Title

Characterizing meiotic crossover designation mechanisms in C. elegans and developing advanced image analyses

Author

Stauffer, Weston Thomas

Publication Date

2020

Peer reviewed|Thesis/dissertation

Characterizing meiotic crossover designation mechanisms in *C. elegans* and developing advanced image analyses

By

Weston Stauffer

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Abby Dernburg, Co-Chair
Professor Michael Shapira, Co-Chair
Professor Peter Sudmant

Summer 2020

Characterizing meiotic crossover designation mechanisms in *C. elegans* and developing advanced image analyses

by
Weston Stauffer

Abstract

Characterizing meiotic crossover designation

mechanisms in *C. elegans*

and developing advanced image analyses

By

Weston Stauffer

Doctor of Philosophy in Integrative Biology

University of California, Berkeley

Professor Abby Dernburg

The concerted acquisition and analysis of microscopy has led to countless advancements in biology. By consistently advancing microscopy and analysis of images with fluorescent reporters, researchers have been able to characterize the interactions, dynamics in terms of movement, and functions of biomolecules within live and fixed cells. In this dissertation, a number of microscopy and advanced image analyses are performed to both build methods towards streamlining the analysis of microscopy images, and applying image analysis to challenging systems to answer fundamental questions in meiosis. Within these aims, this dissertation presents works that establish a powerful tool in quantifying the localization of proteins and colocalization between proteins. Building on this, additional analysis techniques are performed to characterize the dynamics of numerous proteins, and concentrations and effects of many of the known crossover designating and meiosis essential proteins in *C. elegans* meiosis. Finally, imaging and analysis of microscopy images is performed to further answer fundamental questions in meiosis and the designation of crossovers. Together this dissertation presents numerous advancements in tools for image analysis, applications of image analysis tools, application of cell biology tools, and the synthesis of these tools towards elucidating the mechanisms which govern a fundamental process in meiosis.

TABLE OF CONTENTS

Introduction to Part I:

Expanding the open-source toolset for microscopy image analysis

Recent years have seen countless advancements in microscopy, image acquisition, and image processing. These advancements are fueled by increased computing speeds, and developments in illumination microscopy modalities which push the limits of spatial and temporal resolution. Techniques such as lattice light-sheet microscopy[1] and AiryScan[2] now allow imaging with greatly reduced photobleaching of fluorescent proteins within live specimens. Other super-resolution microscopy techniques such as STORM[3], STED[4], and SIM[5] frequently push the spatial resolution of fixed specimens to below 100 nm. On the computing side, image processing such as the CBSDeep system[6] uses deep learning to process images taken under non-ideal conditions to achieve useful image reconstructions. Together these advancements have enabled researchers to answer diverse biological questions.

Many additional technical developments have also advanced imaging of different biomolecules in a wide range of cellular contexts. For example, the development of fluorescently tagged RNA-binding proteins such as MS2[7] made it possible to track specific RNA species, and more recently, the development of CRISPR/Cas9-based genome editing has greatly facilitated the ability to tag endogenous proteins with small epitope or fluorescent proteins in living eukaryotic cells[8]. Virtually any molecule of interest within a cell can potentially be made visible. However, while fluorescence imaging has expanded dramatically, the platforms and software to efficiently analyze these forms of data have been outpaced.

Many commercial tools have been developed for image analysis and display (e.g. Imaris, Metamorph), but the cost of these tools, some of which require ongoing subscriptions to maintain, can be prohibitive for many researchers. The growing movement towards reproducibility and “open science” has also driven many researchers to develop open-source, and often free solutions. With sufficient explanation, annotation, adaptability, and user-friendliness these open-source tools provide researchers a means to analyze the myriad of new data types

A key platform for open-source analysis is ImageJ/Fiji[9]. This Java-based image analysis tool has a staggering number of tools prepacked within it to allow for image analysis. Further, it is open-source and allows for expansion in the form of plugins and macros. Within this framework, a researcher can develop their analysis tools into a more user-friendly tool, requiring no knowledge or coding effort, which is freely available to others.

I encountered this lack of suitable analysis tools firsthand in the process of publishing other research projects with co-authors. In beginning colocalization analysis on real-world data we generated or downloaded, it was easiest to perform analysis within Matlab. Although a number of colocalization tools were available within ImageJ, they lacked key functionality to make them useful. Moreover, we found that the available metrics as tools to quantify colocalization performed poorly on nearly all real-world

images. Specifically, we found that the relationship of pixel values with respect to one another was nearly never linear or monotonic, because pixel values of two different reporters each labeling a different biomolecule often do not vary or scale in similar ways. Because of this relationship of pixel values, existing metrics such as Pearson's correlation coefficient (PCC) and Spearman's rank correlation coefficient (SRCC) could not identify simple patterns of anticolocalization or colocalization[10], due to their use of a linear correlation or rank correlation to non-linear and non-monotonic. Coefficients that require a set of thresholds such as Manders's colocalization coefficient (MCC) typically are used in conjunction with the Costes method[10] for an automated threshold identification. The Costes method depends upon a linear correlation, and when applied to real non-linear data regularly identified all or no pixels as above threshold within cells. Because of this incorrect thresholding, MCC also performed poorly[10]. In order to alleviate this, we developed the Threshold Overlap Score (TOS) and showed it to be a robust tool to quantify colocalization.

With a sound and reliable method to quantify colocalization, we investigated the localization of sRNA within *E. coli* cells[11]. However, we found that we needed additional tools to automate cell identification and segmentation, remove non-cell objects, and perform colocalization analysis. Thus, we implemented TOS analysis[10] with a workflow that included Image to identify suitable cells and Matlab to perform metric calculations. By using this approach, we were able to show that the localization of RNAs within the *E. coli* cell depended upon the size of an RNA with smaller RNAs readily entering the nucleoid and larger RNAs being excluded from the nucleoid[11]. While preparing to publish our workflow, we realized our approach was far from user-friendly. Because these tools required uses to be familiar with Matlab as well as ImageJ, we looked for a way to consolidate these approaches by developing an ImageJ plug-in.

We wanted this tool to be widely applicable, easy-to-use, open-source, customizable and modifiable by others. The results of our efforts is the EzColocalization plug-in. EzColocalization has gone through a significant number of iterations and tests which have resulted in a more robust and useful tool openly available to the research community. It is the general hope of the community that others will continue efforts to develop further tools to advance the toolbox of image analysis and as a result, expedite the findings of research around the globe.

REFERENCES

1 Chen, B.-C. *et al.* Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998, doi:10.1126/science.1257998 (2014).

2 Huff, J. The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. *Nature Methods* **12**, i-ii, doi:10.1038/nmeth.f.388 (2015).

3 Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* **3**, 793-796, doi:10.1038/nmeth929 (2006).

4 Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt Lett* **19**, 780-782, doi:10.1364/ol.19.000780 (1994).
5 Gustafsson, M. G. L. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy* **198**, 82-87, doi:10.1046/j.1365-2818.2000.00710.x (2000).
6 Weigert, M. *et al.* Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature Methods* **15**, 1090-1097, doi:10.1038/s41592-018-0216-7 (2018).
7 Querido, E. & Chartrand, P. in *Methods in Cell Biology* Vol. 85 273-292 (Academic Press, 2008).
8 Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096, doi:10.1126/science.1258096 (2014).
9 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
10 Sheng, H., Stauffer, W. & Lim, H. N. Systematic and general method for quantifying localization in microscopy images. *Biol Open* **5**, 1882-1893, doi:10.1242/bio.019893 (2016).
11 Sheng, H., Stauffer, W. T., Hussein, R., Lin, C. & Lim, H. N. Nucleoid and cytoplasmic localization of small RNAs in Escherichia coli. *Nucleic Acids Res*, doi:10.1093/nar/gkx023 (2017).

Chapter 1: EzColocalization: An ImageJ plugin for visualizing and measuring colocalization in cells and organisms

Originally published as: Stauffer, Weston et al. “EzColocalization: An ImageJ plugin for visualizing and measuring colocalization in cells and organisms.” *Scientific reports* vol. 8,1 15764. 25 Oct. 2018, doi:10.1038/s41598-018-33592-8

By Weston Stauffer, Huanjie Sheng, and Han N. Lim

ABSTRACT

Insight into the function and regulation of biological molecules can often be obtained by determining which cell structures and other molecules they localize with (*i.e*. colocalization). Here we describe an open source plugin for ImageJ called EzColocalization to visualize and measure colocalization in microscopy images. EzColocalization is designed to be easy to use and customize for researchers with minimal experience in quantitative microscopy and computer programming. Features of EzColocalization include: (i) tools to select individual cells and organisms from images; (ii) filters to select specific types of cells and organisms based on physical parameters and signal intensity; (iii) heat maps and scatterplots to visualize the localization patterns of reporters; (iv) multiple metrics to measure colocalization for two or three reporters; (v) metric matrices to systematically measure colocalization at multiple combinations of signal intensity thresholds; and (vi) data tables that provide detailed information on each cell in a sample. These features make EzColocalization well-suited for experiments with low reporter signal, complex patterns of localization, and heterogeneous populations of cells and organisms.

INTRODUCTION

Advances in microscopy equipment and labeling techniques make it possible for researchers to image a variety of biological molecules in almost any cell, tissue, or organism[1,2,3,4,5,6,7]. However, researchers often find it difficult to rigorously evaluate and interpret the images. In particular, it is often challenging to determine whether the different molecules of interest occur in the same locations, different locations or independent locations (*i.e*. colocalization, anticolocalization and noncolocalization respectively) in cells, tissues or organisms[8].

Several factors limit the use of current software for visualizing the localization of reporters in biological samples and measuring colocalization[9,10,11,12]. One factor is that customization of the software is often required for the equipment, reporters and samples[13,14], and for automated analyses. A second factor is that the software is often not suited to experiments that push the limits of detection, where the intensity of the

intracellular signal is similar to the extracellular signal (*i.e*. “background”)15, and where there are high levels of non-specific signal in cells8. The latter can occur because the probes or reporters are not sufficiently specific16, are not adequately removed from cells or organisms17, or have low signal relative to endogenous compounds (*i.e*. “autofluorescence”)18. That is, software tools are needed to distinguish intracellular pixels from extracellular pixels, and to select signal intensity thresholds to limit analyses to a subset of intracellular pixels. A third factor is that there are often mixed localization patterns within cells and different localization patterns among cells in a sample8,11,19. When this heterogeneity is present, software is need to provide measurements for each cell or defined subsets of cells in samples.

It is often possible to address the above challenges by combining multiple existing software programs and customizing them8,15. However, combining and customizing software requires proficiency in programming, experience with quantitative microscopy, comfort with mathematics and statistics, and other support. Many researchers do not have these skills or resources, and this is a likely reason that many studies evaluate colocalization by the simple, but often misleading, method of overlaying red and green color images10,11. Therefore there is a pressing need for a single application that provides all the tools for start to finish analysis of colocalization and can be easily customized.

In this study, an open source plugin for ImageJ called EzColocalization was developed so that researchers at all levels of proficiency can visualize the localization of signals and measure colocalization via an easy-to-use graphical user interface (GUI). The first part of the study describes EzColocalization, and the second part demonstrates its use for different sample types and for resolving common issues that prevent rapid and robust quantitative measurements of colocalization. EzColocalization can measure colocalization in cells, tissues, and whole organisms (*e.g. Caenorhabditis elegans* and *Drosophila* embryos); and the software is especially helpful where automation and customization is required, to obtain individual cell measurements in samples with many cells, and for reporters with low signal or low specificity.

METHODS AND MATERIALS

EzColocalization development

The code for EzColocalization was written in Eclipse Java Integrated Development Environment (IDE) release 4.3.020, which is a workspace for writing code and detecting compiling errors in Java™. EzColocalization incorporates ImageJ Application Program Interfaces (APIs) available from the National Institutes of Health, U.S. Department of Health and Human Services. An environment builder was used so that code written in the IDE ran in an instance of ImageJ as a plugin. This builder was implemented with Java Development Kit 821 and the ImageJ source code within the IDE. The WindowBuilder22 plugin for the IDE was used to design and generate the code for the

GUI, and the code produced was restructured and revised to improve readability, and add listeners, which obtain user inputs from the GUI for running the plugin.

The basic level of organization of the code for EzColocalization are “classes”. Classes are separated blocks of code that represent a set of methods and variables; a class may be devoted to performing calculations which share code or calculations that are most conveniently performed together. Classes with related operations are grouped into a higher level of organization termed “packages”. For example, a class that generates heat maps and a class that displays heat maps may be bundled into the same package. The classes and packages are described in detail in the Supplementary Information. Many processes within EzColocalization are performed as background computing, and thus the results of some classes, which are intermediates in longer methods, are not displayed and cannot be interacted with via the GUI.

Testing of EzColocalization

EzColocalization was tested on images from experiments and on modified images created to test specific issues (*e.g*. misalignment). Unpublished images of bacterial cells (HL6187) were used to illustrate the different modules of EzColocalization (Figs 1–4). These bacteria had plasmid pHL1392 in strain HL333823. pHL1392 has the ampicillin resistance gene, ColE1 origin, and the green fluorescent protein (GFP) fused to part of the *sodB* gene and transcribed from the PLlacO-1 promoter. The sources of the images used for the application experiments (Figs 5–8) are stated in the relevant Results section. Note: images presented in the figures are cropped so that it is easier to see individual cells.

Download and installation

For users without ImageJ, the first step is to download and install the ImageJ application from: https://imagej.nih.gov/ij/download.html. The next step is to download the EzColocalization plugin from: http://sites.imagej.net/EzColocalization/plugins/. When saving the file, the user should delete the timestamp at the end of the name of the EzColocalization file. For example, a version named “EzColocalization_.jar-20180716210728” should be renamed as “EzColocalization_.jar”. Once the plugin has been renamed “EzColocalization_.jar” it can be moved into the “plugins” folder of ImageJ to install it. Alternatively, users can install it by running ImageJ, selecting “Install…” from the “Plugins” menu of the menu bar, and then selecting the renamed file to install. To use EzColocalization, run the ImageJ application (open “ImageJ.exe” in the ImageJ folder) and choose “EzColocalization” from “Plugins” on the menu bar. For those using Fiji, the EzColocalization update site can be followed according to the instructions at https://imagej.net/Following_an_update_site.

RESULTS

Overview of EzColocalization workflow

The workflow for EzColocalization is divided into four modules each with its own tab on the GUI. The tabs are: (i) “Inputs” where images, masks or regions of interest (ROI) lists are selected and aligned; (ii) “Cell Filters” where cells can be selected based on physical features and signal intensity; (iii) “Visualization” where heat maps, scatterplots, and metric matrices (defined below) are created; and (iv) “Analysis” where the colocalization metrics and outputs are chosen. Not all modules and not all processes within a module have to be used. Some tabs have a “Preview” button to run a specific module instead of the “Analyze” button which runs all selected processes in all modules.

Inputs

Image files, which are chosen in the “Inputs” tab (Fig. 1A), must be: (i) monochromatic (*i.e*. not RGB or CMYK formats); (ii) 8-bit, 16-bit, or 32-bit; and (iii) in a format such as TIFF that retains the original pixel intensity values. Large images may be compressed for file transfer using a lossless format such as ZIP or LZW, and then decompressed for analyses. In addition to images, EzColocalization can accept masks and ROI lists for cell identification (see below). If there are multiple images for each channel, the images should be stacked for more efficient analysis in the “Stack” menu (see ImageJ guide for further details24). Images in a stack may be different fields of view or a time series, but must have the same dimensions, magnification and image order for each channel. The input tab also provides options for setting thresholds for signal intensity and aligning misaligned images from different channels (Fig. 1B and Supplementary Information). Recommendations for acquiring suitable images for colocalization analysis are provided in the Supplementary Information. Note: alignment operates on the assumption that an appropriate threshold for signal intensity can be chosen to distinguish pixels inside and outside of cells; if thresholding includes areas outside the cell or only a limited area within cells, then the alignment may not function properly. For this reason, all alignments should be checked by visually by examining the ROIs to confirm that appropriate cell areas are selected.

EzColocalization is primarily designed for one “cell identification” channel and two or three “reporter” channel images. However, it can operate with other input combinations (Table S1). The cell identification channel is used to identify individual cells, and consequently to distinguish intracellular and extracellular pixels. The cell identification channel can be any type of image that permits identification of the cell boundaries including: light microscopy images (*e.g*. phase contrast25,26 and bright-field), images with a reporter that labels the cell membrane or that is throughout the cytoplasm (*e.g*. Cy5, Fig. 1B), and images with an extracellular dye that outlines cells. Differential interference contrast (DIC) images create shadows that make it difficult for automated selection of cells using threshold methods27; therefore for DIC images we recommend that ROIs be created using the “selection tools” in ImageJ to manually outline cell areas, and then adding them to a list by choosing “Add to Manager” (in “Selection” submenu of the “Edit” menu). Once the ROIs for all cells of interest in an image are selected, a binary mask can be created using the “Clear Outside” and “Autothreshold” functions of ImageJ.

Cell Filters

The “Cell Filters” tab is used to help select cells in images (Fig. 2A) and distinguish intracellular and extracellular pixels. Cells are identified by: (i) choosing one of the ImageJ threshold algorithms24, or manually selecting the thresholds (which is done by selecting “*Manual*” from a drop-down list in the Inputs tab and pressing the “Show threshold(s)” button), to identify regions corresponding to cells in the cell identification channel (Fig. 2B); (ii) using watershed segmentation to separate touching objects in the cell identification channel images (optional) (Fig. 2B); (iii) selecting objects from the cell identification channel images based on physical parameters (Fig. 2C) and signal intensity (Fig. 2D). EzColocalization will attempt to automatically detect whether input images have dark or light background using skewness. Assuming there are more pixels in the background than in the cells, an image with positive skewness indicates a dark background and negative skewness indicates a light background. Users can also manually select whether the input images have dark or light background in the “Parameters…” options of the “Settings” menu. Cells that are only partly within an image, and therefore could provide misleading values, are automatically removed from analyses.

EzColocalization has one optional “Pre-watershed filter” and eight optional post-watershed filters (with the option to select more). Watershed segmentation can aid the separation of dividing and touching cells28 but it can also divide large objects such as aggregates of extracellular material into smaller fragments that are the same size as cells. To avoid the latter, the Pre-watershed filter can be used to exclude objects with large areas from the analysis. The Preview button in the Cell Filters tab allows users to see which objects on the current image will be selected when the minimum and maximum bounds of all the filters are adjusted. There are two classes of parameters for the post-watershed cell filters (Table S2): (i) physical parameters based on measurements from the cell identification channel; and (ii) signal intensity parameters from the reporter channels. Physical parameters apply to all channels whereas signal intensity parameters apply only to the reporter channel for which they are selected (because reporters may have very different levels of signal). In addition to filtering based on predefined options in ImageJ, EzColocalization has filters for the “MeanBgndRatio” or “MedianBgndRatio”, which are calculated by dividing the mean or median signal intensity of pixels inside an object by the respective mean or median signal intensity of extracellular pixels.

Visualization

The “Visualization” tab displays signals or metrics in cells as: (i) “heat maps”; (ii) scatterplots; and (iii) “metric matrices” (Fig. 3A).

Heat maps are pseudocolor images that show the relative magnitude of reporter signals (Fig. 3B). They are generated by normalization and rescaling so that the minimum and maximum pixel values are 0 and 255 respectively in each cell, image, or stack. There are eight options for coloring the heat maps, and the intensity values for each color are

obtained from the “Show LUT” function (within the “Color” submenu of the “Image” menu in ImageJ). Cell heat maps are suited for determining where each reporter occurs with highest intensity in cells. Image heat maps can show if different cells within a field of view have substantially different intensities, which may indicate biological heterogeneity or unevenness in labeling. Stack heat maps can show if cells in different images have substantially different levels of signal intensity, which may indicate unevenness in labeling or measurements across a slide (*e.g*. due to photobleaching) or changes in signal over time (if the stack is a time series). Note: heat map appearances are affected by brightness and contrast settings.

Scatterplots show the relationship between the signal intensity for two or three reporter channels for individual cells and images (Fig. 3C). This relationship is important in choosing the appropriate colocalization metric (Supplementary Information). Scatterplots can also reveal heterogeneity in the localization patterns8, which may require removal of background pixels or separate analyses for different cell types.

Metric matrices provide an overview of localization patterns by showing the calculated values of a colocalization metric for many threshold combinations. Metric matrices for the threshold overlap score (TOS) have been shown to be useful for the analysis of localization patterns for two reporter channels8,15 (Fig. 3D). For completeness, EzColocalization has the option to calculate metric matrices for two reporter channels using five other metrics: threshold overlap score with logarithmic scaling8, Pearson correlation coefficient (PCC), Manders’ colocalization coefficients (M1 and M2), Spearman’s rank correlation coefficient (SRCC), and intensity correlation quotient (ICQ)8,15. Colocalization for three channels can also be measured using ICQ, Manders’ colocalization coefficients and TOS29(Supplementary Information).

Thresholds for all metrics are measured as the top percentile (F_T) of pixels for signal intensity8,15. For example, $F_T = 0.1$ is the 10% of pixels with the highest signal. For the metric matrices, F_T is also used to specify the step size for the threshold combinations. That is, $F_T = 0.1$ also selects thresholds for the 10%, 20%, …, and 100% of pixels with the highest signal. If F_T does not divide evenly into 100, then the remaining percent is the last step size. For metrics that do not need a threshold (*i.e*. PCC, SRCC, and ICQ) the values are calculated assuming that only the pixels above the thresholds exist. The metric matrix window has options for the results to be saved as text or image, for changing the F_T or type of metric, viewing individual cell metric values as a list, and calculating the mean, median or mode of the metric for each threshold combination. The “Proc” (processed) and “Raw” button determines whether the list of data displayed, copied, or saved with the “List”, “Copy”, or “Save…” buttons respectively is the average value for the sample for each threshold combination (*e.g*. median value) or all values for each cell in the sample for all threshold combinations.

Analysis

The “Analysis” tab has three subtabs (“Analysis Metrics”, “Metrics Info” and “Custom”). The Analysis Metrics subtab has six metrics for measuring colocalization for two reporters (Fig. 4A) and three metrics for three reporters (see previous section). Users may choose a threshold or no threshold for PCC, SRCC and ICQ. TOS and Manders’ colocalization coefficients must have a threshold to be calculated. The Metrics Info subtab contains information and resources about the metrics used in the Analysis Metrics subtab (more details in the Supplementary Information). Thresholds can be selected using Costes’ method30 or manually. In the Custom subtab (see Supplementary Information for additional information), users can write their own code in Java™ to analyze images (note: the example provided is for calculating PCC) (Fig. 4B). The “Compile” button tests the code and creates a temporary file in the Java temporary directory and displays the outcome of the compiling with a “Succeeded” or “Failed” label. If successful, the compiled custom code is read to the memory again and applied to the selected cells.

The output of every analysis is a table that specifies the image and an identifier number for every cell (Fig. 4C), and for each cell, values are provided for: (i) the selected metric; (ii) physical parameters; and (iii) average signal intensity for each channel (if selected). Note: “NaN” in the output table indicates the failure to calculate a value. Users can also generate summary windows (with the cell number, mean, median and standard deviation for the selected metric) (Fig. 4D), histograms of metric values (Fig. 4E), binary mask images, and a list of ROIs that represent each cell’s position and number on each image in the ROI manager. ROI lists and binary mask images can be saved for re-analysis of the same cells.

