C and D do not co-occur to a high degree, the intensity values are closely related wherever they overlap in the two images. Thus, the Manders' and Pearson's coefficients can, in certain circumstances, give different indications of image similarity. Although seemingly contradictory, such a result can provide powerful biological insight, as well as emphasizes the inherent differences between co-occurrence and correlation.

The sensitivity of a PCC value to changing SNR and the addition of unwanted low spatial frequency signal differs considerably from co-occurrence approaches. To demonstrate this, Fig. 5A shows how the PCC values for three green-red image pairs are affected by progressively decreasing signal-to-noise ratio (SNR). This decrease in SNR is accompanied by a clear broadening of the corresponding scatter plots shown below each image pair, with an ∼35% reduction in PCC value. This behavior is in contrast to that of the Manders' coefficients, where relatively large changes in SNR have little effect on the MOC value (refer to Fig. 2A).

Likewise, correlation and co-occurrence methods also respond differently to unwanted low-frequency signal. Whereas the MOC requires careful subtraction of such signals for accurate measurements (refer to Fig. 2B), the PCC is less sensitive to such artefacts. This can be illustrated by simply adding a constant offset to the images in Fig. 5A. Although this will shift the data points in the scatterplot along the *x*- and *y*-axes, the predictability of their relationship will remain the same (ignoring saturated pixels).

Therefore, a potential weakness of the PCC method lies in its sensitivity to SNR. Decreasing SNR results in the relationship between the pixel intensities of each image becoming less predictive and increases the chances of an inadvertent inclusion of dim unwanted signal during threshold calculation. Imaging the same field of view multiple times to measure noise characteristics can lessen the impact of low SNR on PCC value, but requires some *a priori* assumptions about the underlying relationship between the images (Adler et al., 2008). To identify an optimal threshold value for correlation analysis, Costes et al. also proposed a means to segment images on the basis of computing the PCC value across a range of threshold values (Costes et al., 2004). In this approach, continually decreasing thresholds for the two images are proposed, and the PCC value is calculated for pixels both above and below those threshold values. When the PCC value for the sub-threshold pixels nears zero, an assumption is made that a suitable segmentation of signal from background has been made. However, care must be taken when using Costes' threshold regression. If the signal of interest is not well correlated compared to the background, there will not be a clear demarcation in the PCC values between the two. It is also important to note that Costes' significance testing scheme is equally applicable to correlation analysis, as illustrated in the co-occurrence analysis in Fig. 3 (Costes et al., 2004). It should, therefore, always be performed to assess the reliability of an MOC or PCC result.

Correlation: Spearman's coefficient

There are cases where PCC can give unexpected results. This includes situations where the two images are well – but not linearly – correlated. In other words, the relationship between two images may be very predictable but the scatterplot describing an image intensity relationship is not well described by a straight line. In these types of case, the PCC can underestimate the correlation between images because the PCC can only approach its maximal magnitude (1 or −1) when the pixel intensity relationship is linear. However, Spearman's rank correlation coefficient (SRCC) can address this issue. In short, the SRCC is equivalent to the PCC, but is applied to pixel intensity ranks rather than to the intensities themselves (Spearman, 1904). The conversion from pixel value to pixel rank proceeds such that the lowest above-threshold pixel intensity in the image is given a value of 1, the next lowest value is assigned a rank of 2, and so on. In cases where multiple pixels have the same intensity, that value is given an average rank. For example, if two pixels are tied for 3rd and 4th lowest value (rank 3–4), they are given an 'average' rank of 3.5. The practical effect of this transformation is that it linearizes the scatterplot from the two images, making the Pearson's analysis applicable to non-linear correlation.

Fig. 5B,C illustrates how the SRCC can be used to properly measure multicolor image correlation. In Fig. 5B, two almost identical images are displayed (first two panels). The corresponding intensity scatterplot is displayed in the third and the ranked intensity scatterplot in the last panel. Note that both the intensity and ranked intensity scatterplots show a clear linear relationship, and the PCC and SRCC are both near unity. However, Fig. 5C illustrates the advantage of SRCC over PCC. Here, although the color 1 image is identical to that in Fig. 5B, the color 2 image has been altered so it produces the intensity scatterplot shown in the third panel. This plot indicates a well-correlated but non-linear relationship between color 1 and color 2. Despite the good predictability of the scatterplot, the PCC value is lower than expected (0.875) due to its non-linearity. However, by ranking the pixel intensity values, the relationship becomes linearized and the resulting SRCC value (0.989; last panel) does more accurately reflect the predictive relationship of the images. Importantly, the SRCC is also useful for reducing the effect