Applications of EzColocalization

EzColocalization is designed to be used in a modular manner to facilitate customization of analyses for a wide variety of experiments and researcher needs. This section focuses on demonstrating specific tools in EzColocalization to solve real-world problems in diverse image sets.

In the first application of EzColocalization, images of rat hippocampal neurons from the Cell Image Library (CIL:8773, 8775–8788, which are attributed to Dieter Brandner and Ginger Withers) are used to demonstrate: (i) the use of a reporter channel for cell identification when an experiment does not have separate non-reporter images for cell identification; (ii) cell filters for selecting cells; and (iii) visualization tools for choosing metrics. The workflow of the analysis is outlined in Fig. 5A. In the first step, two reporter image stacks were created: one stack with images where F-actin is labelled (using a phalloidin peptide conjugated to rhodamine); and the second stack with images where tubulin is labelled (using an antibody conjugated to Alexa 488) (Fig. 5B). The interaction of F-actin and tubulin is important for the growth and migration of neurons31,32. We used the F-actin images for cell identification because it is present in all cells and it shows the cell boundaries8. Individual cells were selected from the F-actin images by

applying a threshold to identify cells24 and using a cell filter to remove cell debris (note: parameter values in Fig. 5A).

After the cells were selected, the intensity of reporter signals were examined using cellular heat maps and scatterplots. We found the reporters did not colocalize at high signal levels and there was a complex relationship between the signal intensities (Fig. 5C,D). Due to the latter, localization was quantified using Manders' M1 and M2 and TOS (Supplementary Information). M1 and M2 were evaluated at thresholds selected by Costes' method for the cell outlined in Fig. 5B, and the values were 0.289 and 0.995 respectively. These values are usually interpreted as indicating that tubulin has high colocalization with F-actin, and F-actin has low colocalization with tubulin. TOS values were evaluated by generating a metric matrix with median TOS values. The matrix showed colocalization, anticolocalization and noncolocalization at different thresholds for the signal intensities of tubulin and F-actin (Fig. 5E). At sites in cells where F-actin and tubulin have the highest intensity signal (top 10% of pixels for each channel), the median TOS value is −0.36 ($n = 20$). This negative value indicates anticolocalization, which is consistent with the impression obtained from the heat maps and scatterplots, and with other reports8.

In the second application, images of *Saccharomyces cerevisiae* undergoing mitosis were obtained from the Cell Image Library33 to demonstrate: (i) cell identification via hand-drawn outlines (for experiments where automated methods of cell identification cannot be applied); and (ii) image alignment. The reporter inputs were an image from a wild type strain ("control"; CIL: 13871) that has the BFA1 protein that loads TEM1 onto the spindle pole body, and an image from a strain without the BFA1 protein (Δ*bfa1* deletion mutant; CIL: 13870). In these reporter images, cells expressed TEM1 protein fused to GFP and the DNA was labelled with DAPI (4′, 6-diamidino-2-phenylindole). TEM1 localizes to spindle pole bodies during mitosis and is implicated in triggering exit from mitosis33. The workflow is shown in Fig. 6A. In this application, ROIs were manually drawn around cells using the "Freehand" selection tool in ImageJ on DIC images. Binary masks, which were used to select cell areas, were created by selecting the ROIs and using the "Clear Outside" and then "Auto Threshold" functions of ImageJ24 (Fig. 6B). The cell areas were used for cell identification and to correct alignment between the DIC images and the reporter channels using the "Default" threshold algorithm (Fig. 6C). Following this cell identification and image alignment, the images are now ready for visualization and analysis as described in the previous example.

In the third application, images of whole adult *Caenorhabditis elegans*obtained from the Broad Bioimage Benchmark Collection (BBBC012v1, M14)34 were used to demonstrate that: (i) EzColocalization can analyze colocalization in whole organisms; and (ii) "cell" filters can select individual organisms based on reporter signal intensity. The images in this example are from the same dataset used in our study describing TOS (but they are not the same images)8. The workflow is shown in Fig. 7A. Outlines of individual *C. elegans* were drawn in ImageJ on bright-field images to create ROIs, and the ROIs

were added to the ROI manager for “cell” identification. GFP expressed from the *clec-60* promoter in the anterior intestine was reporter 1 and mCherry expressed from the *myo-2* promoter within the pharynx, which is an organ next to the anterior intestine35, was reporter 2. Cell filters for physical parameters were unnecessary because only those objects considered to be suitable *C. elegans* had outlines drawn around them in the first place. However, cell filters for signal intensity were necessary because some *C. elegans* had low GFP signal, possibly due to transgene silencing36,37 (Fig. 7B). Subsequent visualization and analysis can be performed as described in the first application.

In the fourth application, we demonstrate the analysis of colocalization for three reporter channels. The workflow was the same as for two reporter channels except “3 reporter channels” was first selected in the “Settings” main menu (Fig. 8A). Images were obtained from the Broad Bioimage Benchmark Collection (BBBC025, Version 1, Image set: 37983, image: p23_s9) of U2OS bone cancer cells ($n = 66$)38. The three reporter images had DNA, endoplasmic reticulum (ER) and mitochondria respectively stained with Hoechst 33342, concanavalin A/Alexa Fluor488 conjugate, and MitoTracker Deep Red (upper row, Fig. 8B). Cell identification was performed with an image of the plasma membrane labeled with wheat germ agglutinin (WGA)/Alexa Fluor 555 conjugate (upper left, Fig. 8B). Note: the image also had the Golgi apparatus and F-actin network labeled38. The plasma membrane was traced using the polygon selection tool in ImageJ to create ROIs for the individual cells, and the ROI manager containing the ROIs was selected for cell identification.

The localization patterns were visualized in the same manner as for two reporters except that: (i) there are three sets of heat maps for the reporters instead of two (lower row, Fig. 8B); and (ii) scatterplots and metric matrices are in three dimensions (Fig. 8C–F). There is the option in the Visualization tab and the Analysis tab (Fig. 8G) to measure colocalization for the three reporters using ICQ, TOS or Manders’ M1, M2 and M3 metrics. Of the three metrics, we found that TOS was the easiest to interpret. TOS has a single value for measuring the colocalization of all three reporter signals, and it clearly showed the reporter signals for the nucleus, mitochondria and ER overlapped at low thresholds (*i.e.* at high F_T values there is colocalization; red color in Fig. 8E) and did not overlap at high thresholds (*i.e.* at low $F_T$values there is anticolocalization; blue color in Fig. 8E). These observations are consistent with the nucleus, mitochondria and ER organelles overlapping at their edges (where the signal from their reporters is typically lower) due to known physical interactions, but not at their centers (where the signal from their reporters is typically higher) because they are distinct structures in cells39,40,41.

DISCUSSION

EzColocalization was designed to make it easier for researchers to determine where particular types of molecules occur in cells and organisms in relation to other types of molecules. In addition, EzColocalization can provide data on colocalization for each cell

or organism in a sample, which is increasingly recognized as being crucial to understanding biological processes such as cell differentiation[42], cancer[43], and microbial pathogenesis[44].

Two of the most widely used applications for colocalization analysis are JACoP and Coloc2[10,12]. JACoP is an ImageJ plugin that can generate pixel intensity scatterplots to visualize localization patterns and measure colocalization with a variety of metrics including PCC (Van Steensel's CCF method or Costes' randomization), Manders' M1 and M2, ICQ, and object based methods[10]. It also permits thresholds to be chosen manually or automatically using Costes' method[10]. Coloc 2 is a plugin for Fiji[12], which builds on the functionality of JACoP by providing options to: analyze selected ROIs within single images, threshold images using a "bisection" algorithm, and measure colocalization with SRCC and Kendall's Tau rank correlation. Unfortunately, JACoP and Coloc 2 do not have built-in options to automate analyses or perform separate colocalization measurements for multiple objects in an image, therefore analyses can be challenging for images with a lot of background pixels or different cell types. The Wright Cell Imaging Facility (WCIF) has helped address these challenges by creating a colocalization plugin that can measure colocalization for individual cells by manually creating individual ROIs[11], but this method cannot be easily automated to analyze many cells across many images.

In addition to the above, software has been reported for measuring colocalization in cells, particularly in cases where the signal is defined to distinct regions or foci. One of these applications is MatCol, which can identify overlapping objects after a threshold is applied, and then calculate if the measured overlap is significantly different to that expected if the same objects were randomly scattered[45]. Another reported script calculates object based colocalization in confocal images[46] from the percent overlap of the objects. A third program measures colocalization for three-dimensional images; it measures the proportion of thresholded objects in one channel that have their center of mass within thresholded objects of another channel[47]. There are practical barriers to the widespread use of these three programs including the need for additional software to identify cell areas and that they are written in Matlab or C++ (therefore users must be familiar with these programming languages to customize them).

To make it easy to optimize analyses, EzColocalization has a simple GUI that requires no programming experience unless a custom metric is created. The GUI template is based on one that is familiar to many microscopists. ImageJ also has a large library of tools that can be used with EzColocalization, and it is open source software[24]. ImageJ has options for creating stacks of images and thresholding images, which were incorporated into EzColocalization for automated analyses. EzColocalization also has tools for the input of images, cell identification, visualizing localization patterns, measuring colocalization, and for displaying and saving results.

EzColocalization can select individual cells from cell identification images using thresholds, ROIs, or binary mask images. Identification of individual cells allows pixels within cells to be discriminated from pixels in the background and non-cell objects. In addition, cell filters can limit analyses to a subset of cells with certain physical parameters and minimum signal levels. Filters are used to select cells instead of more advanced techniques for cell detection[48] because: (i) they do not require assumptions about cell features (therefore diverse cell types can be analyzed); and (ii) they are intuitive, which makes it easier for researchers to tailor settings for their experiments and identify if patterns of localization are associated with specific cell features.

The visualization tools (heat maps, scatterplots, and metric matrices) can help with choosing the appropriate metrics and thresholds for the analyses. The metric matrices are particularly useful for samples with non-specific binding or localization of probes. These matrices display colocalization values for multiple combinations of thresholds for signal intensity, which facilitates the selection of thresholds so the analysis includes pixels from cellular regions with high signal (due to specific localization) and excludes pixels from regions with low signal (due to non-specific localization).

EzColocalization can not only measure colocalization for two reporters but also for three reporters. The latter is a useful feature that is unavailable for most software applications for measuring colocalization. In addition, custom metrics can be programmed in EzColocalization.

The data table generated by the colocalization analysis is an important feature of EzColocalization. Because the value of the colocalization metric for each cell is provided, and not just the average measurement of colocalization for the sample, it is possible to examine the distribution of metric values, perform statistical analyses, calculate receiver operating characteristic curves, and analyze subsets of cells in heterogeneous samples[8]. The data table also lists the specific image and a unique identifying number for each cell, therefore researchers can examine the images to determine why different cells have different measurements. The data tables can be downloaded and used in any spreadsheet application, which makes the data accessible to researchers without programming experience. Furthermore, the values for the physical parameters, signal intensity, and colocalization metrics can be retrieved from the tables (if the check box is selected) for more sophisticated multivariate analyses, including clustering, classifying and ordination methods.

In conclusion, EzColocalization is an ImageJ plugin with a user-friendly GUI, tools for start-to-finish analysis of colocalization, and many options to customize analyses. The tools are provided to select specific types of cells or organisms, visualize and measure colocalization, and automate analyses. The analysis generates a data table with measurements of colocalization, signal intensity and physical parameters for each cell, which allow users to delve deep into their data. Together these features make

EzColocalization ideal for researchers at all levels, and for analyzing heterogeneous samples and complex patterns of localization.

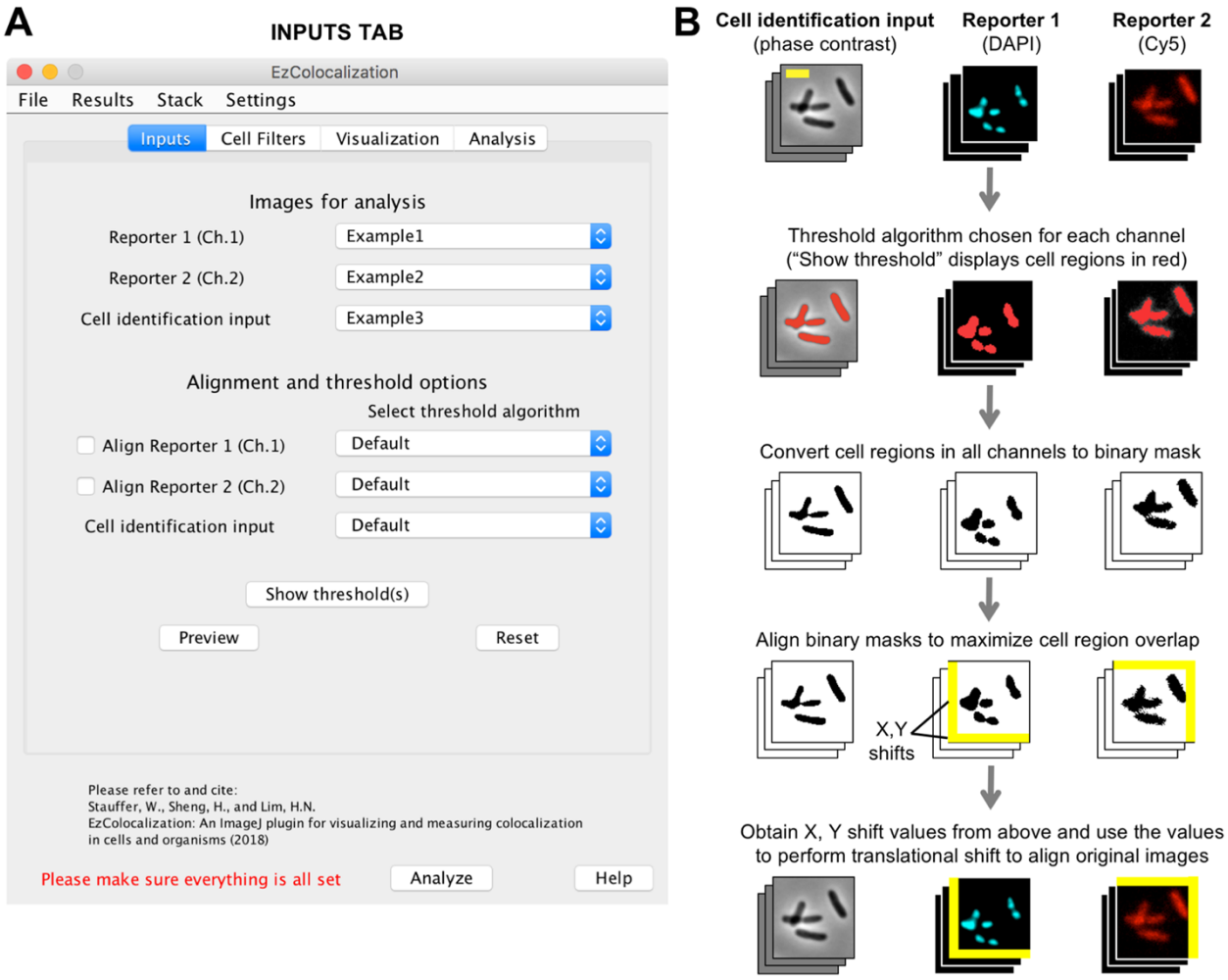

Figure 1: Inputs and alignment tab. (**A**). Inputs tab in the GUI. (**B**) General steps for the alignment of images. The cell identification image stack (phase contrast; left column), reporter 1 image stack (DAPI staining of DNA; center column), and reporter 2 image stack (Cy5; right column) are images of a previously reported bacterial strain (HL6320)15. Scale bar is 2 µm. Reporters 1 and 2 images are pseudocolored. Red coloring in the second row of images indicates the objects identified by thresholding of the signal in each channel ("Default" algorithm in ImageJ). Following alignment of the images, pixels that overhang are removed and gaps are filled with pixels with zero value (yellow areas) so that all images have the same area in the common aligned region.

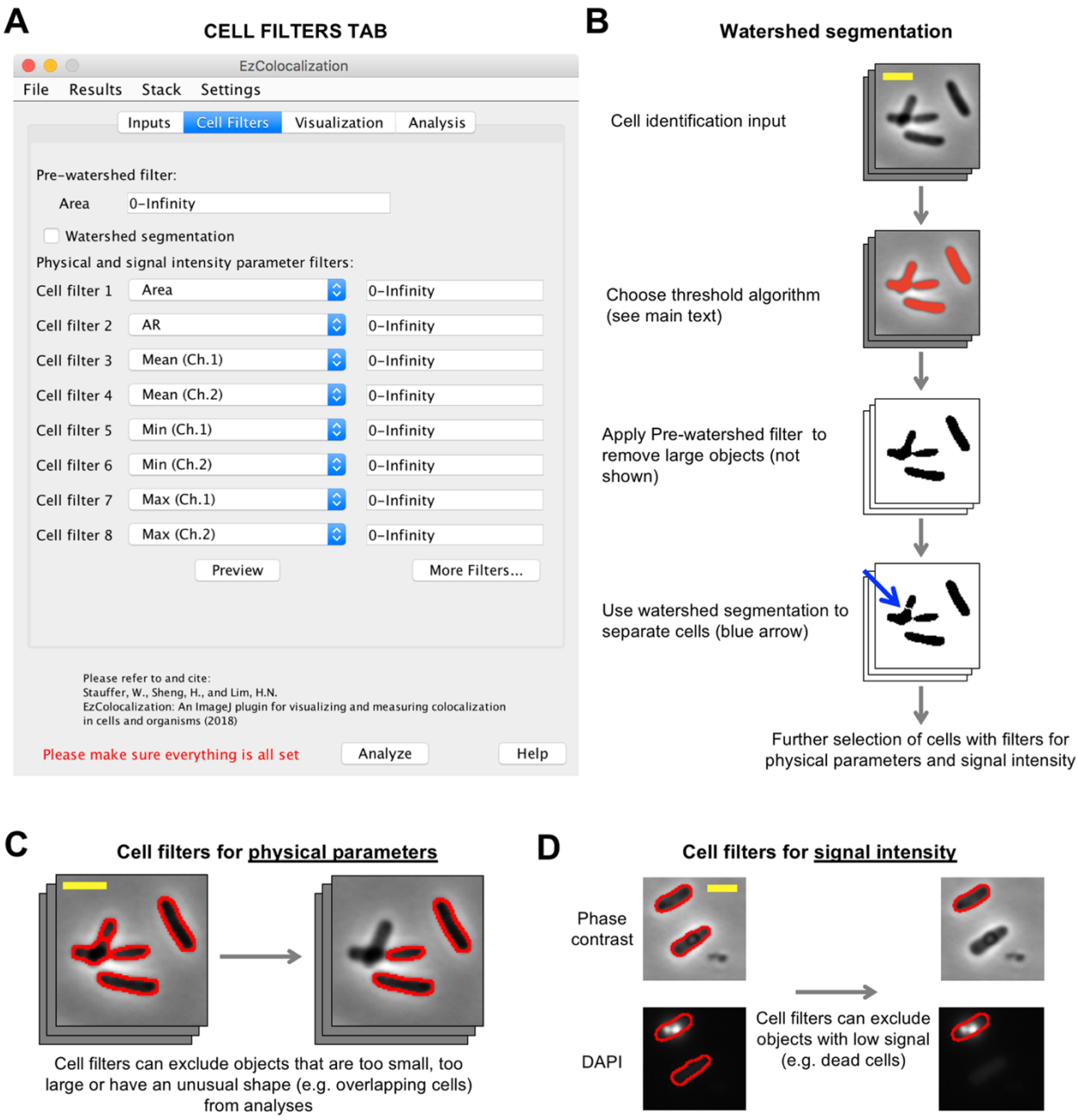

Figure 2: Cell identification and cell filters tab. (**A**) Cell Filters tab in the GUI. (**B**) Cell selection and watershed segmentation. Red coloring in the image in the second row indicates objects identified by thresholding of the signal in the cell identification channel ("Default" algorithm in ImageJ). Cells are the same as in Fig. 1. (**C**) Selection of cells based on physical features using the cell filters. Scale bar is 2 µm. Phase contrast image from Fig. 1. Red outline indicates the objects that were identified by thresholding (Panel B), and in the case of the right image, are within the parameter range(s) selected by the filter. (**D**) Selection of cells based on signal intensity using the cell filters. Phase contrast (cell identification image) and DAPI stain (reporter channel) images of bacteria (HL6187). Scale bar is 2 µm. Note: the lower of the two cells (no red border) has been removed from the analysis by the cell filter (that is, it no longer has the red cell outline).

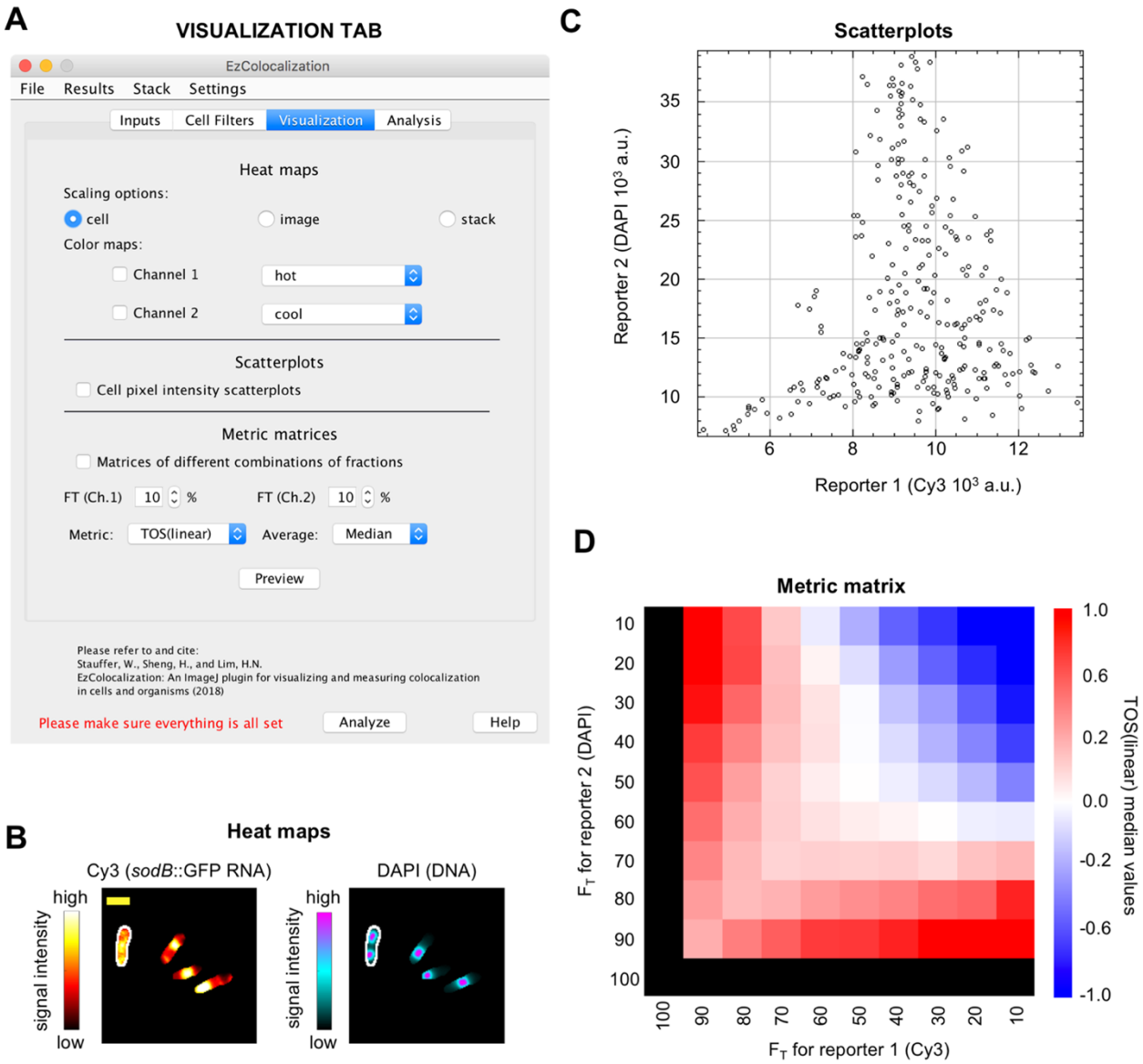

Figure 3: Visualization tab. Data are from bacteria (HL6187) with labeled *sodB::gfp* RNA (Cy3 channel) and DNA (DAPI). (**A**) Visualization tab in the GUI. (**B**) Heat maps of Cy3 and DAPI signals for bacteria with "cell scaling" (defined in main text). Scale bar is 2 μm. (**C**) Scatterplot of Cy3 and DAPI for the cell on the left and outlined in white in Fig. 3B. (**D**) Metric matrix for TOS (linear scaling) for the cell on the left and outlined in white in Fig. 3B. F_T is the top percentage of pixels in the channel; for example, if F_T for Cy3 is 80% then it refers to the 80% of pixels with the highest Cy3 signal. Black color on the left column and bottom row indicate that TOS values are not informative when one threshold is 100%; that is, the overlap of two reporters can only be 100% if 100% of pixels are selected for at least one channel.

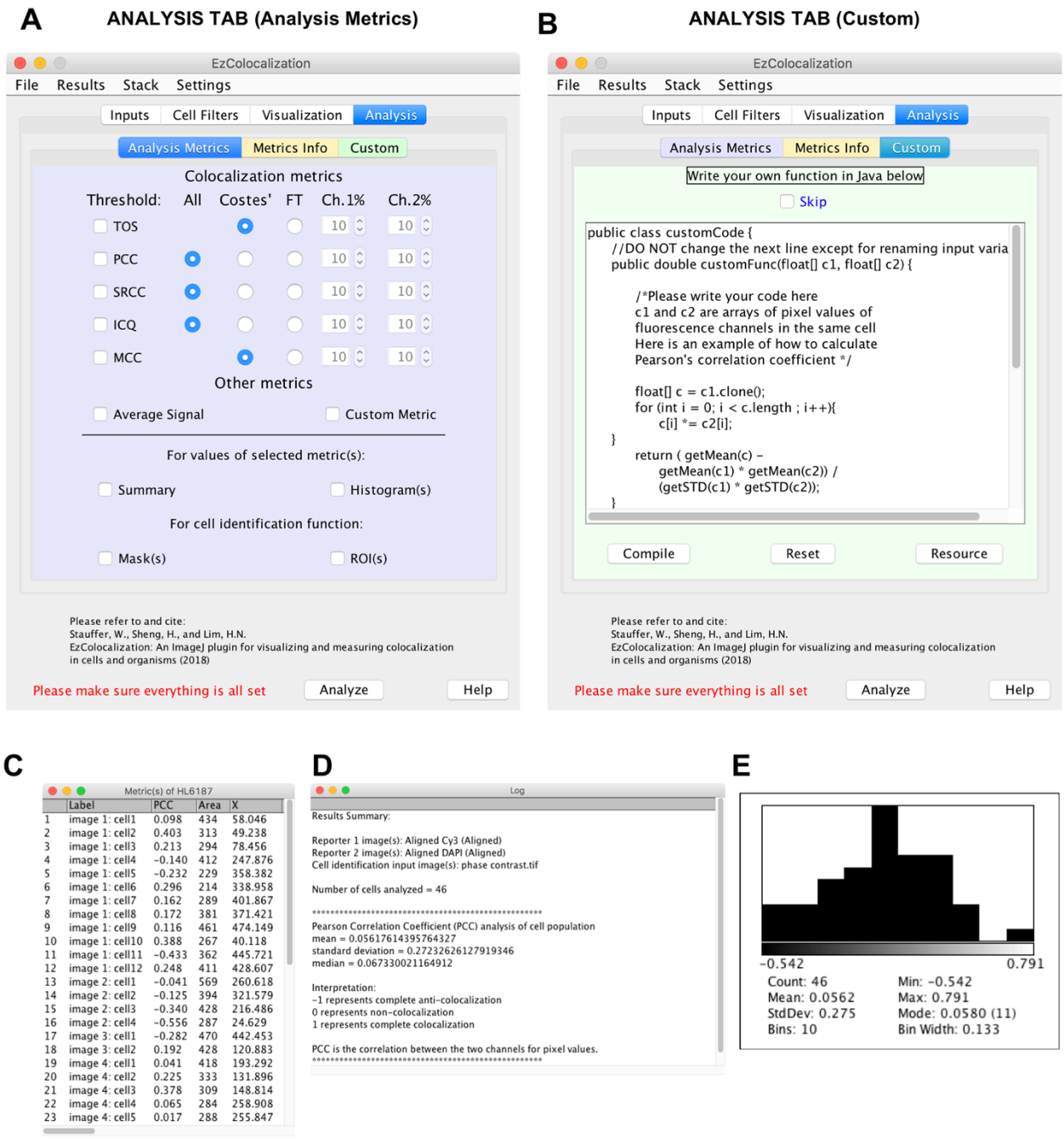

Figure 4: Analysis tab. (**A**) Analysis tab in the GUI for selecting default metrics. Note: this example is for two reporter channels (see Fig. 8G for 3 reporter channels). (**B**) Analysis tab in the GUI for users to code custom metrics. The example code provided is for measuring colocalization by Pearson correlation coefficient. (**C**) Example of a data table showing metric values for Pearson correlation coefficient (PCC) and some of the parameter values for some of the cells in the analysis. Label = the image and unique cell number to identify individual cells; Area = area of each cell in pixels; and X = the average x-value of all pixels in a cell. Data is from the example used in Fig. 3. (**D**) Summary report ("Log") of the results in Fig. 4C. (**E**) Histogram generated from the results in Fig. 4C. The height of each bin is the relative frequency. The Count is the

number of cells. Mean is the mean value. StdDev is the standard deviation. Bins is the number of bins. Min and Max are the minimum and maximum values of the lowest and highest bin respectively (which are shown immediately under the histogram). Mode is the mode value. Bin Width is the width of each bin within the histogram.

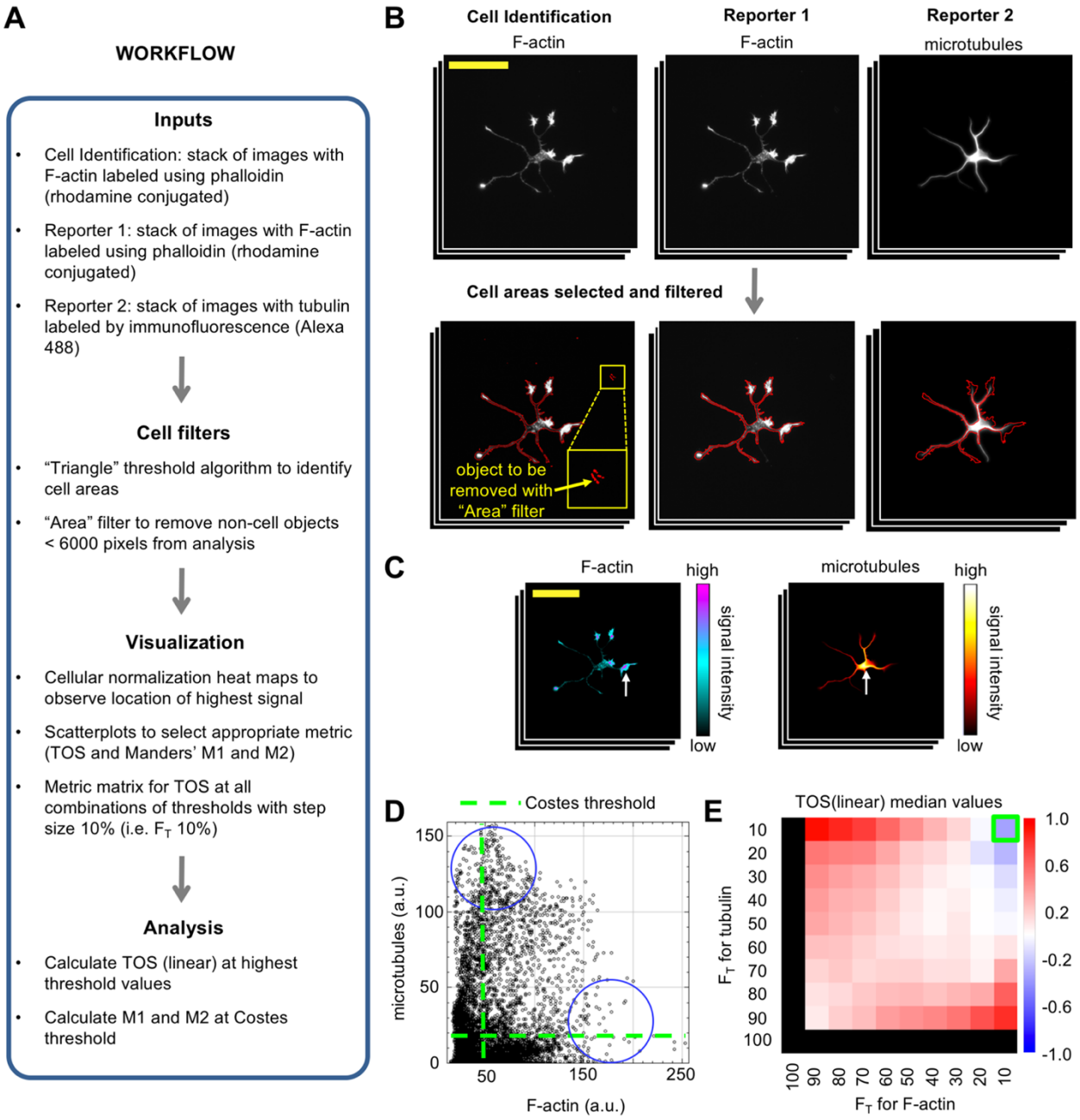

Figure 5: Application 1: Cell selection using reporter images and physical parameters. Images are rat hippocampal neurons labelled with an F-actin probe and anti-tubulin antibody visualized by fluorescence microscopy (see main text). (**A**) Workflow of the analysis. (**B**) Cell identification using the F-actin reporter and filters to remove small non-cell objects (yellow arrow) based on their size (*i.e.* Area option from the cell filters). Large yellow box in left panel is a zoomed in view of the smaller yellow box. Red outline

of the neuron indicates it has been identified as an object (*i.e*. a cell) for analysis. Scale bar is 100 µm. (**C**) Heat maps with cellular normalization showing localization regions of signal intensity for the cell shown in panel B. Scale bar is the same as panel B. (**D**) Scatterplot showing relationship between the signal intensity for two reporter channels for a random cell in the sample. Pixels with the highest intensity signal for each reporter channel have the lowest intensity signals for the other reporter, which indicates anticolocalization (blue circles). Green dash lines indicate thresholds selected by Costes' method. (**E**) Metric matrix for the median TOS (linear) value for all cells in the sample ($n = 20$). Green box indicates the threshold combination where F-actin and tubulin have the highest intensity signal (top 10% of pixels for each channel); the median TOS value is −0.36.

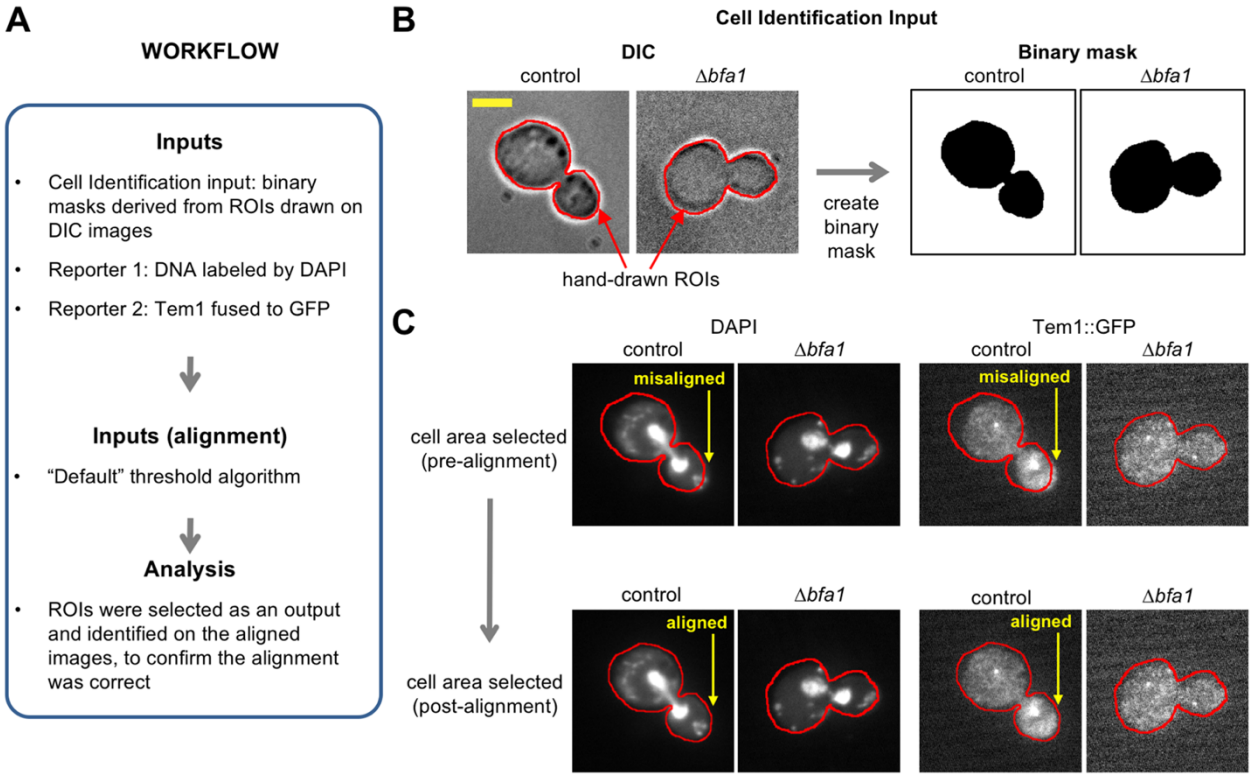

Figure 6: Application 2: Image alignment. Images are *S. cerevisiae* with TEM1 translationally fused to GFP and DAPI staining visualized by DIC microscopy and fluorescence microscopy (see main text). (**A**) Workflow of the analysis. (**B**) Cell identification by hand-drawn ROIs on a DIC image and creation of a binary image mask. Red outline indicates the boundary of the hand-drawn ROI. Scale bar is 3.5 µm. (**C**) Alignment of the reporter images using the binary mask image. Arrows indicate areas of misalignment that are corrected. Red outline is the same as for Panel B.

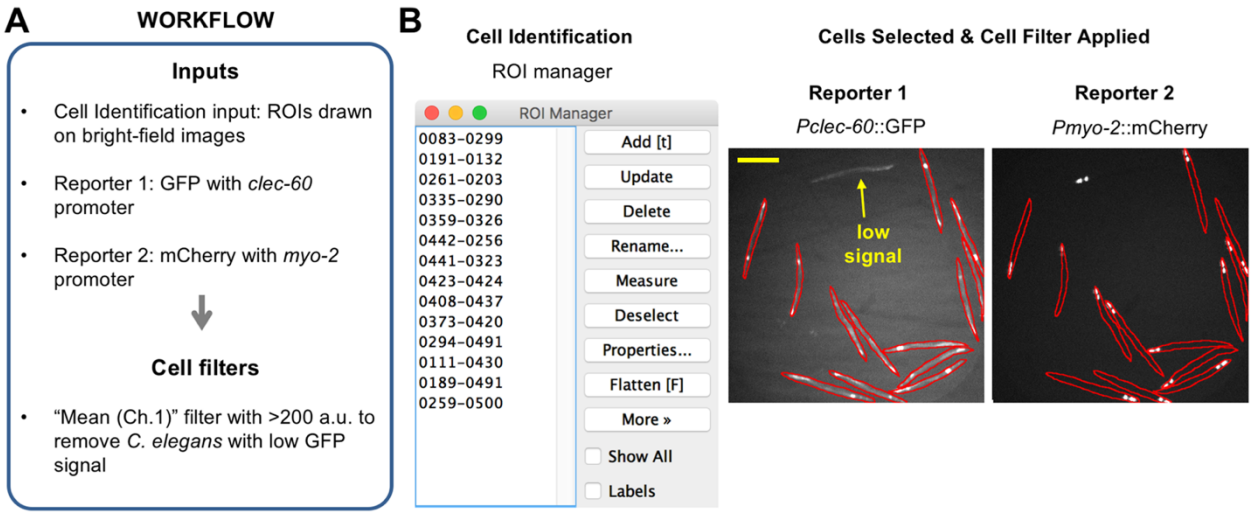

Figure 7: Application 3: Cell selection using signal intensity parameters. Images are whole adult *C. elegans* with GFP expressed from the *clec-60* promoter and mCherry expressed from the *myo-2* promoter that are visualized by bright-field microscopy and fluorescence microscopy (see main text). (**A**) Workflow of the analysis. (**B**) Selection of *C. elegans* so that only those individuals with an average intensity for the reporter signal that is above a threshold level are included in analyses. Left image shows the ROI manager with a list of ROIs that were hand-drawn around each *C. elegans*. Right image shows the reporter channel images with red outlines indicating the boundaries of the ROIs. *C. elegans* below the threshold level were excluded (yellow arrow) from the analyses by using the cell filters for signal intensity. Scale bar is 250 μm.

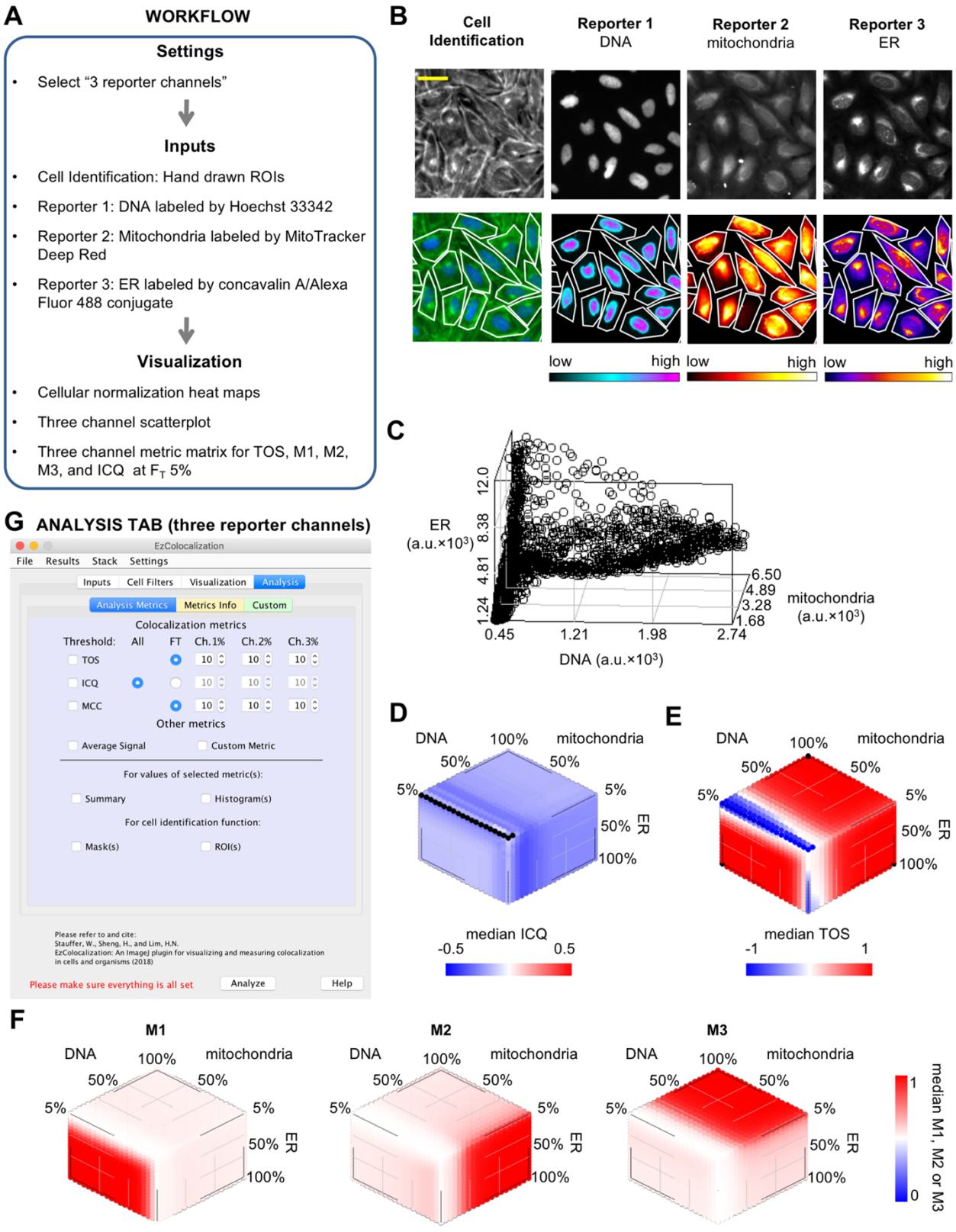

Figure 8: Application 4: Measurement of colocalization for three reporter channels. Images are of human bone cancer cells (U2OS) labelled as described in the main text. (**A**) Workflow of the analysis. (**B**) Images of cells in the cell identification and reporter

channels. Top row are raw images. Bottom row, left image is the cell identification with pseudocolor (blue is the signal from Hoechst 33342 signal and green is the signal from phalloidin/Alexa Fluor 568 conjugate and wheat germ agglutinin/Alexa Fluor 555 conjugate) and boundaries of the ROIs in white (see main text). Bottom row (except left image) are heat maps for each of the three reporters with the boundaries of the ROIs shown. Signal intensity is indicated by the bar below each reporter image. Scale bar is 20 μm. (**C**) A three channel scatterplot for a single cell is shown for illustrative purposes only. (**D–F**) Metric matrices of median values for ICQ (**D**) TOS (**E**) and Manders' colocalization coefficients M1, M2 and M3 (**F**) for all cells in the analysis ($n = 66$). Note: black color on metric matrix for ICQ indicates there were no pixels above all three thresholds for some cells, and therefore ICQ could not be calculated. (**G**) Analysis Metrics subtab for the Analysis tab for three reporter channels.