Journal of Cell Science (2018) 131, jcs211847. doi:10.1242/jcs.211847

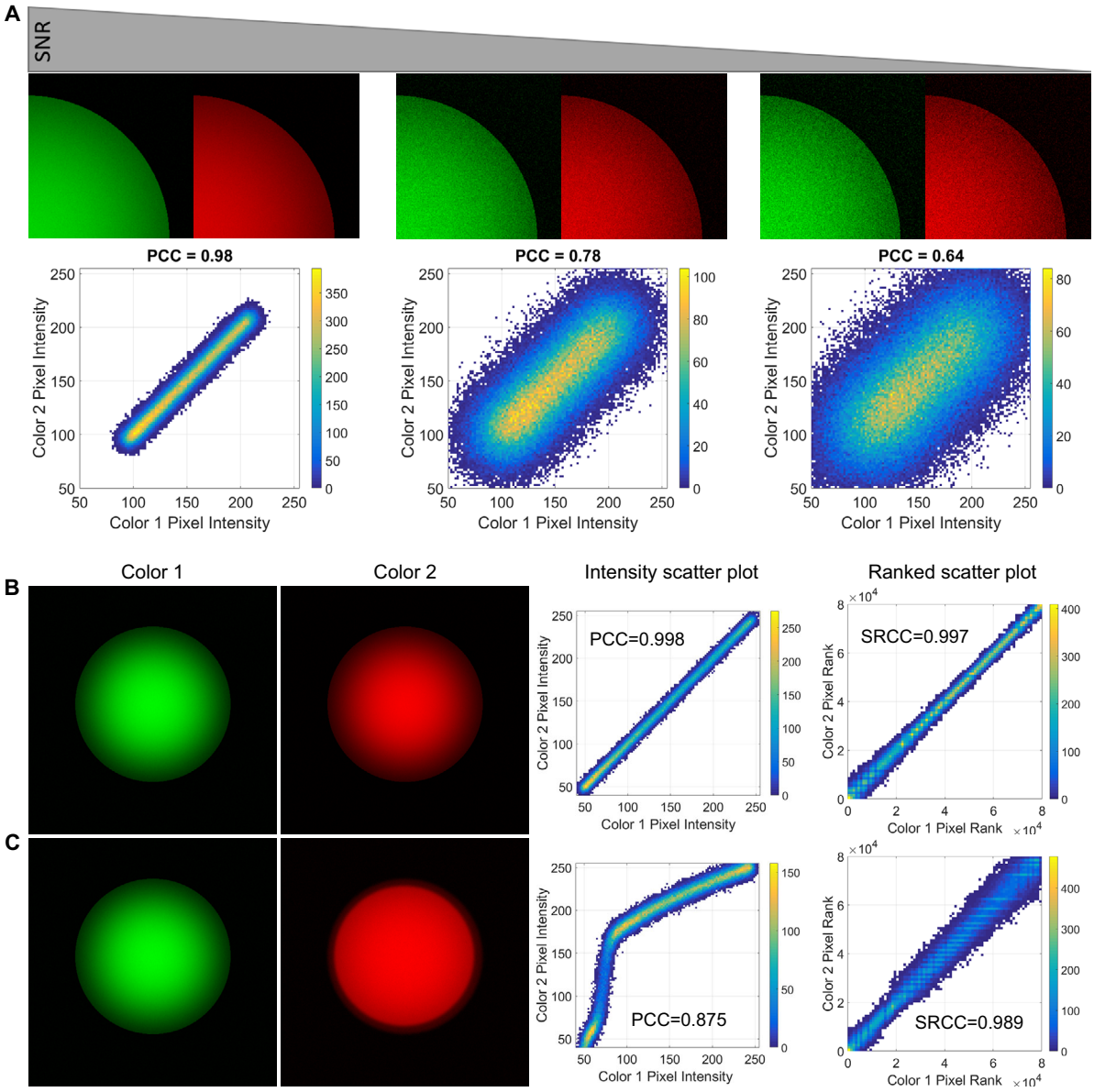


Fig. 5. The Pearson's correlation coefficient is sensitive to both signal-to-noise ratio and scatterplot non-linearity. (A) Shown here is a series of three green and red image pairs whose signal-to-noise ratio (SNR) is progressively decreasing from 50 to 13 (scale not shown) from left via middle to right image. The corresponding scatterplots are shown below each image pair; they illustrate progressive broadening, and a decrease in the Pearson's correlation coefficient (PCC) of ∼35%. (B) Two, almost identical, images shown in the green (left) and red (right) channel, and the corresponding intensity and ranked intensity scatterplots, both of which are linear. (C) The same green image as shown in B (left) compared to a different red image (right). This image pair exhibits a non-linear relationship in their intensity scatterplot with a lower PCC value of 0.875. However, a ranked intensity scatterplot (see main text) recovers the linear relationship, with the resulting SRCC value almost equaling one (SRCC=0.989).

of pixel saturation as, in this case, the ranked value will generally deviate less from the corresponding mean than the absolute intensity value.

Unlike the Manders' approach, correlation-based analyses do not measure the extent to which one image overlaps with another. Indeed, high values of PCC or SRCC can occur even when the fractional co-occurrence is low (see Fig. 4). This can be especially apparent when two imaging targets are anti-correlated or when they only rarely interact. Correlation analysis simply measures how well the intensities in one image predict those in the other image, when signal is present in the same pixel. This can be a powerful method to suggest a functional relationship, such as – among many others – those shown for the assembly of endocytosis regulators (Teis et al., 2008), association of focal adhesion components (Roca-Cusachs et al., 2013; Carisey et al., 2013) or viral replication machinery assembly (Hsu et al., 2010).

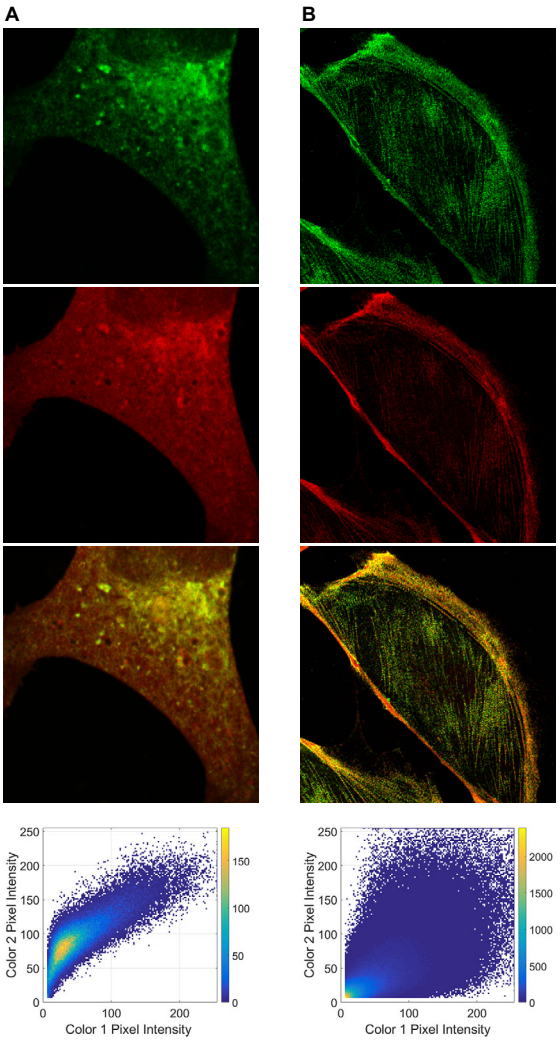
Taken together, Figs 4 and 5 illustrate the requirements and limitations of PCC or SRCC as a measure of co-localization. Further, these illustrations indicate the fundamental differences between co-occurrence and correlation, with an eye towards inferring the type of biological analysis that may be better suited to one approach than another. For these correlation-based methods to give meaningful results, the SNR in both images should be maximized to the extent that is practical. As shown in Fig. 5, the PCC can fail to give an expected result if the correlation between two images is non-linear. Interestingly, however, calculating the

Journal of Cell Science (2018) 131, jcs211847. doi:10.1242/jcs.211847

SRCC for linearly correlated images still provides an accurate result. For these reasons, the SRCC should always be used in favor of the PCC. Fundamentally, however, both correlation and co-occurrence measurements yield different interpretations of co-localization. Yet, these interpretations can often be complementary.

To further illustrate these considerations, Fig. 6 shows two examples of biological associations. In A (top panel), a confocal image of an U2OS cell reveals the distribution of epidermal growth factor receptor (EGFR) (Herbst, 2004), tagged with GFP. The middle image shows Rab13 (Zerial and McBride, 2001) expressed as an mCherry-tagged fusion protein in the same cell. Fig. 6B shows confocal images of a Ptk2 cell immunostained for both total myosin (top panel) and phosphorylated myosin (middle panel), respectively. The last image in A and B, each displays the corresponding pixel intensity scatterplots for each of the protein pairs. Appropriate signal threshold values for each image were calculated by using Otsu's method (Otsu, 1975). Then, the Manders', Pearson's and Spearman's coefficients were calculated as described previously, and their results are summarized in Fig. 6C.

Qualitatively, the image pairs appear to have comparable levels of similarity and the MOC values are, indeed, almost the same. However, the individual Manders' coefficients reveal important differences. Although almost all of the EGFR signal overlaps with that of Rab13, not all Rab13 co-occurs with EGFR. This suggests that, although Rab13 may associate with EGFR, it may also be associated with other molecules at different cellular locations. Total and phosphorylated myosin, however, co-occur with each other nearly equally. The level of correlation between these two examples is strikingly different, as suggested by their corresponding scatterplots. The PCC and SRCC values confirm this, with a two-fold difference in their value. The intersecting EGFR and Rab13 concentrations predict each other relatively well, indicating a concentration-dependent relationship between these molecules (Ioannou and McPherson, 2016). However, the overall abundance of myosin does not predictably determine the concentration of phosphorylated myosin (and vice versa), suggesting the presence of other regulatory mechanisms (Goeckeler et al., 2000). This example illustrates the value to co-localization analysis in a holistic approach that considers both overall and individual co-occurrence as well as correlation, in order to gain a broader range of biological insights.