REFERENCES

1 Kannaiah, S. & Amster-Choder, O. Methods for studying RNA localization in bacteria. *Methods* (2015).

2 Kocaoglu, O. & Carlson, E. E. Progress and prospects for small-molecule probes of bacterial imaging. *Nat Chem Biol* **12**, 472-478, doi:10.1038/nchembio.2109 (2016).

3 Xue, L., Karpenko, I. A., Hiblot, J. & Johnsson, K. Imaging and manipulating proteins in live cells through covalent labeling. *Nat Chem Biol* **11**, 917-923, doi:10.1038/nchembio.1959 (2015).

4 Gautam, S., Gniadek, T. J., Kim, T. & Spiegel, D. A. Exterior design: strategies for redecorating the bacterial surface with small molecules. *Trends Biotechnol* **31**, 258-267, doi:10.1016/j.tibtech.2013.01.012 (2013).

5 Gruskos, J. J., Zhang, G. & Buccella, D. Visualizing Compartmentalized Cellular Mg2+ on Demand with Small-Molecule Fluorescent Sensors. *J Am Chem Soc* **138**, 14639-14649, doi:10.1021/jacs.6b07927 (2016).

6 Perry, J. L., Ramachandran, N. K., Utama, B. & Hyser, J. M. Use of genetically-encoded calcium indicators for live cell calcium imaging and localization in virus-infected cells. *Methods* **90**, 28-38, doi:10.1016/j.ymeth.2015.09.004 (2015).

7 Kervrann, C., Sorzano, C. Ó. S., Acton, S. T., Olivo-Marin, J. C. & Unser, M. A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy. *IEEE Journal of Selected Topics in Signal Processing* **10**, 6-30 (2016).

8 Sheng, H., Stauffer, W. & Lim, H. N. Systematic and general method for quantifying localization in microscopy images. *Biol Open* **5**, 1882-1893, doi:10.1242/bio.019893 (2016).

9 Adler, J. & Parmryd, I. Quantifying colocalization: thresholding, void voxels and the H(coef). *PLoS One* **9**, e111983 (2014).

10 Bolte, S. & Cordelieres, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* **224**, 213-232 (2006).

11 Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *Am J Physiol Cell Physiol* **300**, C723-742 (2011).
12 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
13 Yao, Z. & Carballido-López, R. Fluorescence imaging for bacterial cell biology: from localization to dynamics, from ensembles to single molecules. *Annual review of microbiology* **68**, 459-476 (2014).
14 Haas, B. L., Matson, J. S., DiRita, V. J. & Biteen, J. S. Imaging live cells at the nanometer-scale with single-molecule microscopy: obstacles and achievements in experiment optimization for microbiology. *Molecules* **19**, 12116-12149 (2014).
15 Sheng, H., Stauffer, W. T., Hussein, R., Lin, C. & Lim, H. N. Nucleoid and cytoplasmic localization of small RNAs in Escherichia coli. *Nucleic Acids Res*, doi:10.1093/nar/gkx023 (2017).
16 Snapp, E. Design and use of fluorescent fusion proteins in cell biology. *Curr Protoc Cell Biol* **Chapter 21**, Unit 21 24, doi:10.1002/0471143030.cb2104s27 (2005).
17 Wallner, G., Amann, R. & Beisker, W. Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* **14**, 136-143, doi:10.1002/cyto.990140205 (1993).
18 Patterson, G. H., Knobel, S. M., Sharif, W. D., Kain, S. R. & Piston, D. W. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys J* **73**, 2782-2790, doi:10.1016/S0006-3495(97)78307-3 (1997).
19 Li, B. & You, L. Predictive power of cell-to-cell variability. *Quantitative Biology* **1**, 131-139 (2013).
20 Wiegand, W. Eclipse: A platform for integrating development tools. *IBM Systems Journal* **43**, 371-383 (2004).
21 Réveillac, J.-M. *Modeling and Simulation of Logistics Flows 2: Dashboards, Traffic Planning and Management*. (John Wiley & Sons, Inc., 2017).
22 Clayberg, E. & Rubel, D. *Eclipse Plug-ins*. Third edn, (Addison-Wesley, 2008).
23 Hussein, R. & Lim, H. N. Disruption of small RNA signaling caused by competition for Hfq. *Proc Natl Acad Sci U S A* **108**, 1110-1115, doi:10.1073/pnas.1010082108 (2011).
24 Ferreira, T. & Rasband, W. ImageJ User Guide — IJ 1.46. (2012). <imagej.nih.gov/ij/docs/guide>.
25 Zernike, F. Phase contrast, a new method for the microscopic observation of transparent objects. *Physica* **9**, 686-698 (1942).
26 Zernike, F. Phase contrast, a new method for the microscopic observation of transparent objects part II. *Physica* **9**, 937-1019 (1942).
27 Obara, B., Roberts, M. A., Armitage, J. P. & Grau, V. Bacterial cell identification in differential interference contrast microscopy images. *BMC Bioinformatics* **14**, 134, doi:10.1186/1471-2105-14-134 (2013).
28 Vincent, L. & Soille, P. Watersheds in Digital Spaces - an Efficient Algorithm Based on Immersion Simulations. *Ieee Transactions on Pattern Analysis and Machine Intelligence* **13**, 583-598 (1991).

29 Manders, E. M. M., Verbeek, F. J. & Aten, J. A. Measurement of colocalization of objects in dual-colour confocal images. *J Microsc* **169**, 375-382 (1993).
30 Costes, S. V. *et al.* Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys J* **86**, 3993-4003 (2004).
31 Brandner, D. a. W., G. *Rattus, multipolar neuron* (The Cell Image Library, 2010).
32 Rodriguez, O. C. *et al.* Conserved microtubule-actin interactions in cell movement and morphogenesis. *Nat Cell Biol* **5**, 599-609, doi:10.1038/ncb0703-599 (2003).
33 Coles, C. H. & Bradke, F. Coordinating neuronal actin-microtubule dynamics. *Curr Biol* **25**, R677-691, doi:10.1016/j.cub.2015.06.020 (2015).
34 Valerio-Santiago, M. & Monje-Casas, F. Tem1 localization to the spindle pole bodies is essential for mitotic exit and impairs spindle checkpoint function. *J Cell Biol* **192**, 599-614, doi:10.1083/jcb.201007044 (2011).
35 Ljosa, V., Sokolnicki, K. L. & Carpenter, A. E. Annotated high-throughput microscopy image sets for validation. *Nat Methods* **9**, 637 (2012).
36 Wahlby, C. *et al.* An image analysis toolbox for high-throughput C. elegans assays. *Nat Methods* **9**, 714-716 (2012).
37 Kelly, W. G., Xu, S., Montgomery, M. K. & Fire, A. Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene. *Genetics* **146**, 227-238 (1997).
38 Hsieh, J. *et al.* The RING finger/B-box factor TAM-1 and a retinoblastoma-like protein LIN-35 modulate context-dependent gene silencing in Caenorhabditis elegans. *Genes Dev* **13**, 2958-2970 (1999).
39 Bray, M.-A. *et al.* Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nature protocols* **11**, 1757-1774, doi:10.1038/nprot.2016.105 (2016).
40 English, A. R. & Voeltz, G. K. Endoplasmic Reticulum Structure and Interconnections with Other Organelles. *Cold Spring Harbor Perspectives in Biology* **5**, a013227, doi:10.1101/cshperspect.a013227 (2013).
41 Marchi, S., Patergnani, S. & Pinton, P. The endoplasmic reticulum–mitochondria connection: One touch, multiple functions. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1837**, 461-469, doi:https://doi.org/10.1016/j.bbabio.2013.10.015 (2014).
42 Prachar, J. Intimate contacts of mitochondria with nuclear envelope as a potential energy gateway for nucleo-cytoplasmic mRNA transport. *General physiology and biophysics* **22**, 525-534 (2003).
43 Potente, M. & Makinen, T. Vascular heterogeneity and specialization in development and disease. *Nat Rev Mol Cell Biol* **18**, 477-494, doi:10.1038/nrm.2017.36 (2017).
44 Meacham, C. E. & Morrison, S. J. Tumour heterogeneity and cancer cell plasticity. *Nature* **501**, 328-337, doi:10.1038/nature12624 (2013).
45 Ackermann, M. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat Rev Microbiol* **13**, 497-508, doi:10.1038/nrmicro3491 (2015).

46 Khushi, M., Napier, C. E., Smyth, C. M., Reddel, R. R. & Arthur, J. W. MatCol: a tool to measure fluorescence signal colocalisation in biological systems. *Sci Rep* **7**, 8879, doi:10.1038/s41598-017-08786-1 (2017).
47 Kreft, M., Milisav, I., Potokar, M. & Zorec, R. Automated high through-put colocalization analysis of multichannel confocal images. *Comput Methods Programs Biomed* **74**, 63-67, doi:10.1016/S0169-2607(03)00071-3 (2004).
48 Fletcher, P. A., Scriven, D. R., Schulson, M. N. & Moore, E. D. Multi-image colocalization and its statistical significance. *Biophys J* **99**, 1996-2005, doi:10.1016/j.bpj.2010.07.006 (2010).
49 Buggenthin, F. *et al.* An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinformatics* **14**, 297, doi:10.1186/1471-2105-14-297 (2013).

ACKNOWLEDGEMENTS

The research was partly supported by funds from UC Berkeley to HNL, and a Summer Grant from the Graduate Division at UC Berkeley to HS. Publication is made possible in part by support from the Berkeley Research Impact Initiative (BRII) sponsored by the UC Berkeley Library.

AUTHOR INFORMATION

Han N. Lim
Present address: Atomwise Inc., San Francisco, CA, USA

Affiliations

Department of Integrative Biology, University of California Berkeley, Berkeley, CA, USA
Weston Stauffer, Huanjie Sheng & Han N. Lim
Atomwise Inc., San Francisco, CA, USA
Han N. Lim

Contributions

W.S., H.S. and H.N.L. conceived the project. W.S. and H.S. performed the experiments (including creating EzColocalization). H.N.L., W.S. and H.S. wrote the paper.

Corresponding author

Correspondence to Han N. Lim.

Competing interests

The authors declare no competing interests

SUPPLEMENTARY INFORMATION

Contents

1. Description of packages and classes for EzColocalization

The first two packages have very basic purposes. The first is the “default” package (by Java™ convention) and its only function is to load the plugin within ImageJ. This package contains a single class, “EzColocalization_”, and outputs from this package are not accessible by other classes in other packages. The second package is “ezcol.files”, which has a single class (“FilesIO”) that loads all emblems and sample images for the GUI.

The third package is “ezcol.main”. It performs shared and general functions, and has six classes (“GUI”, “ImageInfo”, “MacroHandler”, “PluginStatic”, “PluginConstants”, and “AnalysisOperator”). GUI creates the GUI. ImageInfo stores information on the formats of the input images. MacroHandler enables use of the recorder in ImageJ so users can run macros that automatically run commands in batches. For example, the recorder can be used to create a macro to automatically modify and analyze a large set of images with particular settings. PluginStatic contains all static parameters (inputs) and static utility methods (common functions) used in analyses. PluginConstants contains all shared constants that are accessed by other classes. AnalysisOperator coordinates the operation of analyses in response to the inputs.

The fourth package is “ezcol.align”, which performs image alignment and has three classes (“BackgroundProcessor”, “TurboRegMod”, and “ImageAligner”). BackgroundProcessor enhances the contrast of images by: (i) subtracting background signal from pixels using the rolling ball algorithm in the “Subtract Background” function of ImageJ [1]; (ii) generating binary images of the reporter and cell identification channels with a user chosen algorithm from the “Auto Threshold” function in ImageJ [1] or thresholds manually set by the user (note: only a single manual threshold can be applied for a stack of images, and this is performed by selecting “*Manual*” and then displaying the thresholds by selecting “Show threshold(s)”; if no manual selection is made, the “Default” algorithm is applied); (iii) converting all pixels above the value identified by the Auto Threshold algorithm to a value of 255, and all those below it to 0; (iv) applying the “Fill Holes” function of ImageJ [1] on the binary images to better select

the entire area of cells; and (v) calculating the average signal of pixels below the threshold in each reporter channel.

TurboRegMod uses the "Translation" alignment algorithm from TurboReg [2] to calculate the required XY coordinate shifts to align the binary images from the output of the BackgroundProcessor class by maximizing the overlap of pixels above the threshold. Note: interpolation of pixel values and other alignment functions that are normally performed by TurboReg are avoided because these functions alter pixel values. ImageAligner performs the image alignment by applying the calculated XY shifts from TurboRegMod to the original images.

The fifth package is "ezcol.cell", which identifies cell areas and obtains pixel values. This package has six classes ("ParticleAnalyzerMT", "CellFinder", "CellFilterDialog", "CellDataProcessor", "DataSorter", and "CellData"). Note: "cell" refers to any objects being analyzed, including subcellular structures or whole organisms. ParticleAnalyzerMT is a customized multithreading version of the "Analyze Particle" function from ImageJ [1] and it is used to identify cell areas above the thresholds, which are pixels of the objects on the binary images produced by BackgroundProcessor (see previous package)). CellFinder takes inputs from the previous class and converts them into a format for the next class, performs watershed segmentation [1], and removes cells based on user defined cell filters. CellFilterDialog opens the window for additional cell filters. CellDataProcessor obtains the values of pixels identified for each cell. DataSorter and CellData sort the pixel values of cells based on intensity and store them so that these steps do not need to be repeated multiple times for later calculations.

The sixth package is "ezcol.metric", which performs colocalization analysis in response to inputs from ezcol.cell, and contains six classes ("BasicCalculator", "MetricCalculator", "CostesThreshold", "MatrixCalculator", "MatrixCalculator3D", and "StringCompiler"). BasicCalculator is an abstract class containing methods and values shared by the other "calculators" (*i.e.* MetricCalculator, MatrixCalculator, MatrixCalculator3D). MetricCalculator uses previously described algorithms to calculate Li's ICQ [3], Manders' colocalization coefficients M1, M2 [4] and M3, PCC [5], SRCC [6], and TOS [7]. "CostesThreshold" uses Costes' method for determining a threshold [8] and the algorithm was optimized for faster operation using ranked pixel values and dynamic programming as follows. The thresholds start at the maximum pixel values for each channel and PCC is calculated. Then the thresholds are decreased to the next highest pixel value, the values above the new threshold are subtracted from the stored sums, and PCC is calculated again from the new stored sums, and so on. During the entire process when all the pixels have been removed, we compare all the PCC values calculated for all thresholds to find the minimum absolute PCC value. MatrixCalculator calculates metric matrices for two reporter channels. MatrixCalculator3D creates metric matrices for three reporter channels. StringCompiler compiles and performs any custom analysis written by the user.

The seventh and eighth packages are “ezcol.visual.visual2D” and “ezcol.visual.visual3D”, which output plots and data from the analyses. These packages are located in the folder called “visual” and both obtain inputs from ezcol.cell for heat maps and scatterplots and from ezcol.metric for histograms and metric matrices.

The ezcol.visual.visual2D package contains nine classes for visualizing two dimensional data and results (“HeatGenerator”, “HeatChart”, “HistogramGenerator”, “HistogramStackWindow”, “ScatterPlotGenerator”, “PlotStackWindow”, “HeatChartStackWindow”, “OutputWindow”, and “ProgressGlassPane”). HeatGenerator normalizes pixels values so the maximum and minimum values are 0 and 255 (8-bit) or 65535 (16-bit) respectively for each cell, image, or stack. The normalized values are assigned colors from ImageJ lookup tables [9], or assigned from Matlab (R2015a, Mathworks, Natick, MA, USA) in the case of “hot” and “cool” colors. HeatChart is a modified version of the class JHeatChart (created by Tom Castle) which takes colors from the previous class to generate heat maps as RGB images, and values from MatrixCalculator to generate two dimensional metric matrices. HistogramGenerator and HistogramStackWindow generate histograms by respectively converting cell based data into histogram data starting with ten bins, and generating a stack of histograms for selected metrics. The number of bins can be increased or decreased in increments of one with the “nBin+” or “nBin−” buttons. ScatterPlotGenerator obtains pixel values from two reporter channels for five random cells per image in a stack. If five or less cells are present in an image, then pixel values are obtained for all cells in the image. PlotStackWindow creates and displays a stack of scatterplots, with each plot containing the pixel values for a single cell. HeatChartStackWindow generates the metric matrices window. OutputWindow generates the analysis summary window and its content. ProgressGlassPane generates the progress bar and presents tips in the GUI.

The ezcol.visual.visual3D package has 14 classes for visualization of three reporter channels in dynamic three dimensional scatterplots and metric matrices (“Arrow3D”, “Cube3D”, “Element3D”, “GraphicsB3D”, “Line3D”, “Point3D”, “Polygon3D”, “Rect3D”, “Renderer”, “ScatterPlot3D”, “ScatterPlot3DWindow”, “Spot3D”, “Square3D”, and “Text3D”). All classes are adopted from the jaytools.jar written by Urah Jay. His original classes are modified particularly for three dimensional scatterplots. Element3D is an abstract class (which means it cannot be initialized or constructed) containing methods and values shared by the other classes, including “Arrow3D”, “Cube3D”, “Element3D”, “Line3D”, “Point3D”, “Polygon3D”, “Rect3D”, “Spot3D”, “Square3D”, and “Text3D”. These classes represent the corresponding 3D elements as their names suggest; for example, “Arrow3D” is a class to indicate an arrow on a 3D graph. Some of these 3D elements (“Arrow3D”, “Polygon3D”, “Rect3D”, and “Square3D”) are not used for the purpose of this plugin but are kept for completeness of the package. Renderer is the 3D graphics process of automatically converting 3D elements into 2D image data, the results of which are feed into GraphicsB3D to paint the 2D image data of the projected 3D elements on the image canvas. ScatterPlot3D and ScatterPlot3DWindow generate 3D scatterplots and metric matrices by respectively converting the cell data into a compatible format for plotting and presenting the data in the plot window. 3D scatterplots are created in the same manner as ScatterPlotGenerator except for

obtaining pixel values from three reporter channels. Usually the first cell is shown when the image window opens, and users can select the next cell or the previous cell by clicking the forward and back buttons on the window. 3D metric matrices are generated in the same way as HeatChart but all color squares are replaced by 3D spheres to enable visualization of deeper layers in the 3D matrices.

The ninth and final package is “ezcol.debug”. It has two classes, and reports errors and warnings within the plugin. It contains a class, “ExceptionReporter”, which handles and reports errors or warnings, and the class, “Debugger”, which was used during development to debug the plugin.

2. Data acquisition guidelines

Accurate colocalization measurements begin with good experimental data, which depends on the samples, the reporters, the imaging system, data collection methods, controls and replicate measurements. Some guidelines the authors have found useful include the following. Samples should be prepared in a manner that: preserves the native spatial organization, minimizes touching of cells or organisms (which makes the identification of individual cells or organisms easier), and minimizes movement of cells or organisms (especially for live imaging). The reporters should be optimized to specifically label the molecules of interest (which includes minimizing excess or non-specifically bound reporter), and to minimize cross-talk (also known as bleed-through) between the reporter signals and between each reporter signal and non-reporter signals in cells and tissues (*e.g.* autofluorescence). In addition, reporters with the highest specific signal should be preferentially paired with targets that have the lowest concentration. Too much reporter is sometimes more problematic for colocalization measurements than too little because of non-specific labeling or aggregation. A reporter that identifies the cell boundary or entire cell should be considered if the cell boundaries are unclear in bright-field imaging to facilitate automated cell identification so that single cell measurements of colocalization can be easily performed. The imaging system should be set-up with: a high quality monochromatic camera to maximize the signal-to-noise ratio, controls to check the settings and reproducibility of measurements on different days, and adjustments to the light source or neutral density filters to prevent oversaturated pixels with artificially low intensity values. It is important to recognize that misalignment between imaging channels often occurs (and may occur after the initial set-up and alignment) therefore images from different channels should ideally be overlaid in each experiment to evaluate the alignment and to correct any misalignment by adjusting the physical apparatus or the analysis. Differential interference contrast (DIC) is not recommended, and users should instead use phase contrast or another method that does not create shadows for identifying cell boundaries. Generally, it is preferable to maximize the resolution, but the scale of the cells and structures must be considered. For example, measuring the colocalization of reporters in intracellular structures will require a higher level of resolution than measuring colocalization at different tissue structures or organs. Additional guidance on the practical aspects of setting up a system for colocalization measurements is available in several reviews [10-12]. The data should be collected at the highest number of bits to maximize the dynamic

range of the signal, and images saved in an appropriate format (see note below). The importance of controls for the proper analysis of the colocalization measurements cannot be overstated. Researchers should not only include appropriate biological controls (*e.g.* deletions strains without the labeled protein) but should also measure some cells with only one of each reporter to quantify and to correct for bleed-through. In addition, independent replicate measurements of controls and samples on different days are important because labeling, microscopy set- up (especially in shared facilities), and any automated settings for image collection can vary dramatically between different experiments and often without the researcher being aware of it until the analysis is performed. As an aside, researchers should only use deconvolution or image corrections that have been proven to provide more accurate representation of localization for their specific reporters, samples, and imaging system.

The format of the images is important. The image file format should be TIFF or another lossless compression format with a single value for pixel intensity. A color camera that records pixel values in RGB can be problematic because it is unknown how the three values contribute to total signal intensity. Pseudocolors can be created for visualization purposes if the pixel intensity values are not changed. RGB and pseudocolor images can be distinguished by looking at the information on top of the image window in ImageJ.

3. Image alignment

The Inputs tab provides the option for the alignment of images from different channels. The alignment is performed by: (i) subtracting background signal from the cell identification and reporter images to enhance contrast using the rolling ball algorithm of the “Subtract Background” function (note: this step can be turned on or off in the “Parameters...” options of the “Settings” menu); (ii) thresholding the resulting images; (iii) creating a binary mask from the thresholded images; (iv) processing the binary mask with the “Fill Holes” function to ensure cell interiors are selected; (v) aligning the reporter channels and binary mask image using the translation alignment algorithm component of the TurboReg plugin [2]; (vi) obtaining the X and Y coordinate offset values from the alignment and using them to align the original cell identification and reporter images; and (vii) removing overhanging pixels and filling-in pixels (with a value of zero) so all images in the stacks have same size (yellow area in **Fig. 1B**). Note: TurboReg functions that interpolate pixel values are not used because they change the original values.

4. Heat maps, scatterplots and metric matrices

Many factors should be considered when performing analyses and selecting a metric for quantifying localization. These factors should include heterogeneity in the data, the specificity of the reporter, the relative intensity of the intracellular and extracellular background signals, and the relationship between the intensities of the reporter signals. EzColocalization provides tools in the Visualization tab to help users evaluate these considerations.