C

	Protein Pair	
Coefficient	EGFR/Rab13	Myosin epitope 1/2
MOC	0.72	0.68
M_1	0.99	0.68
M_2	0.44	0.71
PCC	0.76	0.36
SRCC	0.74	0.29

Fig. 6. Interpreting co-occurrence and correlation analysis in two examples of biological data. (A) Confocal images of an U2OS cell that reveal the distribution of epidermal growth factor receptor (EGFR) (top) and Rab13 (middle), expressed as GFP- and mCherry-tagged fusion proteins, respectively. The two-color overlay image is shown in the bottom image and the corresponding pixel intensity scatterplot is shown underneath. (B) Confocal images of a Ptk2 cell immunostained for total myosin (#3674, Cell Signaling) (top) and phosphorylated myosin (#922701, BioLegend) (middle), after application of secondary antibodies conjugated to Alexa Fluor 488 and Alexa Fluor 594 (Thermo Fisher), respectively. The corresponding two-color overlay is shown in the bottom image and the corresponding pixel intensity scatterplot is shown underneath. (C) Summary of the co-occurrence and correlation coefficients, including Manders' overlap (MOC), individual Manders' coefficients (M_1 and M_2), Pearson's correlation coefficient (PCC) and Spearman's rank correlation coefficients (SRCC). As can be seen, both image pairs show similar levels of total overlap, with almost identical MOC values. However, the fractional overlap values M_1 and M_2 indicate important differences between the two biological systems. Furthermore, the PCC and SRCC values also indicate that, while EGFR and Rab13 have a clear concentration relationship, the abundance of phosphorylated mysosin is not dependent on the total amount of mysosin, suggesting a secondary effector. Thus, careful consideration of both correlation and co-occurrence metrics can reveal different aspects of these biological systems.

Super-resolution imaging

As discussed initially, any image similarity measure is subject to the resolving power of the imaging system being used. Super-resolution imaging circumvents the diffraction-limit and may offer detail that is otherwise unavailable. Techniques that improve resolution by ∼1.5 to 2-fold, such as structured illumination microscopy (SIM) (Gustafsson et al., 2008), or image scanning microscopy and its derivatives (Müller and Enderlein, 2010; York et al., 2013), offer the possibility of increased accuracy when using pixel-based image similarity measurements – as long as the interacting structures of interest both occur within in the same image pixels. Indeed, imaging targets that appear co-localized under conventional imaging, may in fact be well separated with even a modest improvement in resolution (Schermelleh et al., 2008). However, the near-molecular level resolution of single-molecule localization (SML) techniques, such as photo-activation and localization microscopy (PALM) (Betzig et al., 2006) or stochastic optical reconstruction microscopy (STORM) (Rust et al., 2006), begin to reveal the effects of the Pauli exclusion principle, which states that no two molecules can occupy the same space at the same time.

For SML techniques, spatial statistics have been utilized to quantitatively measure molecular associations without spatial overlap (Coltharp et al., 2014; Lagache et al., 2015; Nicovich et al., 2017; Rossy et al., 2014; Georgieva et al., 2016; Malkusch et al., 2012). Malkusch et al. proposed a method termed coordinate-based co-localization (CBC) analysis. The procedure first calculates two functions,

$$D_{x_i,x}(r) = \frac{N_{x_i,x}(r)}{N_{x_i,x}(R_{max})} \cdot \frac{R_{max}^2}{r^2} \quad (7)$$

and

$$D_{x_i,y}(r) = \frac{N_{x_i,y}(r)}{N_{x_i,y}(R_{max})} \cdot \frac{R_{max}^2}{r^2}, \quad (8)$$

where $N_{x_i,x}(r)$ indicates the number of molecules of type x (or color 1) that are found within a radius r surrounding a given single localization, x_i, of the same molecule type. R_{max} refers to the maximum search radius, which should be larger than the expected interaction distance. In the second expression, $N_{x_i,y}(r)$ is defined analogously, but counts the number of molecules of type y (or color 2) within the same radius r and maximum distance R_{max}. For each x_i, an SRCC is then calculated between $N_{x_i,x}(r)$ and $N_{x_i,y}(r)$. In this way, a correlation value can be assigned to each molecule within the SML image. However, instead of a measure of intensity correlation, it reflects how well two types of molecule are correlated in space. By assigning each localization a CBC value, a 'map' can be constructed that shows areas of both high and low spatial correlation within the image.