Heat maps created by EzColocalization can be normalized for each cell, each image, or each stack ("cell heat maps", "image heat maps" and "stack heat maps" respectively). Cell heat maps can help visually identify the locations in cells where reporters have the highest and lowest intensity, and the localization patterns (*i.e.* colocalization, anticolocalization and noncolocalization of the reporters). Image heat maps can show whether different cells have different average signal intensities within each image. The cell and image heat maps should be carefully inspected for evidence of heterogeneity among cells with respect to: the locations of reporters within cells, the localization pattern (*i.e.* relative positions of the reporters), and the average signal intensity. If there is heterogeneity, then it may be appropriate to limit analysis to a subpopulation of cells by using the cell filters in EzColocalization so that measurements are not an average of multiple populations. Image heat maps should also be examined to determine if the pixels with the highest signal (likely containing reporter) have similar levels of intensity to the pixels with the lowest signal ("background"). If so, then analysis may be improved by selecting individual cells from the image so that the only intracellular pixels are analyzed or by selecting thresholds so that only pixels with signal greater than background levels are analyzed (see metric matrices below) [6,10].

Scatterplots reveal the relationship between the signal intensities for different reporters. Evaluating this relationship is important because different assumptions about the relationship of the reporter signals are central to the calculation, interpretation and selection of the metrics for colocalization (see next section). Scatterplots may also reveal if different cells or organisms within a sample have very different intensities or different relationships between pixel intensity. If there is heterogeneity, cell filters may be able to limit analysis to a more homogeneous population. In addition to cell-to-cell heterogeneity, there may be heterogeneity within each cell; that is, different relationships between the signals at different levels of signal intensity. For example, a cell may have pixels with low signal for two reporters that have no correlation and pixels with high signal for the same two reporters that have a positive correlation (due to specific binding to a protein) [11]. In cases where there are different relationships between the pixels at different levels of signal, it may be possible to select thresholds for the reporter signals so that colocalization is only measured for a subset of pixels.

Metric matrices can be calculated for six different colocalization metrics in EzColocalization: TOS with linear or logarithmic scaling [7], PCC [12], SRCC [13], Manders' colocalization coefficients M1 and M2 [4,12] and ICQ [3]. Each metric matrix calculates the value of the selected metric at every combination of the thresholds chosen (**Fig. S1**). Metric matrices can quickly determine whether there are general patterns of colocalization, anticolocalization or noncolocalization that depend on signal intensity [7,14]. A metric matrix can also help to select a threshold that provides a better measure of colocalization for a subset of pixels with different intensities in a cell. That is, the selection of thresholds via the metric matrices can provide more targeted analysis. Because the thresholds in the metric matrix are measured in terms of the percentage of pixels rather than absolute signal level, the metric matrix is well-suited to comparing and

aggregating values in a groups of cells where there may be some differences in average signal intensity and cell size.

In relation to selecting thresholds, EzColocalization provides two options: Costes' method and manual selection. Costes' method chooses the thresholds automatically [8]. The advantage of automatic selection of the thresholds is that it decreases the potential for user bias. However, the method often does not work well if the signal intensities of the intracellular and background pixels are not clearly distinguishable, the reporter signals do not have similar levels of intensity or a monotonic relationship, or there are outlier pixels with high signal [12]. Further, changing the order of reporter channels can result in different Costes' thresholds due to asymmetry in the linear regression used by the Costes' algorithm. Manual selection of thresholds by the user is more flexible but it requires care to ensure they are chosen appropriately. The heat map and scatterplots, as well as the metric matrices, can guide the manual selection of the thresholds. Metric matrices can help ensure the thresholds are chosen so that they are representative of broad trends and the results are robust (*i.e*. a small change in the values of the thresholds should not substantially alter the result). Two notes of caution in regard to the selection of thresholds: (i) the metric matrix should not be used to "fish" for a metric and threshold values to give a result that is not broadly consistent with all the data; and (ii) the selection of thresholds must balance the need to eliminate pixels with background or non-specific signal against the need to keep as many pixels as possible so the results of the analysis are broadly representative and not fluctuating due to the noise associated with having a small number of values. Additional guidance on the selection of thresholds is provided in previous publications [7,14].

Colocalization metrics for two reporter channels

This section provides brief and general guidelines for selecting a colocalization metric. More detailed information on colocalization metrics is published elsewhere [7,10,11,15]. As mentioned in the previous section, it is important to examine the scatterplots to determine the relationship between the signal intensities of the reporters before choosing a colocalization metric.

Pearson's correlation coefficient (PCC) is the covariance of two variables divided by the product of their standard deviations. It is typically used to measure the linear correlation of the signal intensity values for two reporters (**Fig. S2**) [5,16]. PCC values can range from −1 which indicates a strong negative correlation between the signals (anticolocalization) to 1 which indicates a strong positive correlation (colocalization). A PCC value of 0 indicates there is no correlation (noncolocalization). Note: PCC could be used to measure nonlinear relationships following a nonlinear transformation of the data, although this is not typically done for measuring colocalization.

Spearman's rank correlation coefficient (SRCC) is calculated by ranking the pixels according to the intensity of signal for each channel and then measuring the correlation in the rankings between two channels [13]. SRCC measures whether the signal

intensities of the reporters have a monotonic relationship (**Fig. S2**), and it is relatively insensitive to outliers because it is based on rankings. Therefore SRCC is suitable for non-linear, monotonic relationships such as power law or logarithmic functions. SRCC values range from −1 (anticolocalization) to 1 (colocalization) [15], and 0 indicates there is no correlation (noncolocalization).

The intensity correlation quotient (ICQ) is the ratio of the total number of pixels where the signal intensity is above the means for both channels or below the means for both channels (*i.e.* excluding pixels that are above the mean in one channel and below the mean in the second channel), divided by the total number of pixels, minus 0.5 [3,6]. ICQ ranges from –0.5 to +0.5. ICQ is essentially a sign test with positive or negative values for pixels that are on a positive or negative slope of a function through the mean of both channels. ICQ, like SRCC, is often used to evaluate whether the signal intensities of two reporters have a monotonic relationship (although it could also be used for some types of non-monotonic relationships) (**Fig. S2**). ICQ is less sensitive to outliers than PCC. ICQ is not an appropriate metric for heterogeneous samples because the mean may not be an appropriate point around which localization should be evaluated.

Manders' colocalization coefficients M1 and M2 are calculated by determining the sum of the intensities of pixels that exceed thresholds for both signals 1 and 2 divided by the sum of the intensities of the pixels that exceed the threshold for signal 1 or by the sum of the intensities of the pixels that exceed the threshold for signal 2, respectively [4,12]. The threshold can be determined by several algorithms including Costes' threshold [8]. Disadvantages of M1 and M2 are that both values are needed to determine whether there is colocalization, and the interpretation of these values is complicated by them being dependent on the threshold values [17]. Manders' colocalization coefficients M1 and M2 (and also the threshold overlap score defined below) tend to be better for evaluating colocalization or anticolocalization in cases where there is not a clear localization pattern, there is a mixed pattern of localization, or there is a non- monotonic relationship (**Fig. S2**).

The threshold overlap score (TOS) is a newer metric that shares some similarity to M1 and M2 in that it calculates the overlap in pixels above a threshold [7,14]. TOS is calculated by determining the number of pixels that exceed thresholds for both signals 1 and 2 and dividing this number by the number of pixels that exceed the threshold for signal 1 or by the number of pixels that exceed the threshold for signal 2 (note: following normalization the same TOS value will be obtained using either denominator [7]). Unlike M1 and M2 there is no weighting for signal intensity. In addition, TOS divides the observed overlap by the overlap expected to occur simply by chance (which is not done for M1 and M2). A result of this normalization is that TOS measures colocalization as a single value which makes it easier to interpret and compare between experiments than Manders' colocalization coefficients [7]. TOS values are rescaled so that −1 corresponds to the minimum possible overlap (anticolocalization), 0 corresponds to the same overlap as would occur by chance (noncolocalization), and 1 corresponds to the maximum

possible overlap (colocalization). The default rescaling option is linear because it is easily interpreted, and its value

reflects the fraction between random distribution and the minimum or maximum values (-1 or +1 respectively). For example, a value of 0.5 represents half the maximum possible overlap. EzColocalization also permits logarithmic rescaling (natural log) for users requiring a metric without a discontinuity in the first derivative, but it is harder to interpret than linear rescaling [7]. As mentioned above, TOS is suitable for the analysis of experiments that have non-monotonic relationships, mixed patterns of localization, or unclear localization patterns (**Fig. S2**). TOS can also be used for monotonic relationships including linear correlations, although in such cases it may not be as sensitive or specific as other metrics (*e.g.* PCC).

In summary, PCC is often used for datasets where the reporters have an approximately linear relationship between the pixel values. SRCC and ICQ are commonly used to evaluate whether the signal intensities of the two reporters have a monotonic relationship, and are generally considered more robust to outliers than PCC. Manders' M1 and M2 or TOS are often preferred in cases where there is no clear monotonic localization pattern, or mixed patterns of localization.

6. Colocalization metrics for three reporter channels

The metrics for two channels were: (i) PCC; (ii) SRCC; (iii) ICQ; (iv) Manders' coefficients; and (v) TOS with linear or logarithmic rescaling. Of these metrics, we extended ICQ, Manders' coefficients and TOS (linear or logarithmic rescaling) to measure colocalization for three reporters, and their derivations are below. PCC and SRCC were not extended for three reporters because their meaning and interpretation becomes much more complicated. Specifically, no single value of PCC (or SRCC) can represent the standardized covariance. Instead there are multiple values, each of which reports the extent two channels (independent variables) can predict the signal in the third channel (dependent variable). The first component of principal component analysis (PCA) should be used to measure linearity without assuming dependency of three channels [18]. However, PCA is difficult to interpret in relation to colocalization analysis and therefore was not included [19].

Li's ICQ [3] can be easily expanded to three (or more) channels.

$$\mathrm{ICQ} = \frac{N_{above}+N_{below}}{N_{total}} - 0.5, \qquad \text{Eq. 1}$$

where N_{above} is the number of pixels above the means of all three channels, N_{below} is the number of pixels below the means of all channels, and N_{total} is the total number of pixels. For two channels, ICQ is a crude measure of the fraction of pixels that are on the positive diagonal; that is, it can be interpreted as the fraction of pixels that are broadly consistent with a monotonic increasing relationship. For three channels, ICQ provides a crude measure of whether pixel values tend to increase in all three channels. However, the interpretation of the value is more complicated because of the combinatorics; a pixel may have values above or below the mean in 8 possible combinations. A value of −0.25 would be expected if the pixel values have a random distribution, and assuming the

median and mean values are approximately equal. In this case, a value >−0.25 may indicate a positive relationship, but it does not exclude the co-presence of a negative relationship. A value of <−0.25 indicates a negative relationship but it does not rule out a positive relationship in a subset of pixels.

The use of Manders' colocalization coefficients for three channels (i.e. M1, M2, and M3) has been previously reported [4]. The derivation of Manders' colocalization metrics for more than two channels is straight forward as it simply evaluates the proportion of overlapping signal. However, Manders' colocalization metric are often used with an automated method of threshold selection, such as Costes' method, and these methods typically do not readily extend to three channels [8]. Therefore, EzColocalization users can either select thresholds manually or by using the metric matrix for Manders' colocalization coefficients with three channels. The thresholds are measured as F_T. M_1, M_2, … M_n can be calculated by Eq. 1 where there are at least two reporters:

$$M_i = \frac{\sum G_{i,coloc}}{\sum G_i}, \qquad \text{Eq. 2}$$

where $G_{i,coloc}$ is the value of each pixel in channel i that is above all thresholds and G_i represents the value of each pixel in channel i that is above the threshold for only channel i. The number of Manders' colocalization coefficients is equal to the number of channels, therefore three values need to be interpreted for three channels. Three values can be difficult to interpret collectively and to compare colocalization between samples. Another challenge is that the interpretation of the Manders' colocalization coefficient depends on the selected thresholds [7].

TOS measures the overlap of the signal above the threshold for each channel accounting for the amount of overlap that would be expected to occur by random chance for different thresholds [7]. One of the first steps in calculating TOS is to determine the number of pixels in each cell that exceed the thresholds for all three reporter channels (A_{coloc}) and the number of pixels that exceed the threshold for one of the reporter channels (A_i, where i is the i^{th} channel). Dividing the former by the latter is the "observed AO". This calculation, is equivalent to calculating the fraction of pixels in the cell that exceed the threshold for all three channels (F_{coloc}) divided by the fraction of pixels that exceed the threshold for the chosen channel i (F_{Ti}). That is,

$$\text{observed AO}_i = \frac{A_{coloc}}{A_i} = \frac{A_{coloc}/A_{total}}{A_i/A_{total}} = \frac{F_{coloc}}{F_{Ti}}, \text{where } i = 1, 2 \text{ or } 3. \qquad \text{Eq. 3}$$

Note: F_{Ti} and F_{coloc} are fractions rather than percentages for all equations in this section, and are defined as greater than zero and less than or equal to one.

The next calculation is the expected AO value assuming uniformly distributed random pixel values. If the pixels above the threshold for the first channel are randomly distributed throughout the cell, then the chance a pixel above the threshold for the second channel overlaps one of the pixels that exceeds the threshold for the first channel, is simply equal to the fraction of pixels above the threshold for the first channel (previously explained elsewhere [7]). Following from this, the chance a pixel that exceeds the threshold for the third channel overlaps a pixel that already exceeds both the first and second channels is simply the product of the fraction of pixels that exceed the first and second reporter channels. Therefore, the

$$\text{expected AO}_i = \frac{F_{T1} \times F_{T2} \times F_{T3}}{F_{Ti}}, \text{where } i = 1, 2 \text{ or } 3. \quad \text{Eq. 4}$$

The observed AO is divided by the expected AO to generate the "AO ratio", which accounts for the increase in overlap that occurs with selection of more pixels (*i.e.* greater F_T).

$$\text{AO ratio} = \frac{F_{coloc}}{F_{T1} \times F_{T2} \times F_{T3}}. \quad \text{Eq. 5}$$

The AO ratio is equal to 1 for cells where the overlap is the same as expected by chance. The value of the AO ratio depends on whether the observed overlap is more or less than expected by chance as well as the selected thresholds. The latter can make interpretation difficult, therefore the AO ratio is rescaled to generate the TOS, which enables easier comparison of analyses with different thresholds.

To rescale the AO ratio, the minimum and maximum value must be determined for the thresholds. The minimum AO ratio can be zero if the sum of the F_T for two channels is less than or equal to 1 (*i.e.* $F_{T1} + F_{T2} \leq 1$). In this case where first two channels do not overlap, the threshold for the third channel is inconsequential. That is, if there is no overlap of pixels above the thresholds for two channels then there can be no overlap of all three channels, even if all the pixels are selected for the third channel (*i.e.* $F_{T3} = 1$) and consequently the minimum AO ratio would be zero. That is, if $F_{T1} + F_{T2} + F_{T3} \leq 2$, then it is possible for the minimum AO ratio to be equal to zero. If $F_{T1} + F_{T2} + F_{T3} > 2$ then overlap of all three channels must occur by at least the amount exceeding 2. In summary,

$$\text{minimum AO ratio} = \begin{cases} \frac{F_{T1} + F_{T2} + F_{T3} - 2}{F_{T1} \times F_{T2} \times F_{T3}}, \text{when } F_{T1} + F_{T2} + F_{T3} > 2 \\ 0, \text{ when } F_{T1} + F_{T2} + F_{T3} \leq 2 \end{cases}. \quad \text{Eq. 6}$$

The limits of the minimum AO ratio are 0 and 1.

The maximum AO ratio occurs when all three channels maximally overlap, and the maximum amount of overlap can be no more than the minimum F_T. For example, if two channels both have thresholds that select 80% of pixels and the third channel only selects 5% of pixels in the cell, then the maximal overlap of the selected pixels can be no more than 5% of the pixels in the cell; that is, the minimum of the three F_T values.

$$\text{Maximum AO ratio} = \frac{\text{minimum}\{F_{T1}, F_{T2}, F_{T3}\}}{F_{T1} \times F_{T2} \times F_{T3}}. \quad \text{Eq. 7}$$

The last step in calculating TOS is to rescale the AO ratio using the limits for the minimum AO ratio and for the maximum AO ratio as previously reported [7]. A TOS value reflects the fraction of the "distance" between random chance (also known as the null distribution) and the minimum or maximum possible overlap for the thresholds. A positive value indicates colocalization, zero indicates overlap that is no more or less than a random distribution, and a negative value is anticolocalization. For example, 0.5 is halfway between a random distribution and maximum TOS value (half-maximal colocalization for the chosen thresholds) and value of -0.5 is halfway between a random distribution and the minimum possible TOS value (half-maximal anti-colocalization for the chosen thresholds). It should be noted that the contribution of each channel to the colocalization measurement is not specified in the TOS value. Therefore,

anticolocalization may be due to one single channel not overlapping with the other two (as opposed all channels not overlapping).

7. Custom analysis

The Custom subtab in the Analysis tab allows users to perform custom mathematical analysis for all pixel intensity values in selected cells without having to directly modify the code for EzColocalization. In brief, custom written code inserted into the Custom tab of the plugin uses the same cells or organisms that would be selected by the cell filters (with any alignment used) for non-custom analyses. Each cell's pixel intensity value for each reporter channel are stored in an array, named c1, c2, and c3 for reporter channels 1, 2, and 3 respectively. The order of the pixels within each array is the same; that is, the same index within each array is the same pixel in each channel, and is the intensity value for that channel. The pixel values in the arrays can be analyzed using code written with standard mathematical functions in Java. Selecting the "Resource" button takes the user to a website with a list of operators and functions in Java for mathematical calculations.

8. Supplementary references

1 Ferreira, T. & Rasband, W. ImageJ User Guide — IJ 1.46. (2012). <imagej.nih.gov/ij/docs/guide>.

2 Thevenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process* **7**, 27-41, doi:10.1109/83.650848 (1998).

3 Li, Q. *et al.* A syntaxin 1, Galpha(o), and N-type calcium channel complex at a presynaptic nerve terminal: analysis by quantitative immunocolocalization. *J Neurosci* **24**, 4070-4081, doi:10.1523/JNEUROSCI.0346-04.2004 (2004).

4 Manders, E. M. M., Verbeek, F. J. & Aten, J. A. Measurement of colocalization of objects in dual-colour confocal images. *J Microsc* **169**, 375-382 (1993).

5 Manders, E. M., Stap, J., Brakenhoff, G. J., van Driel, R. & Aten, J. A. Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. *J Cell Sci* **103 (Pt 3)**, 857-862 (1992).

6 Adler, J. & Parmryd, I. Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. *Cytometry A* **77**, 733-742, doi:10.1002/cyto.a.20896 (2010).

7 Sheng, H., Stauffer, W. & Lim, H. N. Systematic and general method for quantifying localization in microscopy images. *Biol Open* **5**, 1882-1893, doi:10.1242/bio.019893 (2016).

8 Costes, S. V. *et al.* Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys J* **86**, 3993-4003 (2004).

9 Harrington, K., Stiles, T. S., Venkatraman, L., Prahst, C. & Bentley, K. in *BioImage Informatics Conference.*

10 Barlow, A. L., Macleod, A., Noppen, S., Sanderson, J. & Guerin, C. J. Colocalization analysis in fluorescence micrographs: verification of a more

accurate calculation of pearson's correlation coefficient. *Microsc Microanal* **16**, 710-724 (2010).
11 Bolte, S. & Cordelieres, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* **224**, 213-232 (2006).
12 Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *Am J Physiol Cell Physiol* **300**, C723-742 (2011).
13 Adler, J., Pagakis, S. N. & Parmryd, I. Replicate-based noise corrected correlation for accurate measurements of colocalization. *J Microsc* **230**, 121-133, doi:10.1111/j.1365-2818.2008.01967.x (2008).
14 Sheng, H., Stauffer, W. T., Hussein, R., Lin, C. & Lim, H. N. Nucleoid and cytoplasmic localization of small RNAs in Escherichia coli. *Nucleic Acids Res*, doi:10.1093/nar/gkx023 (2017).
15 Adler, J. & Parmryd, I. Colocalization analysis in fluorescence microscopy. *Methods Mol Biol* **931**, 97-109, doi:10.1007/978-1-62703-056-4_5 (2013).
16 Cordelieres, F. P. & Bolte, S. Experimenters' guide to colocalization studies: finding a way through indicators and quantifiers, in practice. *Methods Cell Biol* **123**, 395-408 (2014).
17 McDonald, J. H. & Dunn, K. W. Statistical tests for measures of colocalization in biological microscopy. *J Microsc* **252**, 295-302 (2013).
18 Andrews, D. T., Chen, L., Wentzell, P. D. & Hamilton, D. C. Comments on the relationship between principal components analysis and weighted linear regression for bivariate data sets. *Chemometrics and Intelligent Laboratory Systems* **34**, 231-244, doi:https://doi.org/10.1016/0169-7439(96)00031-7 (1996).
19 Suhr, D. D. Principal Component Analysis vs. Exploratory Factor Analysis. 203-230 (2005).

Inputs			Outputs						
Cell identification images*	1 reporter channel	2 or 3 reporter channels	Report of physical features of cells	Visualization: scatterplots, & metric matrices of cell	Visualization: heat maps of cell	Visualization: scatterplots & metric matrices of image or stack	Visualization: heat maps of image or stack	Analysis: measure colocalization in cells	Analysis: measure colocalization in images & stacks #
Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Yes	Yes	No	Yes	No	Yes	No	Yes	No	No
Yes	No	No	Yes	No	No	No	No	No	No
No	Yes	Yes	No	No	No	Yes	Yes	No	Yes
No	Yes	No	No	No	No	No	Yes	No	No
No	No	No	No	No	No	No	No	No	No

Table S1. Input images and possible outputs from IntracellularJ. * Cell identification channel may be reporter images as discussed in main text. Therefore it is possible to perform all possible analyses with two sets of images (with one set being used as both a cell identification image and a reporter image). # Cell identification images are required to distinguish intracellular and extracellular signal therefore without them any analysis or normalization must be for whole images or image stacks.

Physical (P) or Signal Intensity (S) parameter	Name	Units	Description
P	Area	pixels*	Number of pixels in a cell.
P	X	pixels*	Average x-coordinate of a cell.
P	Y	pixels*	Average y-coordinate of a cell.
P	Perimeter length	pixels*	Length of the outside boundary of a cell.
P	Width	pixels*	Width of a cell in the x-axis.
P	Height	pixels*	Height of a cell in the y-axis.
P	BX	pixels*	Top left x-coordinate of the smallest rectangle enclosing a cell.
P	BY	pixels*	Top left y-coordinate of the smallest rectangle enclosing a cell.
P	Major	pixels*	Primary axis of the best fit ellipse for a cell.
P	Minor	pixels*	Secondary axis of the best fit ellipse for a cell.
P	Circularity	unitless	Circularity of a cell calculated by $4 \pi \times$ area $\div$ perimeter^2. A value of 1 is a perfect circle and <1 is an ellipse.
P	Angle	degrees	Angle between the main axis of an ellipse fit to a cell and x-axis of the entire image containing the cell.
P	Feret's diameter	pixels*	Longest possible distance between any two points on a cell boundary.
P	FeretX	pixels*	Starting x-coordinate of the Feret's diameter of a cell.
P	FeretY	pixels*	Starting y-coordinate of the Feret's diameter of a cell.
P	Feretangle	degrees	Angle between a cell's Feret's diameter and its images x-axis.
P	MinFeret	pixels*	Minimum caliper diameter of a cell.
P	AR	unitless	Aspect ratio of a cell calculated by major axis $\div$ minor axis.
P	Round	unitless	Roundness of a cell calculated by $4 \times$ Area $\div$ $\pi \times$ major axis^2.

P	Area fraction	unitless	Percentage of pixels in an image which are included in a cell.
P	Solidity	unitless	Solidity of a cell calculated by its area ÷ area of its convex hull.
S	Mean (Ch1), (Ch2), or (Ch3)	arbitrary	Mean of pixel values for a cell in reporter channels 1, 2, or 3.
S	Mode (Ch1), (Ch2), or (Ch3)	arbitrary	Mode of pixel values for a cell in reporter channels 1, 2, or 3.
S	Median (Ch1), (Ch2), or (Ch3)	arbitrary	Median of pixel values for a cell in reporter channel s1, 2, or 3.
S	Minimum (Ch1), (Ch2), or (Ch3)	arbitrary	Minimum pixel value for a cell in reporter channels 1, 2, or 3.
S	Maximum (Ch1), (Ch2), or (Ch3)	arbitrary	The maximum pixel value for a cell in reporter channels 1, 2, or 3.
S	StdDev (Ch1), (Ch2), or (Ch3)	arbitrary	Standard deviation of pixel values for a cell in reporter channels 1, 2, or 3.
S	Skew (Ch1), (Ch2), or (Ch3)	unitless	Skewness of pixel values for a cell in reporter channels 1, 2, or 3.
S	Kurt (Ch1), (Ch2), or (Ch3)	unitless	Kurtosis of pixel values for a cell in reporter channels 1, 2, or 3.
S	RawIntDen (Ch1), (Ch2), or (Ch3)	arbitrary	Sum of all pixel values for a cell in reporter channels 1, 2, or 3.
S	IntDen (Ch1), (Ch2), or (Ch3)	arbitrary	Product of pixel number and average pixel value for a cell in reporter channels 1, 2, or 3.
S	Mean BgndRatio (Ch1),	fold background	Fold change of average pixel value for a cell versus the average pixel value for all pixels outside of cells in reporter channels 1, 2, or

	(Ch2), or (Ch3)		3. For example, “1x-2x” for Ch1 would select cells with a mean pixel value one to two fold the mean pixel value outside cells for reporter channel 1.
S	Median BgndRatio (Ch1), (Ch2), or (Ch3)	fold background	Fold change of median pixel value for a cell versus the median pixel value for all pixels outside of cells in reporter channel 1, 2, or 3. For example, “1x-2x” for a Ch1 filter input would select cells with a median pixel value one to two times the value of the median pixel value outside cells in reporter channel 1.

Table S2. Physical and signal intensity parameters for cell features. *Units for pixel size are arbitrary units unless users set a scale on the images.

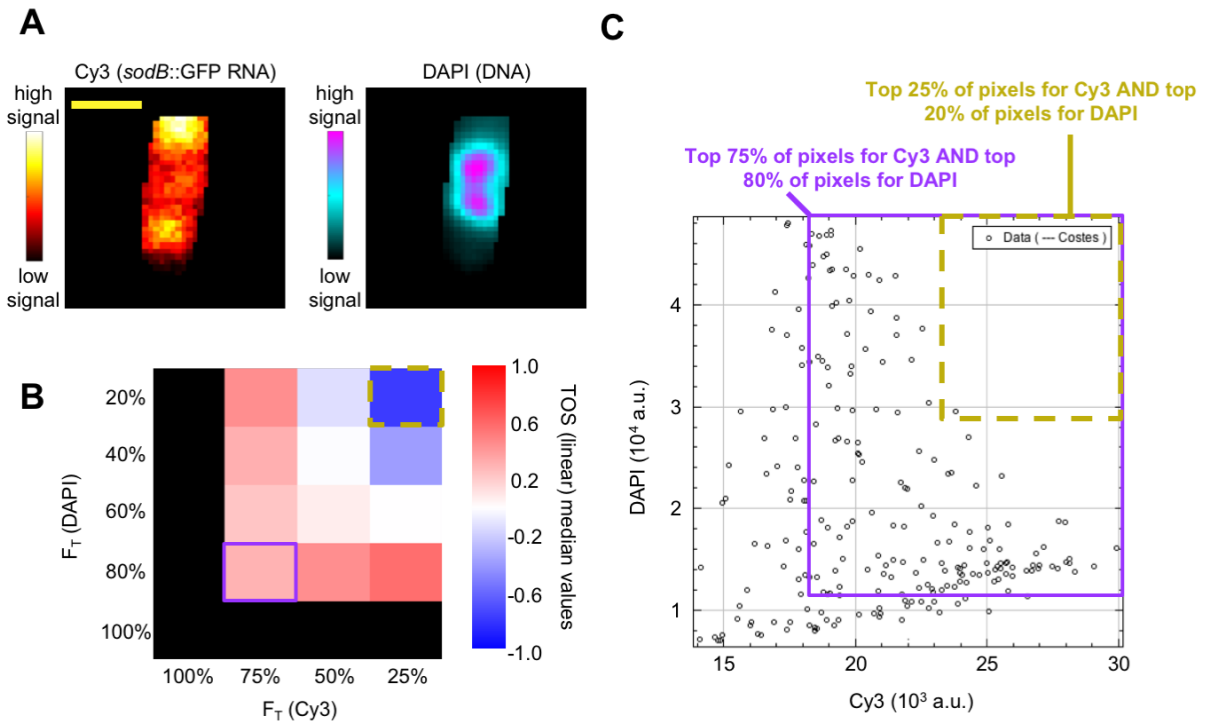

Fig. S1. Metric matrices and selected fractions**.** (**A**) Heat maps showing the intensities of Cy3 and DAPI signal for *sodB::gfp* RNA and DNA respectively in a bacterial cell. The *sodB::gfp* RNA was labeled with Cy3 labeled probes by RNA fluorescence in-situ hybridization. Scale bar is 1 µm. (**B**) Metric matrix with TOS values (linear) for the cell in **Panel A**. Each box in the matrix is the TOS value calculated for the pixels that are above the threshold for each channel. The thresholds are measured as the percentage of pixels with the highest signal for each channel (FT). For this example, the chosen FT

are the top 100%, 75%, 50% and 25% for Cy3 and the top 100%, 80%, 60%, 40% and 20% for DAPI. The calculated value of TOS is shown for every combination of thresholds and the approximate value is displayed in the bar to the right. The box is colored black when at least one threshold is 100% because in such cases TOS values are not informative; that is, when 100% of pixels are selected for at least one reporter then the overlap with the other channel must always be 100%. Threshold combinations indicated by the purple box and gold dash line box are discussed in **Panel C**. (**C**) Scatterplot of the pixels in the cell in **Panel A**. The purple box has pixels that are both in top 75% and the top 80% of values for Cy3 and DAPI respectively, which are used to calculate the TOS value shown in the purple box in the metric matrix (**Panel B**). The gold dash line box has pixels that are both in top 20% and the top 25% of values for Cy3 and DAPI respectively, which are used to calculate the TOS value shown in the gold box in the metric matrix (**Panel B**).

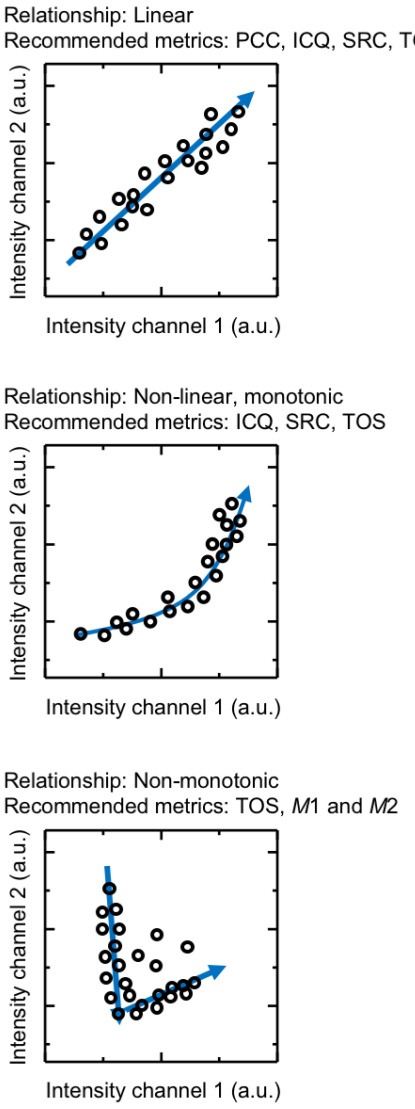

Fig. S2. Scatterplots identify the relationship between signal intensities. Scatterplots reveal the relationship between the intensities of different reporters, which is important for selecting an appropriate colocalization metric. Three relationships and the recommended metric for measuring colocalization for each are shown (see text of Supplementary Information). The blue line and the circles indicate the hypothetical relationship and hypothetical data points respectively.

Conclusion:

In this section we developed a robust tool with a great deal of functionality to perform microscopy analysis as discussed above. In the following sections we expand upon this advancement in microscopy image analysis by applying multiple techniques to challenging applications and systems, specifically live and fixed imaging of the *C. elegans* hermaphrodite gonad. By using these microscopy analysis techniques, along with manipulations of cell biology we are able to suggest some profound properties of the crossover designation patterning which is essential to meiosis.

Transition to Part II:

Meiosis and crossover designation

Meiosis is the form of cell division in which diploid cells undergo two successive divisions following only a single round of DNA replication to produce haploid cells. These haploid cells then develop into gametes (ova, sperm, or spores), and thus enable sexual reproduction. The two meiotic divisions are known as meiosis I, in which homologous chromosomes separate, and meiosis II, during which sister chromatids separate. The meiosis I division is preceded by a long period known as meiotic prophase, during which chromosomes undergo a set of unique events to enable each chromosome to separate from its homologous partner. Although often referred to as prophase I, this stage differs fundamentally from the canonical prophase of mitosis. During meiotic prophase, chromosomes are restructured[1-3], they undergo programed double-strand breaks in their DNA, which lead to homologous recombination. Through this process, sections of homologous chromosomes are exchanged, and these "crossover" events also lead to the formation of physical links between chromosomes that ensure their proper segregation. Only a subset of breaks become crossovers, while others are repaired through non-CO pathways that give rise to small regions of "gene conversion" without exchange of extensive chromosomal regions.

Meiotic prophase is divided into substages. These were first defined by chromosome cytology, and are now known to correspond to specific molecular events. During the leptotene stage, chromosomes become highly elongated as linear arrays of chromatin loops attached to a linear core, or axis. Concomitant with this reorganization or shortly thereafter, in zygotene, homologous chromosomes begin to pair and physically associate together. Double-strand breaks (DSBs) are catalyzed by the SPO-11 protein complex[4]. Pairing is accompanied by assembly of a proteinaceous structure, known as the synaptonemal complex (SC), between chromosomes. Pachytene is defined as the stage during which pairing and synapsis (SC formation) are complete. In *C. elegans*, pachytene is also the stage during which most recombination intermediates are repaired as non-COs, while a subset become designated for CO formation[5]. During the diplotene stage, the SC begins to disassemble, but homologous chromosomes remain associated through linkages called chiasmata. Finally, at diakinesis, homologous chromosomes remain associated by chiasmata and further condense into compact structures known as bivalents.

Following these events, there is a round of chromosome segregation and cell division that separates homologs. Faithful segregation of chromosomes at this stage is dependent on formation of a regular number of COs form between homologs[6], as a disproportionate number of COs leads to missegregation of chromosomes[7]. One indication of this dependence is that inefficient CO formation is linked to aneuploidy in humans[8,9]. This aneuploidy causes a number of genetic disorders[10], and aneuploidy in human embryos is the primary cause of first trimester miscarriages[11]. Because of their significance for human health, and their importance to meiosis, how COs are regulated has remained a fundamental unanswered question in the field of meiosis research.

To maintain the correct number of COs from an excess of DSBs, meiosis exhibits two different regulatory features. At least one CO per set of homologs always forms, through CO assurance[12]. Concurrently, the influence of CO interference[12] ensures that COs occur at a regular interval along the length of a set of homologous chromosomes. Near all COs display these effects, and are known as class I[13]. Notably, in some species of organisms a minimal number of COs also occur through a class II pathway, that shares the use of some class I proteins, but is regulated through different mechanisms[14] and does not display CO interference. Although limited in number, the presence of a class II CO confounds the study of COs, as there is known to be interplay between the regulation of class I and class II pathways when disruption of one class occurs[15,16].

The nematode *C. elegans* has become a powerful experimental system for studying meiotic mechanisms, including CO regulation. Genetic studies first revealed that each chromosome pair undergoes only a single crossover during meiosis. Class II COs appear to be absent in this organism[13], simplifying the analysis of crossover regulation. Additionally, *C. elegans* are transparent, and live imaging in adult animals can be used to visualize fluorescently tagged proteins throughout meiosis. The physical organization of the germline as an “assembly line” of meiotic stages greatly facilitates analysis. Specifically, the gonad allows for the observation of both live imaging and fixed cells in pachytene, the stage at which COs are designated and formed from the pool of DSBs. The development of genome editing techniques, particularly CRISPR/Cas9-based editing, has greatly expanded researchers’ ability to fluorescently tag proteins and to modify them at will[17]. Nevertheless, C. elegans has some shortcomings as an experimental system. For example, it still lacks a robust inducible expression system such as the GAL4-UAS system used in *Drosophila*. Additionally, transgenes can be silenced unpredictably by small RNA pathways, and there is no reliable system for overexpression of proteins in the germline[18].

Significant progress in understanding CO formation and regulation has been made within *C. elegans* or inferred from homologous model systems. Following pairing of homologous chromosomes and DSB formation, the RTR protein complex which contains several proteins including a helicase, binds to cut DNA and supports resection[19] by promoting the activity of HIM-6 helicase and working in conjunction with the MRN/X protein complex[20]. In addition to this function, the RTR complex has activity to form non-COs. The protein, RPA-1 binds during this process to keep single-stranded DNA from reannealing[20]. Following the resectioning of a DSB to contain a 3’ overhang, RAD-51 binds, which performs the homology search for both COs and non-COs[20].

The MutSγ complex, a heterodimer formed by Msh4 and Msh5, binds to strand-invasion intermediates[21]. Additionally, a cyclin-like protein, COSA-1 associates at low levels with these sites[21]. The number of these sites for which the CO indispensable MutSγ complex and COSA-1 proteins initially associates to outnumber the final count of COs. Because these sites contain CO factors, and thus are likely potential CO sites, but not all of which will eventually form COs, we refer to these as CO intermediates.

Concomitant with these steps in recombination, the SC, a proteinaceous structure, begins to form between homologous chromosomes and has assembled along their full-lengths when CO intermediates form. Currently, the SC in *C. elegans* is known to contain at least six structural proteins, SYP-1, SYP-2, SYP-3, SYP-4[22], SYP-5, and SYP-6 [23,24]. All of these have predicted coiled-coil domains. The SC assembles on the meiotic axis, a cohesin containing complex which forms the core of meiotic chromosomes[25]. In C. elegans, axis proteins include HTP-1, HTP-2, HTP-3, and HIM-3[25].

As the SC assembles, four additional proteins essential for CO formation localize along its length. These are a family of four paralogs known as ZHP-1, ZHP-2, ZHP-3, and ZHP-4[26]. These proteins share an N-terminal RING domain, a central coiled-coil domain, and a C-terminal region predicted to be unstructured. RING domain proteins are often components of E3 ubiquitin ligases. Some of the ZHPs in other organisms have been hypothesized to be SUMO ligases[27]. However, both SUMO (SMO-1) and its activating enzyme (UBC-89) are dispensable for ZHP functions in *C. elegans*[28].

The ZHPs interact pairwise to form two complexes with distinct localization dynamics and functions. ZHP-3 and ZHP-4 act as a heterodimer and are essential in CO formation[26]. While ZHP-1 and ZHP-2 act as a heterodimer to concentrate ZHP-3 and ZHP-4 at CO sites[26]. As meiosis progresses, the number of CO intermediates decreases to just one site per pair of homologous chromosomes. The concentration of ZHP-3 and ZHP-4 to a single site per chromosomes enables the loss of CO intermediates[26], and components including the MutSγ, COSA-1, and HIM-6 are enriched at CO sites[21]. This culling function of ZHP-3/4 is demonstrated in the case when ZHP-3/4 are not concentrated to CO sites[26], and COSA-1 maintains foci at multiple sites along chromosomes (i.e. CO intermediates). Following the formation of a CO, for each chromosome, the length of chromosome shorter to a chromosome end retains SC, referred to as the “short-arm”, while the “long-arm” between the chiasma and the more distant ends disassembles its SC, keeping only its axis assembled along chromosomes. Also of note in *C. elegans*, is the enrichment of DSBs along the arms of chromosomes, while the center and ends of chromosomes have much lower rates of DSB formations[29]. This bias in DSB placement likely leads to a bias of COs along chromosome arms.

Despite the identification of many key factors required for CO designation and eventual resolution, the mechanism(s) that ensure(s) designation of a single CO along each chromosome pair remains unclear. Although many models of CO designation have been proposed, they all have fundamental shortcomings. Robin Holliday proposed that the SC acts as a compartment with an essential CO factor in limited concentrations, the limitation of which results in formed COs inhibiting the presence of others[30]. By others, models have been proposed in which a polymer spreads out along chromosomes from a designated CO to knock off other nascent COs[31], and in another similar model a conformational change at a CO has a domino effect along chromosomes making the same conformational change of some nature to suppress additional COs[32]. A reaction-diffusion model with hypothetical molecules which interact and diffuse has been shown

to recapitulate CO designation[33], and a descriptive "counting model" exists which can recreate the spacing of COs in some organisms[34]. Unfortunately, all of these models listed do not identify the real agents or biological molecules which function within the models. A past model exists of CO formation which asserts that SC assembly in fact halts the formation of COs[35], and so SC initiation is at a CO site, and all other potential sites are suppressed. However, in some organisms, SC assembly does not begin at CO sites[36]. A "chromosome oscillatory movement" model proposed that the rapid meiotic movement of chromosomes is responsible for the formation of COs and non-COs[37], but studies have shown inhibition of such movements only modestly impacts CO formation[38]. A "beam-film" model, which posits that tension stored within DNA enables CO interference, has been developed and can be adjusted to match experimental CO distributions[6]. However, this beam-film model fails to explain many observed aspects of CO regulation. COs are determined by the presence of ZHP-3/4 at sites along the SC[26] (i.e. ZHP-3/4 retained along the full SC promotes additional COSA-1, sites thus determining COs). CO number is determined by SC length and not DNA length[39], while DNA length should determine COs if tension in DNA determines COs. There is dosage effects of RNF212 (a ZHP-3 ortholog in mammals) on CO number[40], wherein the number of COs formed is determined by RNF212 concentration[41]. This protein concentration dependence of COs is not clearly explained by the beam-film model. Finally, CO proteins are able to enrich to a single site, demonstrating the capacity to regulate CO number, in the absence of DNA and thus DNA tension[42]. Together, these findings suggest that protein levels, activities, and interactions can determine COs independent of DNA properties, and so further investigation is necessary to understand how COs are regulated.

Several discoveries have hinted towards useful avenues to investigate this fundamental unknown. First, the SC has been shown to be a liquid crystal[42]. A liquid crystal is a material with unique properties. The components have a regular structure of its components as in a crystal, but with mobility of the components as in a liquid. Additionally, the SC is phase-separated membraneless compartment, wherein its components demix from the nucleoplasm due to energetically favorable interactions amongst themselves to form a separate phase from the nucleoplasm and concentrate numerous crossover factors[42]. Together this suggests the SC could act as a compartment or medium with its concentrated mobile components for a mechanism, reaction, or signal of CO interference to inhibit COs within one set of homologs[42]. By having one compartment per set of homologs, there can be regulation on a chromosomal basis, allowing mechanisms within the compartment to ensure that a CO occurs and that only one CO forms. In support of this possibility, the SC is heavily implicated in CO designation. As noted above, the length of the SC [39], concentration of RNF212[40], an SC localized protein, and enrichment of ZHP proteins within the SC to CO sites have been shown to determine CO number[26]. Related, decreased concentration of the SC structural protein, SYP-1, has been shown to increase the number of COs[41]. However, contrary to these suggestions, while the SC is consistently essential to CO formation across sexual organisms, the model organism *S. pombe* is known to not form SC while still having COs[43], and in the model organism *S. cerevisiae*

it seems that the initial sites of SC assembly may be the sites of CO formation[44], in both cases possibly excluding the determination of COs through processes within the SC.

As a second major hint of what to investigate, the CO essential protein COSA-1 has a cyclin-like structure[45]. Because of this, along with recent findings in orthologous systems[46-48], we hypothesized that a cyclin-dependent kinase may exhibit the activity to establish COs in *C. elegans*. In the following projects within this thesis, we present work towards discovering the mechanism by which COs are regulated and established within *C. elegans* meiosis using a variety of techniques including advanced microscopy imaging and analysis to probe the dynamics, interactions, and functions of CO designation proteins.

REFERENCES

1 Ding, D.-Q. *et al.* Meiotic cohesins modulate chromosome compaction during meiotic prophase in fission yeast. *The Journal of cell biology* **174**, 499-508, doi:10.1083/jcb.200605074 (2006).

2 Chen, D. *et al.* Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. *The Journal of cell biology* **168**, 41-54, doi:10.1083/jcb.200407182 (2005).

3 Belmont, A. S. Mitotic chromosome scaffold structure: New approaches to an old controversy. *Proceedings of the National Academy of Sciences* **99**, 15855, doi:10.1073/pnas.262672799 (2002).

4 Dernburg, A. F. *et al.* Meiotic Recombination in C. elegans Initiates by a Conserved Mechanism and Is Dispensable for Homologous Chromosome Synapsis. *Cell* **94**, 387-398, doi:https://doi.org/10.1016/S0092-8674(00)81481-6 (1998).

5 Storlazzi, A., Xu, L., Cao, L. & Kleckner, N. Crossover and noncrossover recombination during meiosis: timing and pathway relationships. *Proceedings of the National Academy of Sciences* **92**, 8512-8516, doi:10.1073/pnas.92.18.8512 (1995).

6 Wang, S., Zickler, D., Kleckner, N. & Zhang, L. Meiotic crossover patterns: obligatory crossover, interference and homeostasis in a single process. *Cell Cycle* **14**, 305-314, doi:10.4161/15384101.2014.991185 (2015).

7 Louis, E. J. & Borts, R. H. Meiotic Recombination: Too Much of a Good Thing? *Current Biology* **13**, R953-R955, doi:https://doi.org/10.1016/j.cub.2003.11.040 (2003).

8 MacLennan, M., Crichton, J. H., Playfoot, C. J. & Adams, I. R. Oocyte development, meiosis and aneuploidy. *Seminars in Cell & Developmental Biology* **45**, 68-76, doi:https://doi.org/10.1016/j.semcdb.2015.10.005 (2015).

9 Wang, S. *et al.* Inefficient Crossover Maturation Underlies Elevated Aneuploidy in Human Female Meiosis. *Cell* **168**, 977-989.e917, doi:https://doi.org/10.1016/j.cell.2017.02.002 (2017).

10 Hassold, T. & Hunt, P. To err (meiotically) is human: the genesis of human aneuploidy. *Nature Reviews Genetics* **2**, 280-291, doi:10.1038/35066065 (2001).