Another class of SML co-localization measures is based on Ripley's K-function (Ripley, 1977, 1976) and can serve as an ensemble measure of interaction distance between two different molecules:

$$K_{ij}(r) = \frac{A}{N_i N_j} \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \frac{I(d_{ij} < r)}{w_{ij}}. \quad (9)$$

Here, A is the total imaging area (or volume in the case of 3D data), while N_i and N_j are the total number of localized molecules in each imaging channel. $I(d_{ij}<r)$ has a value of 1 if the distance between the i^{th} particle in channel one and the j^{th} particle in channel two (denoted by d_{ij}) is less than r and zero otherwise. A correction factor, w_{ij}, is included to account for undercounting particles near the edge of the image (Haase, 1995). If two molecules are randomly distributed with respect to each other, then $K_{ij}(r)=\pi r^2$. Thus, we can define an L-function, such that

$$L_{ij}(r) = K_{ij}(r) - \pi r^2. \quad (10)$$

Finding local maxima (or minima) in $L_{ij}(r)$ can, therefore, yield characteristic interaction (or repulsion) distances (Kiskowski et al., 2009). A variant of Ripley's K-function is termed pair-correlation function. Here, only the molecules contained within a ring with inner radius r and outer radius $r+dr$ are included. Pair correlation is more sensitive to local changes in molecule density around a given localization, which can result in a more accurate measure of characteristic interaction distance. However, it also depends on a high localization density within the image.

Importantly, Ripley's K-function (or pair correlation function) and CBC can be used in tandem to gain complementary information. CBC assigns a correlation value to each localization in the image, thereby creating a map that can highlight areas of high or low association. Ripley's K-function (and the related pair correlation function), while only offering a global measure of association, can provide a characteristic association distance between two different molecules that is missing in the CBC analysis.

Conclusions and future perspectives

Quantitative image co-localization analyses, while utilized in numerous life science studies, represent a toolbox that is prone to flawed usage and misinterpretation. Contributing to this problem is the fact that co-localization can be interpreted in at least two distinct ways – co-occurrence or correlation. Both interpretations have inherent, and often opposing, strengths and weaknesses. On one hand, measuring co-occurrence by using the Manders' coefficients can offer an intuitive accounting of the concentrated weighted overlap between two imaging targets, with relative insensitivity to imaging noise. Yet, MOC values can be artefactually inflated by unwanted signal, such as out-of-focus light or endogenous fluorescence. On the other hand, correlation-based analysis offers a means to evaluate the intensity interdependence between images and, thus, is better able to distinguish both molecular attraction and repulsion from random association. However, correlation measures such as PCC and SRCC work best when both images display a high SNR and are more sensitive to changes in threshold values. Importantly, Spearman's rank correlation (Spearman, 1904) should always be preferred over Pearson's correlation (Pearson, 1896) to account for possible non-linear associations between images, and to guard against any pixel pair outliers, such as those occurring due to image saturation. The statistical significance of any MCC, PCC or SRCC value should be assessed by Costes' randomization approach (Costes et al., 2004) wherever possible.

Previous treatments of this subject have given recommendations on the appropriate approach based on image characteristics (Dunn et al., 2011; Bolte and Cordelières, 2006) advocated for the superiority of correlation over co-occurrence (Adler and Parmryd, 2010) or attempted to incorporate both concepts into a merged metric of overall co-localization (Zinchuk et al., 2013). We argue that the specific biological question at hand should guide both the image acquisition and analysis strategy. As illustrated in Figs 4 and 6, co-occurrence and correlation can occur independently of each other, and the extent of either phenomenon is largely determined by the biological behavior being probed. Considering both types of metric can also yield complementary information.

In addition, the effect of so-called 'global bias' has not been historically discussed in the context of image co-localization. Global bias refers to any external factor that can affect the interaction between two molecules, apart from their affinity for each other. Zaritsky and colleagues have proposed a powerful means to separate local interactions from global relationships that can confound co-localization measurements (Zaritsky et al., 2017). Factors that can induce a global bias are numerous and can be biological or non-biological, such as cell shape, cell cycle state, spatially correlated noise or offset in the detector, out-of-focus signal, or any combination thereof. For this reason, any imaging modality that does not suppress out-of-focus light might lead to inaccurate results due to its inclusion of signal from above and below the focal plane. Confocal and TIRF microscopy are, therefore, better imaging methodologies for such analyses, as they confine excitation or detection to within the depth of field of the objective lens (Conchello and Lichtman, 2005; Schneckenburger, 2005). Furthermore, the advent of super-resolution techniques can, in principle, offer greater fidelity when inferring molecular interactions. However, structures that might spatially overlap