11 Lathi, R. B., Westphal, L. M. & Milki, A. A. Aneuploidy in the miscarriages of infertile women and the potential benefit of preimplanation genetic diagnosis. *Fertility and Sterility* **89**, 353-357, doi:https://doi.org/10.1016/j.fertnstert.2007.02.040 (2008).
12 Shinohara, M., Oh, S. D., Hunter, N. & Shinohara, A. Crossover assurance and crossover interference are distinctly regulated by the ZMM proteins during yeast meiosis. *Nature Genetics* **40**, 299-309, doi:10.1038/ng.83 (2008).
13 Saito, T. T. & Colaiácovo, M. P. Regulation of Crossover Frequency and Distribution during Meiotic Recombination. *Cold Spring Harb Symp Quant Biol* **82**, 223-234, doi:10.1101/sqb.2017.82.034132 (2017).
14 Youds, J. L. & Boulton, S. J. The choice in meiosis – defining the factors that influence crossover or non-crossover formation. *Journal of Cell Science* **124**, 501-513, doi:10.1242/jcs.074427 (2011).
15 Gray, S. & Cohen, P. E. Control of Meiotic Crossovers: From Double-Strand Break Formation to Designation. *Annual review of genetics* **50**, 175-210, doi:10.1146/annurev-genet-120215-035111 (2016).
16 Sym, M. & Roeder, G. S. Crossover interference is abolished in the absence of a synaptonemal complex protein. *Cell* **79**, 283-292, doi:https://doi.org/10.1016/0092-8674(94)90197-X (1994).
17 Dickinson, D. J. & Goldstein, B. CRISPR-Based Methods for Caenorhabditis elegans Genome Engineering. *Genetics* **202**, 885-901, doi:10.1534/genetics.115.182162 (2016).
18 Kelly, W. G., Xu, S., Montgomery, M. K. & Fire, A. Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene. *Genetics* **146**, 227-238 (1997).
19 Jagut, M. *et al.* Separable Roles for a Caenorhabditis elegans RMI1 Homolog in Promoting and Antagonizing Meiotic Crossovers Ensure Faithful Chromosome Inheritance. *PLOS Biology* **14**, e1002412, doi:10.1371/journal.pbio.1002412 (2016).
20 Koury, E., Harrell, K. & Smolikove, S. Differential RPA-1 and RAD-51 recruitment in vivo throughout the C. elegans germline, as revealed by laser microirradiation. *Nucleic acids research* **46**, 748-764, doi:10.1093/nar/gkx1243 (2018).
21 Woglar, A. & Villeneuve, A. M. Dynamic Architecture of DNA Repair Complexes and the Synaptonemal Complex at Sites of Meiotic Recombination. *Cell* **173**, 1678-1691 e1616, doi:10.1016/j.cell.2018.03.066 (2018).
22 Schild-Prüfert, K. *et al.* Organization of the synaptonemal complex during meiosis in Caenorhabditis elegans. *Genetics* **189**, 411-421, doi:10.1534/genetics.111.132431 (2011).
23 Hurlock, M. E. *et al.* Identification of novel synaptonemal complex components in C. elegans. *Journal of Cell Biology* **219**, doi:10.1083/jcb.201910043 (2020).
24 Zhang, Z. *et al.* Multivalent weak interactions between assembly units drive synaptonemal complex formation. *Journal of Cell Biology* **219**, doi:10.1083/jcb.201910086 (2020).
25 Köhler, S., Wojcik, M., Xu, K. & Dernburg, A. F. Superresolution microscopy reveals the three-dimensional organization of meiotic chromosome axes in intact

Caenorhabditis elegans tissue. *Proceedings of the National Academy of Sciences* **114**, E4734-E4743, doi:10.1073/pnas.1702312114 (2017).
26 Zhang, L., Kohler, S., Rillo-Bohn, R. & Dernburg, A. F. A compartmentalized signaling network mediates crossover control in meiosis. *Elife* **7**, doi:10.7554/eLife.30789 (2018).
27 de Carvalho, C. E. & Colaiacovo, M. P. SUMO-mediated regulation of synaptonemal complex formation during meiosis. *Genes Dev* **20**, 1986-1992, doi:10.1101/gad.1457806 (2006).
28 Bhalla, N., Wynne, D. J., Jantsch, V. & Dernburg, A. F. ZHP-3 acts at crossovers to couple meiotic recombination with synaptonemal complex disassembly and bivalent formation in C. elegans. *PLoS Genet* **4**, e1000235, doi:10.1371/journal.pgen.1000235 (2008).
29 Yu, Z., Kim, Y. & Dernburg, A. F. Meiotic recombination and the crossover assurance checkpoint in Caenorhabditis elegans. *Seminars in Cell & Developmental Biology* **54**, 106-116, doi:https://doi.org/10.1016/j.semcdb.2016.03.014 (2016).
30 Holliday, R., Lewis, K. R. & Hultén, M. Recombination and Meiosis [and Discussion]. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **277**, 359-370 (1977).
31 King, J. S. & Mortimer, R. K. A polymerization model of chiasma interference and corresponding computer simulation. *Genetics* **126**, 1127-1138 (1990).
32 Kaback, D. B., Barber, D., Mahon, J., Lamb, J. & You, J. Chromosome size-dependent control of meiotic reciprocal recombination in Saccharomyces cerevisiae: the role of crossover interference. *Genetics* **152**, 1475-1486 (1999).
33 Fujitani, Y., Mori, S. & Kobayashi, I. A Reaction-Diffusion Model for Interference in Meiotic Crossing Over. *Genetics* **161**, 365 (2002).
34 Foss, E., Lande, R., Stahl, F. W. & Steinberg, C. M. Chiasma interference as a function of genetic distance. *Genetics* **133**, 681 (1993).
35 Egel, R. Synaptonemal complex and crossing-over: structural support or interference? *Heredity (Edinb)* **41**, 233-237, doi:10.1038/hdy.1978.92 (1978).
36 Rog, O. & Dernburg, A. F. Direct Visualization Reveals Kinetics of Meiotic Chromosome Synapsis. *Cell reports* **10**, 1639-1645, doi:10.1016/j.celrep.2015.02.032 (2015).
37 Hultén, M. A. On the origin of crossover interference: A chromosome oscillatory movement (COM) model. *Molecular Cytogenetics* **4**, 10, doi:10.1186/1755-8166-4-10 (2011).
38 Alleva, B. & Smolikove, S. Moving and stopping: Regulation of chromosome movement to promote meiotic chromosome pairing and synapsis. *Nucleus* **8**, 613-624, doi:10.1080/19491034.2017.1358329 (2017).
39 Kleckner, N., Storlazzi, A. & Zickler, D. Coordinate variation in meiotic pachytene SC length and total crossover/chiasma frequency under conditions of constant DNA length. *Trends in Genetics* **19**, 623-628, doi:https://doi.org/10.1016/j.tig.2003.09.004 (2003).
40 Reynolds, A. *et al.* RNF212 is a dosage-sensitive regulator of crossing-over during mammalian meiosis. *Nat Genet* **45**, 269-278, doi:10.1038/ng.2541 (2013).

41 Libuda, D. E., Uzawa, S., Meyer, B. J. & Villeneuve, A. M. Meiotic chromosome structures constrain and respond to designation of crossover sites. *Nature* **502**, 703-706, doi:10.1038/nature12577 (2013).
42 Rog, O., Kohler, S. & Dernburg, A. F. The synaptonemal complex has liquid crystalline properties and spatially regulates meiotic recombination factors. *Elife* **6**, doi:10.7554/eLife.21455 (2017).
43 Cromie, G. & Smith, G. R. Meiotic Recombination in Schizosaccharomyces pombe: A Paradigm for Genetic and Molecular Analysis. *Genome Dyn Stab* **3**, 195-195, doi:10.1007/7050_2007_025 (2008).
44 Fung, J. C., Rockmill, B., Odell, M. & Roeder, G. S. Imposition of Crossover Interference through the Nonrandom Distribution of Synapsis Initiation Complexes. *Cell* **116**, 795-802, doi:https://doi.org/10.1016/S0092-8674(04)00249-1 (2004).
45 Yokoo, R. *et al.* COSA-1 reveals robust homeostasis and separable licensing and reinforcement steps governing meiotic crossovers. *Cell* **149**, 75-87, doi:10.1016/j.cell.2012.01.052 (2012).
46 Wijnker, E. *et al.* The Cdk1/Cdk2 homolog CDKA;1 controls the recombination landscape in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **116**, 12534, doi:10.1073/pnas.1820753116 (2019).
47 Ashley, T., Walpita, D. & de Rooij, D. G. Localization of two mammalian cyclin dependent kinases during mammalian meiosis. *Journal of Cell Science* **114**, 685 (2001).
48 Viera, A. *et al.* CDK2 is required for proper homologous pairing, recombination and sex-body formation during male mouse meiosis. *Journal of Cell Science* **122**, 2149, doi:10.1242/jcs.046706 (2009).

Chapter 2: Meiotic protein dynamics suggest potential mechanisms of crossover regulation in *C. elegans*

ABSTRACT

In meiosis, programmed double-stranded breaks in DNA are induced to enable recombination between chromosomes. A subset of these breaks are eventually processed to become crossovers between homologous chromosomes, wherein sections of chromosomes are exchanged. The majority are repaired as non-crossovers. However, the mechanism by which cells regulate the number and assortment of crossovers amongst an excess of double-stranded breaks remains unclear. Prior work in our lab has suggested proteins may move along chromosomes through the SC, and that this movement may determine crossover regulation. Here, we characterize the movement of numerous chromosomal proteins in live *C. elegans* using fluorescence correlation spectroscopy (FCS) and fluorescence recovery after photobleaching (FRAP). We see a wide variation in protein mobility, and specifically find that synaptonemal complex proteins are dynamic, and that four RING domain containing proteins that regulate crossovers, the ZHPs, are all dynamic in contrast to other crossover factors. We also tested the dosage dependence of these ZHPs. Using mutants and knockdowns we increase and decrease the effective concentrations of these ZHPs, demonstrating their concentration sensitivity in *C. elegans*. Together, our results are consistent with a model in which crossovers are regulated by a system of dynamically interacting ZHPs and other factors that establish a pattern of crossovers based on their interactions and concentrations.

INTRODUCTION

Meiosis is the form of cell division during which cells decrease their ploidy by half, enabling the production of gametes in sexually reproducing organisms. Vital to this process is the formation of crossovers (COs) between homologous chromosomes, enabling the recombination of genes by an exchange of chromosomal sections, and the formation of bridges (chiasmata) between chromosomes. These bridges allow the proper alignment and segregation of homologous chromosomes during meiosis I, the first division round in meiosis. Before CO formation, an excess of programmed double-stranded breaks (DSBs) are made in DNA by meiotic cells. A small subset of these DSBs are processed into COs, while the majority are repaired as non-COs[1].

The locations of the majority of COs, known as class I COs[2], are determined by two patterning processes, known as CO assurance and CO interference. CO assurance ensures that each set of homologous chromosomes receives at least one CO. Due to CO interference, COs display decreased incidence in the formation of additional COs in a proximity-dependent manner along the chromosome[3]. The inhibition leads to DSBs close to designated COs being repaired as non-COs. Class II COs also exist which use some redundant and some unique proteins from class I, but most notably do not display

any CO interference[2]. The model system *C. elegans* only undergoes class I COs, greatly facilitating their study in this system, and leading to every set of homologous chromosomes receiving exactly one CO in wildtype meiosis, or six COs per nucleus. Localization of proteins are often used to visually demarcate the location and number of COs. While significant progress has been made to characterize the results and contributing proteins to both CO assurance and CO interference, the underlying mechanism to establish these patterning processes and orchestrate CO proteins remains unclear.

Numerous models of CO regulation already exist. Many primarily focus upon the physical mechanisms which determine the pattern. This includes a model in which an unknown polymer spreads from a designated CO to knock off proteins and inhibit other COs[4]. In another proposed model, an unknown conformational or property change at COs has a domino effect along their length to inhibit formation of other COs[5]. Other models have focused on using mathematical simulations and principles to show that they can recapitulate CO patterning. This includes a reaction-diffusion model of theoretical molecules[6], and a "counting model" which can be used to recreate the spacing of COs. Finally, other models have synthesized physical processes to explain the patterning and mathematical simulations to demonstrate their ability to recreate CO patterning. This includes the "chromosome oscillatory movement" model, in which rapid movements directed by centromere or telomere tethers are proposed to regulate COs[7]. Although notably the inhibition of such movements has minimal effects on COs[8]. It also includes the "beam-film" model, which predicts that tension stored within chromatin is released by CO formation, and that this release of tension propagates to stifle further CO formation[9]. Although these models are useful for predicting and explaining CO patterning, recent discoveries have suggested alternative mechanisms may contribute to or determine crossovers.

Central to these discoveries is the synaptonemal complex (SC), a proteinaceous structure which forms between paired homologous chromosomes during prophase of meiosis I. The presence of SC characterizes a substage of meiotic prophase known as pachytene (**Fig. 1A**), and in this stage COs are patterned. The disassembly of the SC signifies an exit out of pachytene, and a progression into the prophase I substage known as diplotene. The SC in *C. elegans* and other organisms, contains numerous structural proteins with coiled-coil domains enabling its assembly[10], and was first identified as a dense striated ladder-like structure in electron micrograph images of meiotic cells[11]. The SC is known to have defined regular architecture of its components[10], and assembles onto the meiotic axis[12], a cohesin containing protein structure which also assembles onto chromatin loops along the length of chromosomes.

Several features of the SC demonstrate its necessity and therefore function in CO patterning. The partial RNAi knockdown of a structural SC protein in *C. elegans*, SYP-1, resulted in the inhibition of CO interference and consequently additional COs[13]. The number of COs is dictated by the length of the SC [14,15], and not by the length of DNA[16], meaning that its length along chromosomes reflects genetic distance instead of physical

distance. The SC has liquid crystalline properties, meaning its components have mobility and structure all within a phase-separated compartment (**Fig. 1B**), although it remains unclear whether SC exchanges with the nucleoplasm. The liquid crystalline properties suggest that the SC is viable as a medium to establish CO patterning[12]. In the absence of DSBs and possibly DNA, SC components assemble into a circular globular structure termed polycomplex. Within polycomplexes of *C. elegans,* CO marker proteins, COSA-1, and ZHP-3, localize interdependently to a single site, as in wildtype SC[12]. The capacity of CO proteins to localize to a single site as on a normal *C. elegans* chromosome, independent of DSBs suggests that CO patterning mechanisms act exclusively with the interaction of proteins within the SC.

A family of SC-associated proteins further demonstrates that CO patterning is determined through the SC in *C. elegans*. These proteins known as the ZHPs (ZHP-1 through ZHP-4) are RING domain-containing proteins, with COs only forming at sites when ZHP-3 and ZHP-4 are present, and ZHP-1 and ZHP-2 act to concentrate ZHP-3/4 from throughout the SC to CO sites[17] (**Fig. 1C**). Notably, in the absence of either ZHP-1 or ZHP-2, ZHP-3/4 remain throughout the SC, and the CO marker COSA-1, becomes enriched at a multitude of sites along the SC[17]. Although most of these sites do not form COs, this demonstrates that ZHP proteins can determine CO patterning. Contrastingly, some evidence suggests that CO patterning can happen in the absence of the SC. This includes the capacity of *S. pombe* to form crossovers without an SC[18], and that *S. cerevisiae* seems to select CO sites prior to SC formation and forms COs at initiation sites of SC formation[19]. Based upon this observation in *S.* cerevisiae, it has been proposed that the presence of SC inhibits CO formation[20].

However, due to the dependence of CO patterning on the SC outlined above, an SC-dependent mechanism model could coherently explain many cases of CO patterning and its protein features. Additionally, the liquid crystalline dynamics and compartmentalized nature of the SC makes it an attractive medium to establish the pattern of COs and non-COs for each set of chromosomes. Key to understanding how a pattern of COs may be established is knowing which CO factors are dynamic and which are not. While mobile proteins can move throughout the SC and select sites for COs, static proteins are more likely to be recruited by proteins that have already established a pattern of COs and non-COs. In this study, by using the tractable meiotic live-imaging of the species *C. elegans*, in conjunction with multiple live-imaging strategies and analyses, we reveal the mobility of numerous CO proteins and the effects of concentration for some CO factors in establishing COs. Together, our results and known features of COs suggest a reaction-diffusion-like system, in which mobile factors diffuse throughout the SC, enrich pro-CO factors at CO sites, and deplete them from non-CO sites, establishes CO patterning.

METHODS AND MATERIALS

FCS immobilization, imaging, and analysis

FCS experiments were performed with 24-hour post-L4 *C. elegans* hermaphrodites. Worms were immobilized with serotonin and polystyrene beads as previously performed[12], however with the addition of 0.2% tricaine and 0.02% tetramisole in the mounting medium. Imaging was performed on a Zeiss LSM 710 with a 63× 1.4 NA oil objective. The "smart setup" filter set for the EGFP channel, and 2% laser power were used. The pinhole was set to image voxels of 1.1 µm thick, recording only a single SC. Imaging of two overlapping SCs would confound obtained measurements. Voxel size was set to the optimal setting with a line-scan zoom of 125× (1.08 µm line length). This resulted in eight voxels with x and y dimensions of 0.135 µm. For acquisition, maximum speed was used resulting in a line time of 473 µs with one thousand lines collected for a total time of 474 ms. These short acquisitions were chosen due to the movement of live *C. elegans* during imaging.

Matlab was used to calculate the autocorrelation function (ACF) of voxels using previously reported equations[21]. The ACF of individual voxels were analyzed, as well as the average of ACF for all voxels in a scan, which increases the total time[22]. Fits to these ACFs were performed in Matlab using a least-squares regression algorithm. A fluorescein solution was used to calculate the value of w_0, as previously reported[23], which determines the PSF and was found to have an average value of 0.29 µm. A two-dimensional diffusion model[24] best fit our scans and was used to estimate diffusion coefficients of voxel-based and average ACF. Paired correlation function (pCF) calculations were made in SimFCS at a distance of 0.135 µm for voxel pairs. The maximum value in the pCF at a given distance is the average time for fluorophores to move that distance[23]. The correlation time to reach the maximum pCF value was extracted and reported as "movement index" for all voxel pairs, as well as for the average pCF calculated from all voxel pairs within a line-scan.

FRAP immobilization, imaging, and analysis

FRAP experiments were all performed on 24-hour post-L4 *C. elegans* hermaphrodites. Worms were immobilized by picking them into a drop of M9 medium containing 0.4% tricaine and 0.04% tetramisole on top of a thin 2% agarose pad. This droplet was covered with a coverslip that was then sealed with VALAP (1:1:1 petroleum jelly: lanolin: paraffin). Nuclei in the gonad closer to the coverslip were chosen for imaging.

FRAP experiments were performed on a Marianas spinning-disc confocal microscope from 3i (Intelligent Imaging Innovations, Inc) using a 100× 1.46 NA oil immersion objective. Z-stacks were acquired with 0.25 µm spacing and a total thickness of 11 µm, and 50% laser power. Areas for photobleaching containing foci (or shorts-arms) of fluorescent protein were selected using the rectangular selection tool of Slidebook 6 software. Photobleaching was performed with a 488 nm laser at 100% power. An image stack was first acquired prior to photobleaching, with subsequent stacks acquired every minute. Halfway through the second image stack, photobleaching was performed and then twenty-three more stacks were acquired for a total imaging time of twenty-five minutes.

Analysis of FRAP images was performed in ImageJ on .TIF files exported from Slidebook 6. For each z-stack, a sum projection was made. The “Subtract Background” function with a rolling ball size of 5 pixels was applied to these projections to minimize background. Foci and short-arms areas were segmented by first using the “Default” threshold of the “Auto Threshold” function of ImageJ, followed by an increase in threshold until the area identified as the focus or short-arm included no diagonally connected pixels. Average fluorescence of identified focus or short-arm area were taken at every time point, and the area was manually realigned if lateral movement of a nucleus occurred. The average fluorescence obtained for every time point was divided by the average fluorescence of the first time point to yield the “relative fluorescence”.

ZHP-3 foci counting, intensity quantification, immobilization, and imaging

To count the number of ZHP-3::AID::GFP foci, and to quantify ZHP-3::AID::GFP foci brightness, live 24-hour post-L4 *C. elegans* hermaphrodites were immobilized in PBS with 20% glycerol and 20% PEG2000 on a microscopy slide with a coverslip.

Images of ZHP-3::AID::GFP were collected as z-stacks on a Marianas spinning-disc confocal microscope (Intelligent Imaging Innovations, Inc. (3i)) at intervals of 0.25 µm, a total thickness of 5.25 µm, with a 100× 1.46 N.A. oil-immersion objective, and both consistent laser power and exposure time in the 488 nm channel. Sum projections were made in ImageJ of the entire z-stacks for quantifying ZHP-3::AID::GFP foci maximum pixel brightness. Foci considered for quantification and counting were in late pachytene through diplotene according to previously reported ZHP-3 appearance[17]. Only nuclei completely encompassed within the z-stack were considered for counting or quantification. Maximum values of foci were found using ImageJ.

Auxin treatment

Auxin treatment to partially knockdown ZHP-3 in *zhp-3::AID::gfp*; $P_{gld\text{-}1}$*::TIR1::mRuby*; *Δzim-3, Δhim-8* was performed as follows: L4 worms were picked onto NGM plates containing 20 µM auxin and grown for 24 hours on these plates. This concentration of auxin was determined by partial knockdowns of similar proteins in another study[17], and was sufficient to decrease, but not eliminate ZHP-3::AID::GFP fluorescence. These plates were seeded with OP50 *E. coli* culture resuspended in minimal media.

CRISPR/Cas9 and strain generation

Split GFP tags (GFP11) on the C-termini of proteins at endogenous sites were made using previously established protocols for the CRISPR/Cas9 system[17] into a strain containing a germline expressed GFP1-10[25]. This protocol generated roller progeny from which wildtype progeny were chosen, PCR checked for DNA insertions, and then screened for fluorescence of protein in the SC to confirm correct insertion of GFP11. Repair template (ssDNA) and crRNA were provided by Integrated DNA Technologies (IDT), and sequences of them are available in Supplementary Information. Strain details are also available in Supplementary Information.

MosDel deletion of *zim-3, him-8*

The *Δzim-3, Δhim-8* deletion was made with the MosDel system[26]. The strain CA870[27] had the deletion generated with a donor plasmid. This plasmid was assembled with GateWay reaction and contained a 2.6 kb genomic fragment to the right of the *Mos1* site in CA870, a 3kb fragment to the left of *zim-3*, and the *C. briggsae unc-119*+ rescue fragment in pRL8. This plasmid was injected into CA870 hermaphrodites, along with pJL43.1, pGH8, pCFJ90, and pCFJ104[26]. Non-unc worms without mCherry were picked, and successful deletion of *zim-3* and *him-8* were confirmed with PCR.