when interrogated with conventional modalities can, in fact, be well-separated when using subject to super-resolution microscopy. In general, as achievable resolution improves, the intersection of two multi-channel images necessarily approaches zero and pixel-based similarity measures are rendered obsolete, favoring spatial statistical approaches instead (Nicovich et al., 2017; Coltharp et al., 2014; Lagache et al., 2015). But, while the ultra-high resolution attainable in PALM and STORM imaging can be attractive, these methods typically preclude imaging dynamic samples. Further, the preferred means to analyze these data is computationally intensive. Thus, careful consideration must be paid to the underlying biological behavior being investigated to select the optimal imaging and analysis method. Any co-localization measurements are most meaningful when expressed as relative changes. Evaluation of a coefficient value relative to stringent controls (both positive and negative) and/or over time will significantly strengthen its ability to draw meaningful biological conclusions. In summary, the complexity of co-localization analysis demands careful consideration to determine the best approach to answer a given biological question. Furthermore, as imaging technology continues to improve, strategies for measuring biomolecular associations will be required to evolve concomitantly.

Acknowledgements
We thank Satya Khuon, Maria Ioannou and Damien Alcor (HHMI Janelia Research Campus) for assistance in preparing and imaging biological samples.

Competing interests
The authors declare no competing or financial interests.

Funding
The authors gratefully acknowledge support from the Advanced Imaging Center at Janelia Research Campus, a facility jointly funded by the Gordon and Betty Moore Foundation and the Howard Hughes Medical Institute.

References
Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Arch. Für Mikrosk. Anat.* **9**, 413-418.
Adler, J. and Parmryd, I. (2010). Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. *Cytometry A.* **77A**, 733-742.
Adler, J., Pagakis, S. N. and Parmryd I. (2008). Replicate-based noise corrected correlation for accurate measurements of colocalization. *J. Microsc.* **230**, 121-133.
Arruda, A. P., Pers, B. M., Parlakgül, G., Güney, E., Inouye, K. and Hotamisligil, G. S. (2014). Chronic enrichment of hepatic endoplasmic reticulum-mitochondria contact leads to mitochondrial dysfunction in obesity. *Nat. Med.* **20**, 1427-1435.
Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J. and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642-1645.
Bolte, S. and Cordelières, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213-232.
Bravo, R., Vicencio, J. M., Parra, V., Troncoso, R., Munoz, J. P., Bui, M., Quiroga, C., Rodriguez, A. E., Verdejo, H. E., Ferreira, J. et al. (2011). Increased ER–mitochondrial coupling promotes mitochondrial respiration and bioenergetics during early phases of ER stress. *J. Cell Sci.* **124**, 2143.
Carisey, A., Tsang, R., Greiner, A. M., Nijenhuis, N., Heath, N., Nazgiewicz, A., Kemkemer, R., Derby, B., Spatz, J. and Ballestrem, C. (2013). Vinculin regulates the recruitment and release of core focal adhesion proteins in a force-dependent manner. *Curr. Biol.* **23**, 271-281.
Chudakov, D. M., Matz, M. V., Lukyanov, S. and Lukyanov, K. A. (2010). Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol. Rev.* **90**, 1103.
Coltharp, C., Yang, X. and Xiao, J. (2014). Quantitative analysis of single-molecule superresolution images. *Curr. Opin. Struct. Biol.* **28**, 112-121.
Conchello, J.-A. and Lichtman, J. W. (2005). Optical sectioning microscopy. *Nat. Methods* **2**, 920-931.
Costes, S. V., Daelemans, D., Cho, E. H., Dobbin, Z., Pavlakis, G. and Lockett, S. (2004). Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys. J.* **86**, 3993-4003.
Dickinson, M. E., Bearman, G., Tille, S., Lansford, R. and Fraser, S. E. (2002). Multi-spectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy. *BioTechniques* **31**, 1272-1278.
Dunn, K. W., Kamocka, M. M. and McDonald, J. H. (2011). A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol. Cell Physiol.* **300**, C723-C742.
Georgieva, M., Cattoni, D. I., Fiche, J.-B., Mutin, T., Chamousset, D. and Nollmann, M. (2016). Nanometer resolved single-molecule colocalization of nuclear factors by two-color super resolution microscopy imaging. *Methods* **105**, 44-55.
Glasbey, C. A. (1993). An analysis of histogram-based thresholding algorithms. *CVGIP Graph. Models Image Process.* **55**, 532-537.
Goeckeler, Z. M., Masaracchia, R. A., Zeng, Q., Chew, T.-L., Gallagher, P. and Wysolmerski, R. B. (2000). Phosphorylation of myosin light chain kinase by p21-activated kinase PAK2. *J. Biol. Chem.* **275**, 18366-18374.
Grimm, J. B., English, B. P., Chen, J., Slaughter, J. P., Zhang, Z., Revyakin, A., Patel, R., Macklin, J. J., Normanno, D., Singer, R. H. et al. (2015). A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat. Methods* **12**, 244-250.
Gustafsson, M. G. L., Shao, L., Carlton, P. M., Wang, C. J. R., Golubovskaya, I. N., Cande, W. Z., Agard, D. A. and Sedat, J. W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys. J.* **94**, 4957-4970.
Haase, P. (1995). Spatial pattern analysis in ecology based on Ripley's K-function: Introduction and methods of edge correction. *J. Veg. Sci.* **6**, 575-582.
Helmuth, J. A., Paul, G. and Sbalzarini, I. F. (2010). Beyond co-localization: inferring spatial interactions between sub-cellular structures from microscopy images. *BMC Bioinformatics* **11**, 372.
Herbst, R. S. (2004). Review of epidermal growth factor receptor biology. *Int. J. Radiat. Oncol.* **59**, S21-S26.
Horner, S. M., Liu, H. M., Park, H. S., Briley, J. and Gale, M. (2011). Mitochondrial-associated endoplasmic reticulum membranes (MAM) form innate immune synapses and are targeted by hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **108**, 14590-14595.
Hsu, N.-Y., Ilnytska, O., Belov, G., Santiana, M., Chen, Y.-H., Takvorian, P. M., Pau, C., van der Schaar, H., Kaushik-Basu, N., Balla, T. et al. (2010). Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. *Cell* **141**, 799-811.
Ioannou, M. S. and McPherson, P. S. (2016). Regulation of cancer cell behavior by the small GTPase Rab13. *J. Biol. Chem.* **291**, 9929-9937.
Kiskowski, M. A., Hancock, J. F. and Kenworthy, A. K. (2009). On the use of Ripley's K-function and its derivatives to analyze domain size. *Biophys. J.* **97**, 1095-1103.
Lagache, T., Sauvonnet, N., Danglot, L. and Olivo-Marin, J.-C. (2015). Statistical analysis of molecule colocalization in bioimaging. *Cytometry A* **87**, 568-579.
Lange, S., Katayama, Y., Schmid, M., Burkacky, O., Bräuchle, C., Lamb, D. C. and Jansen, R.-P. (2008). Simultaneous transport of different localized mRNA species revealed by live-cell imaging. *Traffic* **9**, 1256-1267.
Leong, F. J. W.-M., Brady, M. and McGee, J. O. (2003). Correction of uneven illumination (vignetting) in digital microscopy images. *J. Clin. Pathol.* **56**, 619-621.
Malkusch, S., Endesfelder, U., Mondry, J., Gelléri, M., Verveer, P. J. and Heilemann, M. (2012). Coordinate-based colocalization analysis of single-molecule localization microscopy data. *Histochem. Cell Biol.* **137**, 1-10.
Manders, E. M. M., Verbeek, F. J. and Aten, J. A. (1993). Measurement of co-localization of objects in dual-colour confocal images. *J. Microsc.* **169**, 375-382.
Müller, C. B. and Enderlein, J. (2010). Image scanning microscopy. *Phys. Rev. Lett.* **104**, 198101.
Nakagawa, Y. and Rosenfeld, A. (1979). Some experiments on variable thresholding. *Pattern Recognit.* **11**, 191-204.
Nakamura, J. (Ed.). (2017). Image sensors and signal processing for digital still cameras. CRC Press, Boca Raton, FL, USA.
Nicovich, P. R., Owen, D. M. and Gaus, K. (2017). Turning single-molecule localization microscopy into a quantitative bioanalytical tool. *Nat. Protoc.* **12**, 453-460.
North, A. J. (2006). Seeing is believing? A beginners' guide to practical pitfalls in image acquisition. *J. Cell Biol.* **172**, 9-18.
Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica* **11**, 23-27.
Pearson, K. (1896). Mathematical contributions to the theory of evolution. III. regression, heredity, and panmixia. *Philos. Trans. R. Soc. Lond. Math. Phys. Eng. Sci.* **187**, 253-318.
Piston, D. W. and Kremers, G.-J. (2007). Fluorescent protein FRET: the good, the bad and the ugly. *Trends Biochem. Sci.* **32**, 407-414.
Polesel, A., Ramponi, G. and Mathews, V. J. (2000). Image enhancement via adaptive unsharp masking. *IEEE Trans. Image Process.* **9**, 505-510.
Pun, T. (1980). A new method for grey-level picture thresholding using the entropy of the histogram. *Signal Process.* **2**, 223-237.
Ripley, B. (1976). The second-order analysis of stationary point processes. *J. Appl. Probab.* **13**, 255-266.

Ripley, B. D. (1977). Modelling spatial patterns. *J. R. Stat. Soc. Ser. B Methodol* **39**, 172-212.
Roca-Cusachs, P., del Rio, A., Puklin-Faucher, E., Gauthier, N. C., Biais, N. and Sheetz, M. P. (2013). Integrin-dependent force transmission to the extracellular matrix by α-actinin triggers adhesion maturation. *Proc. Natl. Acad. Sci.* **110**, E1361-E1370.
Rossy, J., Cohen, E., Gaus, K. and Owen, D. M. (2014). Method for co-cluster analysis in multichannel single-molecule localisation data. *Histochem. Cell Biol.* **141**, 605-612.
Russ, J. C. (2016). *The Image Processing Handbook*. (6th Ed.). CRC Press, Boca Raton, FL, USA.
Russ, J. C. and Russ, J. C. (1987). Automatic discrimination of features in grey-scale images. *J. Microsc.* **148**, 263-277.
Rust, M. J., Bates, M. and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793-796.
Schermelleh, L., Carlton, P. M., Haase, S., Shao, L., Winoto, L., Kner, P., Burke, B., Cardoso, M. C., Agard, D. A., Gustafsson, M. G. L. et al. (2008). Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* **320**, 1332.
Schneckenburger, H. (2005). Total internal reflection fluorescence microscopy: technical innovations and novel applications. *Anal. Biotechnol.* **16**, 13-18.
Seibler, P., Graziotto, J., Jeong, H., Simunovic, F., Klein, C. and Krainc, D. (2011). Mitochondrial park in recruitment is impaired in neurons derived from mutant PINK1 induced pluripotent stem cells. *J. Neurosci.* **31**, 5970.
Sezgin, M. and Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. **13**, 146-165.
Shaner, N. C., Steinbach, P. A. and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nat. Methods* **2**, 905-909.
Spearman, C. (1904). The proof and measurement of association between two things. *Am. J. Psychol.* **15**, 72-101.
Spira, F., Mueller, N. S., Beck, G., von Olshausen, P., Beig, J. and Wedlich-Söldner, R. (2012). Patchwork organization of the yeast plasma membrane into numerous coexisting domains. *Nat. Cell Biol.* **14**, 640-648.
Stelzer, E. H. K. (1998). Contrast, resolution, pixelation, dynamic range and signal-to-noise ratio: fundamental limits to resolution in fluorescence light microscopy. *J. Microsc.* **189**, 15-24.
Sternberg, S. R. (1983). Biomedical image processing. *Computer* **16**, 22-34.
Teis, D., Saksena, S. and Emr, S. D. (2008). Ordered assembly of the ESCRT-III complex on endosomes is required to sequester cargo during MVB formation. *Dev. Cell* **15**, 578-589.
Truong, K. and Ikura, M. (2001). The use of FRET imaging microscopy to detect protein–protein interactions and protein conformational changes in vivo. *Curr. Opin. Struct. Biol.* **11**, 573-578.
Waters, J. C. (2009). Accuracy and precision in quantitative fluorescence microscopy. *J. Cell Biol.* **185**, 1135-1148.
Wu, Q., Merchant, F. and Castleman, K. (2010). *Microscope Image Processing*. Academic Press, Burlington, MA, USA.
Wysocki, L. M. and Lavis, L. D. (2011). Advances in the chemistry of small molecule fluorescent probes. *Mol. Imaging* **15**, 752-759.
Yeung, T., Gilbert, G. E., Shi, J., Silvius, J., Kapus, A. and Grinstein, S. (2008). Membrane phosphatidylserine regulates surface charge and protein localization. *Science* **319**, 210.
York, A. G., Chandris, P., Nogare, D. D., Head, J., Wawrzusin, P., Fischer, R. S., Chitnis, A. and Shroff, H. (2013). Instant super-resolution imaging in live cells and embryos via analog image processing. *Nat. Methods* **10**, 1122-1126.
Young, I., Gerbrands, J. J. and Van Vliet, L. (1998). *Fundamentals Of Image Processing*. Delft University of Technology, Delft, The Netherlands.
Zaritsky, A., Obolski, U., Gan, Z., Reis, C. R., Kadlecova, Z., Du, Y., Schmid, S. L. and Danuser, G. (2017). Decoupling global biases and local interactions between cell biological variables. *eLife* **6**, e22323.
Zerial, M. and McBride, H. (2001). Rab proteins as membrane organizers. *Nat. Rev. Mol. Cell Biol.* **2**, 107.
Zinchuk, V. and Grossenbacher-Zinchuk, O. (2008). Quantitative colocalization analysis of confocal fluorescence microscopy images. In *Current Protocols in Cell Biology*. 4-16.
Zinchuk, V., Zinchuk, O. and Okada, T. (2007). Quantitative colocalization analysis of multicolor confocal immunofluorescence microscopy images: pushing pixels to explore biological phenomena. *Acta Histochem. Cytochem* **40**, 101-111.
Zinchuk, V., Wu, Y. and Grossenbacher-Zinchuk, O. (2013). Bridging the gap between qualitative and quantitative colocalization results in fluorescence microscopy studies. *Sci. Rep.* **3**, 1365.
Zitová, B. and Flusser, J. (2003). Image registration methods: a survey. *Image Vis. Comput.* **21**, 977-1000.