RESULTS

Line-scan FCS to estimate diffusion coefficients of meiotic proteins

The SC is established to have at least some dynamic components. However, detailed analysis of the dynamics of individual proteins is still lacking. To uncover which components are dynamic, we sought to use a technique to estimate the diffusion coefficients of meiotic proteins within the SC. Many techniques exist to that can quantitatively describe the dynamics of biomolecules including single particle tracking (SPT), fluorescence recovery after photobleaching (FRAP), and fluorescence correlation spectroscopy (FCS), and assorted versions of these techniques[28]. Amongst these and other options, we determined that line-scan FCS would suit our goals and exhibited several key advantages for its use in the *C. elegans* gonad. First, line-scan FCS techniques are known to be relatively insensitive to photobleaching[29]. Second, it enables rapid data acquisition, which is important because immobilized live worms still undergo some movement[30], and the cells within the gonad also move relative to the animal as a whole. Third, rapid line-scans take advantage of the structure of the SC, refraining from unnecessary photobleaching by rapidly imaging only the region of interest, limiting the exposure of fluorescent proteins to light. Line-scan FCS has been used within the *C. elegans* gonad before[31,32], has been used on cell membranes, another liquid crystal[29], and allowed us to take advantage of multiple tagged meiotic proteins.

Validating diffusion estimates from ACF

We performed line-scan imaging (details in **Methods and Materials**) on proteins within the gonads of live immobilized *C. elegans* hermaphrodites. Using this data of rapid scans, we performed ACF analysis, as previously performed[31], to estimate diffusion coefficients of proteins. Measurements were taken in cells during late pachytene, chosen by imaging similar regions within the gonads of 24hr post L4s. To validate this approach, we began by imaging and analyzing HIS-72::GFP, a histone component. These estimates are summarized (**Fig. 2B**), and analogous analysis performed on the average autocorrelation function[22] of each line-scan have been included (**Fig. 2S**), although these results show effectively no difference from the voxel-based analysis. We found that HIS-72 had an average diffusion coefficient of 1.51×10^{-2} $\mu m^2/s$ with a standard error (SE) of 2.62×10^{-3} $\mu m^2/s$. Although not previously measured within meiotic cells in the pachytene stage, this relatively low diffusion coefficient agreed with estimations of histone diffusion[33-36]. It should be noted that the global movement of histones is also occurring within the background of other diffusion coefficient measurements, because the structures containing other measured proteins are

attached indirectly or directly to chromatin. This global movement could lead to an increase some diffusion coefficient measurements.

Previous studies by our lab using a photoconvertible fluorescent tag have shown that a protein within the meiotic axis, HTP-3, does not have significant movement[12]. The meiotic axis assembles onto chromosomes, and provides a scaffold for the SC to assemble onto, and contains multiple other proteins. To test if our ACF analysis agreed with our previous findings, and whether additional axis components are static, we performed the same ACF analysis on the axis components HTP-3 and HIM-3 (**Fig. 2B** and **Fig. 2S**). ACF analysis on these proteins yielded an average of 2.68×10^{-2} $\mu m^2/s$ with an SE of 6.26×10^{-3} $\mu m^2/s$ for HTP-3::GFP, and 4.87×10^{-2} $\mu m^2/s$ with an SE of 8.44×10^{-2} $\mu m^2/s$ for GFP::HIM-3. These relatively low diffusion coefficients are comparable to other diffusion coefficient measurements made of axis components using single particle tracking[37], and are consistent with the previous findings that HTP-3 within the axis does not have appreciable movement. Although the diffusion coefficients are higher than histone movement, this is possibly due to the addition of global chromosome movement to protein diffusion.

Measuring diffusion of SC structural components

Previous discoveries by our lab have shown SYP-3, a structural protein of the SC, to have movement throughout the SC using a photoconvertible tag[12], and other groups have also demonstrated this mobility using FRAP experiments[38]. We questioned whether additional known SC components display mobility, and what are the diffusion estimates of these components. To answer this, we applied performed the same FCS experiments and ACF analysis on GFP tagged versions of SYP-1, SYP-2, and SYP-3. Results of these diffusion coefficients are shown in **Fig. 2B** and **Fig. 2S**. All three proteins showed diffusion, which was appreciably higher than the histone and axis proteins. SYP-1 had an average diffusion coefficient of 3.39×10^{-1} $\mu m^2/s$ and an SE of 3.04×10^{-2} $\mu m^2/s$. SYP-2 had slightly lower diffusion, with an average of 1.76×10^{-1} $\mu m^2/s$ and an SE of 2.56×10^{-2} $\mu m^2/s$. Confirming that SYP-3 is mobile, it showed an average diffusion coefficient of 4.89×10^{-1} $\mu m^2/s$ and an SE of 2.70×10^{-2} $\mu m^2/s$. All three components with their coiled-coil domains were mobile with higher diffusion coefficients than the axis and histone proteins, which is consistent with evidence that these are discrete structures.

Other labs have found that early pachytene SYP-3 had more extensive recovery in FRAP experiments than SYP-3 in late pachytene[38]. Based on this result, it was concluded that the SC progresses from a more labile state in early pachytene, to a more stable state in late pachytene[38]. We examined whether this was supported by ACF analysis by performing the same measurements of SYP-3, but in early pachytene. Higher diffusion coefficients obtained from early pachytene than late pachytene SYP-3 would support the possibility of a transition of SC properties from labile to stable. Conversely, we found no significant change in the diffusion coefficient (**Table S1**), with an average diffusion coefficient of 4.17×10^{-1} $\mu m^2/s$ and an SE of 2.68×10^{-2} $\mu m^2/s$. Merging this result with FRAP results of SYP-3 suggests that the protein does not change its state of dynamics, and that the additional relative recovery of SYP-3 comes

from another source. One major possibility is that the additional FRAP recovery comes from nucleoplasmic SYP-3. GFP::SYP-3 is more concentrated in the SC and brighter in late pachytene[38], indicating more protein has loaded from the nucleoplasm, it may be this nucleoplasmic GFP::SYP-3 which is in the process of loading during early pachytene which leads to additional recovery.

Measuring a presumably dynamic crossover factor

Although we showed that SYP proteins are mobile, they function primarily to assemble the SC and recruit CO factors which in turn pattern COs. To learn more about possible ways the SC regulates COs, we sought to quantify the dynamics of mobile CO factors which exploit the liquid crystal properties of the SC. The RING finger proteins ZHP-1 through ZHP-4 have a dynamic pattern of localization within the SC that suggests mobility, and are essential for CO formation[17]. Because of this, we tested the mobility of ZHP-3 (**Fig. 2B** and **Fig. 2S**). This protein is thought to possibly be an E3 ubiquitin ligase[17] and as one of the ZHPs is an essential component in establishing COs. First, we measured the diffusion coefficient of ZHP-3 in mid-pachytene, while it is still throughout the entire length of SC along chromosomes and not yet concentrated to CO sites. The average diffusion coefficient was found to be 2.80×10^{-1} $\mu m^2/s$ with an SE of 2.01×10^{-2} $\mu m^2/s$. This is comparable to the SC proteins, and establishes that ZHP-3 is dynamic. To be consistent with other measured proteins, we also measured ZHP-3 in late pachytene. During this stage, ZHP-3 is enriched at localized to CO sites[17]. We anticipated the protein to go from a dynamic state to a stable, nonmobile state in which the attachment of ZHP-3 protein forms CO sites. Contrary to this expectation, ZHP-3 was shown to still be mobile while confined to CO sites, with an average diffusion coefficient of 2.03×10^{-1} $\mu m^2/s$ and SE of 1.44×10^{-2} $\mu m^2/s$. This finding reveals an important characteristic of the ZHP-3 protein. In order for it to form its essential foci at COs, it likely reaches an equilibrium of interactions with other factors to enrich and maintain enrichment at the CO site.

pCF analysis supports results and conclusions from ACF analysis

An additional analysis that can be performed with line-scan FCS data is pCF analysis[23]. This technique measures the average time it takes for a fluorescent particle to move a given distance[23]. These calculations are dissimilar to ACF analysis in measuring movement between spaces, and thus quantifying movement in a given direction instead of diffusion within any direction. We performed these calculations for the proteins identified as mobile, at a given distance of 0.135 μm, to quantify the local movement of proteins. This included SYP-1, SYP-2, SYP-3 in early and late pachytene, and ZHP-3 in middle and late pachytene. Results from this analysis are shown with pCF analysis performed for each set of voxels at a distance of 0.135 μm within a line-scan (**Fig. 3**). Alternatively, the same analysis has been performed for the average pCF function for all voxel pairs at a distance of 0.135 μm within a line-scan (**Fig. 3S**), but the results from both analysis strategies were consistent. The “movement index” as reported elsewhere[39], is the average time for a protein to move 0.135 μm, so a lower value indicates faster movement of a protein.

Several conclusions can be made from these results. First, SYP-3 does not have significantly different movement within the SC between early and late pachytene (**Table S2**), supporting the lack of a property change in terms of dynamics for the SC as pachytene progresses. Next, this fundamentally different analysis agrees with the results of ACF analysis, with the rank-order of proteins with most to least movement within the SC having the same order as proteins estimated to have the highest to lowest diffusion coefficients. Finally, this analysis reconfirms that when concentrated at CO sites, ZHP-3 moves. Further it shows that ZHP-3 actually moves throughout the site and doesn't just diffuse into and out of the site.

FRAP analysis of meiotic proteins

To corroborate that ZHP-3 is dynamic when confined to foci using a complementary technique, we used fluorescence recovery after photobleaching (FRAP) on live immobilized *C. elegans* hermaphrodite gonads. Because this technique does not necessitate the use of line-scans, which may be too large to capture the fluorescence of foci, we were also able to use it to characterize the dynamics of CO proteins with punctate localization. We performed FRAP (detailed in **Methods and Materials**) on ZHP-3 foci, as well as COSA-1 foci, both of which are essential to and demarcate CO sites[40]. Examples are shown in **Fig. 4A**. The example demonstrates that ZHP-3 recovers from photobleaching indicating movement into and out of CO sites, while COSA-1 does not recover, indicating a lack of movement. Also, in these images, prior to photobleaching, fluorescence for ZHP-3 is completely contained within foci. Seemingly, some portion of the recovery for the photobleached foci comes from other foci which are on separate SCs, although it does not confirm this possibility. In other words, because the fluorescence is only in ZHP-3 foci, it seems likely that protein exchanges into and out of each SC, allowing fluorescent protein in nonphotobleached ZHP-3 foci to migrate to the photobleached foci.

In **Fig. 4B**, recovery or non-recovery of proteins in FRAP are quantified. By normalizing average fluorescence of foci by their average initial fluorescence, proteins with variable brightness can be compared[41,42] as the relative fluorescence. ZHP-3 fluorescence recovers gradually following photobleaching, while there is a decrease in the fluorescence of nonphotobleached foci. This decrease is likely due to the photobleaching of fluorophore from imaging over acquisition time. It should also be considered that this photobleaching from imaging is also occurring in the photobleached foci, meaning that in its absence, the foci would show greater fluorescence recovery. ZHP-3 fluorescence is contained in foci on separate SCs, however it displays near complete recovery from photobleaching. This again indicates that exchange between ZHP-3 foci, and thus separate SCs for separate chromosomes may occur.

To compare the recovery of ZHP-3 to other meiotic proteins, we performed FRAP with quantification for a number of other proteins that form foci at CO sites within *C. elegans* meiosis (**Fig. 4B**). As in ZHP-3, these measurements were performed in late pachytene. To start, we measured the known CO markers COSA-1, then CDK-2, an ortholog of which was recently reported to be important in *Arabidopsis* meiosis and CO formation[43], and the MutSγ component MSH-5, which localizes to CO intermediates and COs[44]. All

three proteins did not have any recovery from FRAP, with the average relative fluorescence line for photobleached foci remaining flat. This indicates that COs are stable in terms of the dynamics of many essential proteins. Next, we tested the ssDNA binding protein RPA-1[45], and found that this protein exhibited very modest recovery which can be seen in the slight positive trend of the average relative fluorescence. It should be noted that we also tested a balanced, fluorescently tagged version of RAD-51[45], but found that it was likely overexpressed as it never formed foci, so we excluded this analysis. Of all the meiotic proteins, only ZHP-3 had recovery signifying mobility of components into and out of foci, although RPA-1 did slightly recover. Overall, it seems ZHP-3 is unique amongst CO proteins in its dynamic properties while in foci, and that many known CO factors are for the most part static at CO sites.

FRAP with split GFP shows a family of RING domain CO factors are all dynamic

After noting that ZHP-3 is unique amongst many CO proteins for being dynamic, we questioned whether its RING domain containing paralogs which have known interactions with it (**Fig. 1C**) and are essential in CO formation, were also dynamic. Initially, we tagged the paralogs with GFP on the C-terminus but found that the resulting proteins were nonfunctional. Because of this, we elected to use split GFP, which had recently been developed for use in the *C. elegans* gonad[25]. We tagged ZHP-1, ZHP-2, ZHP-3, and ZHP-4 with the short GFP11 sequence on the C-terminus, which permanently associates with free GFP1-10 within the gonad to fluorescently label a protein. We intended to also measure the split GFP tagged ZHPs with FCS, but the signal and photostability were found to be too low for those uses.

We measured ZHP-1 and ZHP-2 using the same FRAP method as other proteins with one exception. In late pachytene, instead of localizing to CO sites as ZHP-3/4, ZHP-1/2 localize to a length of the SC termed the “short-arm”, and this whole area was photobleached and measured for recovery. Because of this difference, the comparison is not completely analogous. Both ZHP-1 and ZHP-2 showed recovery. As seen in **Fig. 4B**, the photobleached short-arms of ZHP-1/2 recovered with average relative fluorescence increasing, while nonphotobleached ZHP-1 and ZHP-2 short-arms decreased substantially. The rapid decrease in fluorescence in nonphotobleached sample is due to the lower photostability of split GFP versus GFP[46], and photobleaching of split GFP occurring during imaging. Photobleached short-arms in both ZHP-1/2 also decrease after they recover, leading to an increase and then decrease in fluorescence. This FRAP recovery reveals ZHP-1 and ZHP-2 as dynamic in addition to ZHP-3.

We measured ZHP-3 and ZHP-4 tagged with split GFP in CO foci (**Fig. 4B**). In the case of ZHP-3, we already knew it was dynamic due to measurements performed on full-length GFP tagged ZHP-3. However, we wanted to confirm that with split GFP it also recovered on a similar scale to full-length GFP. When measuring ZHP-3 with split GFP, we again found that it recovered from photobleaching, but also had significant photobleaching during imaging, leading to recovery and then a decrease in fluorescence with continued imaging. Overall, the recovery of ZHP-3 split GFP fluorescence takes place over minutes, similar to the recovery of full-length GFP. Similar to the other ZHP proteins, ZHP-4 also recovers during FRAP (**Fig. 4B**), but

again has the characteristic decrease in fluorescence from imaging of both photobleached and nonphotobleached foci. Together, our FRAP results show that all ZHP proteins are dynamic when within foci or the short-arm of the SC.

Altering ZHP concentration alters CO patterning

Since the ZHPs are dynamic within the SC and have known collaboration to concentrate ZHP-3/4 at CO sites, we reasoned that the concentration must be important in establishing this pattern. In other words, altering their concentrations may alter the capacities of ZHP-1/2 or ZHP-3/4 to promote concentration or to concentrate at CO sites respectively, and, therefore, modify CO patterning. Further, both mammalian and plant orthologs of ZHP-3 have been shown to have a dosage-dependence on CO formation number[47-49]. We pursued a strategy to increase ZHP concentration by decreasing the amount of SC available. As depicted in **Fig. 5A**, we used *a Δzim-3, Δhim-8* mutant which lacks pairing between the I, IV, and X chromosomes. A lack of pairing between half of the chromosomes results in only three SCs forming in meiotic cells, which is in contrast to the normal six. While there are fewer SCs in this mutant, ZHP protein is likely regulated at a cellular level, making more ZHP proteins available for the SC on the II, III, and V chromosomes.

To assess this approach to increase ZHP concentration, we crossed ZHP-3 with a GFP tag into the *Δzim-3, Δhim-8* background. We quantified the maximum fluorescence value of ZHP-3 foci in late pachytene/diplotene for live *C. elegans* hermaphrodites (**Fig. 5B**). While it did not reveal the total difference in protein level, a higher maximum fluorescence indicates a higher concentration of ZHP-3 at the most concentrated region of the focus. ZHP-3 fluorescence values of the mutant had significantly higher fluorescence than the wildtype background ($p=7.34\times10^{-12}$; Mann-Whitney U test). This suggests that the *Δzim-3, Δhim-8* mutant indeed has additional ZHP available for concentration in the SC.

To quantify how an increase in available ZHPs (1-4) would impact CO number, we counted ZHP-3 foci of the mutant and wildtype background to see how the additional ZHPs would designate COs. By taking the number of crossovers per cell and then dividing that number by the number of SCs, we obtain the ZHP-3 foci number per synapsed chromosomes. For example, six ZHP-3 foci in a mutant background will have two foci per chromosome, while six ZHP-3 foci in a wildtype background will have one focus per chromosome. We found that the mutant had a variable number of COs with sometimes having additional or fewer COs per chromosome, while the wildtype nearly always had one CO per chromosome (**Fig. 5C**). This difference was significant ($p=1.20\times10^{-3}$; Mann-Whitney U test), and reveals instability in CO patterning with an increase in ZHP concentration. This instability actually is consistent with what we now know about ZHPs. Firstly, they are mobile and interact. If the concentration of all ZHPs is increased as here, they may begin to concentrate ZHP-3/4 at additional sites. Additionally, the increased concentration of factors may result in ZHP-3/4 at no sites because the additional ZHP-1/2 are able to inhibit their recruitment to any sites. Although the mobility and interactions of the ZHPs has remained the same, the

concentration of all these factors is higher, and destabilized a multicomponent concentration dependent pattern.

To extend this indication of concentration influencing the patterning of COs, we performed a partial knockdown of ZHP-3 using the auxin-inducible degron (AID) system[50] in the mutant background. By using a low dose of auxin, we only partially degraded ZHP-3::AID::GFP, and were able to still count the number of foci. Due to ZHP-3's function in a heterodimer, this decrease in effect also decreases ZHP-4. We hypothesized this decrease in ZHP-3 concentration may shift the CO patterning towards wildtype results. However, the partial knockdown of ZHP-3 only eliminated the cases where more than one CO per SC were formed (**Fig. 5D**) and significantly increased cases with less than one ZHP-3 focus per SC ($p=1.34\times10^{-10}$; Mann-Whitney U test) when compared to the mutant without depletions. This reflects an inability of ZHP-3 to concentrate at one site, if its concentration becomes decreased, while the inhibitory ZHP-1/2 likely remained at increased concentrations.

A major caveat in experiments with the *Δzim-3, Δhim-8* mutant is that it has a delay in meiosis, which can disturb meiosis checkpoints and events[51]. The shift in COs per chromosome we observed in the *Δzim-3, Δhim-8* mutant could have been due to some effect of meiotic delay. To examine this possibility, we used two different biological controls. If they also had a destabilization of CO number, that would suggest that the *Δzim-3, Δhim-8* mutant destabilization of CO number are not due to ZHP concentration. First, HIM-5 is a protein with a role in formation of DSBs on the X chromosome, and its loss results in a delay in meiosis[52]. However, its loss results in minimal other differences from wildtype meiosis and thus was excellent to observe the specific effects of meiotic delay on CO number per chromosome. We performed an analogous experiment counting the ZHP-3 foci of a *Δhim-5* mutant, and despite the meiotic delay, the *Δhim-5* mutant did not recapitulate the unstable number of COs in the *Δzim-3, Δhim-8* mutant (**Fig. 5E**) ($p=5.23\times10^{-12}$; Mann-Whitney U test). Second, we performed parallel experiments in the meT7 background, which has a chromosomal fusion between chromosomes III, X, and IV[53]. Not only did this strain have a meiotic delay, it also only has four SC compartments, although with one compartment being roughly as long as three SC compartments. Because of this, we could examine if the perturbations in *Δzim-3, Δhim-8* mutant ZHP-3 foci number is also possibly due to a variable number of SCs and some unknown mechanism. However, again this strain was not able to recapitulate a similar variation of ZHP-3 foci number with its majority of COs still at one. There was some variation in CO number but it was still significantly different from the *Δzim-3, Δhim-8* mutant ($p=5.20\times10^{-3}$; Mann-Whitney U test) and not as severe. Although our results do not exclude any influence of meiotic delay in affecting CO regulation, they do show that *Δzim-3, Δhim-8* mutants have higher concentrations of ZHPs (**Fig. 5B**), unstable CO numbers (**Fig. 5C**), and that CO patterning is sensitive to concentration of all the ZHPs (**Fig. 5D**).

DISCUSSION

CO designation occurs through a robust system in *C. elegans* which consistently forms a single CO per pair of homologs from among multiple DSBs. Towards revealing how this regulation occurs, we have characterized the dynamics of numerous CO and meiosis proteins. With multiple techniques, we have confirmed that the CO factors ZHP-1 through ZHP-4 are dynamic and that they are unique amongst CO factors in this quality. Additionally, we have confirmed a sensitivity of the system in CO and non-CO patterning due to the concentration of these factors within the SC. Synthesizing the results summarized here and known features of COs suggests a novel model of how COs are patterned amongst DSBs.

Based on the interactions and dynamic properties of ZHPs, class I COs may be patterned by a reaction-diffusion (R-D) system. These systems are a means for biological patterns to form by mobile components and are proposed to form patterns in a number of biological systems[54]. R-D systems involve the diffusion and interaction of "morphogens". The first morphogen is termed an "activator", has positive feedback to recruit itself, and, in this case, are ZHP-3/4. The second morphogen is termed an "inhibitor", inhibits the presence of the activator, and, in this case, are ZHP-1/2 acting along the SC. By this model, it is the diffusion shown here and interaction of ZHPs that generates a pattern[17] of one CO in the SC. In such a system, ZHPs would move throughout the SC acting to promote localization of ZHP-3/4 to potential CO sites. As ZHP-3/4 localizes to CO sites with positive feedback, its limited concentration in the cell and SC, as well as the action of ZHP-1/2 to inhibit its presence along the SC, would produce CO interference by resulting in the growth of ZHP-3/4 at a single CO site and loss from any other potential CO sites. In terms of the SC as a phase-separated compartment, this framework would classify SYP proteins as "scaffolds" which function to form the SC, and ZHPs as "clients" which are concentrated into the SC to perform the function which requires dynamics as well as concentration[55]. Such a system is not a new idea with similar hypotheses formed decades prior to the discovery of the liquid crystal qualities of the SC. Robin Holliday hypothesized the existence of a protein that "stabilizes crossovers", of limited concentration within cells, that once in the "SC diffuse along the length" and binds cooperatively to sites "making it unlikely that there can be stabilization of any other crossover nearby."[56] Some of these ideas even predate Holliday with Michael White noting that CO formation "depends on some kind of key molecules (enzymes or DNA) which are in short supply in the nucleus."[57]. More recently, as noted before, a "physical model" using theoretical CO factors and reaction-diffusion simulations has been shown to recapitulate CO interference in *Drosophila*[6].

Our dynamics findings support an R-D model of CO designation and several other significant discoveries are consistent with this model. First, in both heterozygous mutants of RNF212 (a mammalian ZHP-3 ortholog)[47] and partial AID depletions of ZHP-3, we see a decrease in CO formation, indicating that some homologs do not form COs. We have shown that ZHP-3 likely exchanges between SCs within the cell with FRAP, and with the positive feedback of ZHP-3 or RNF212 to establish COs, once some sites start to accumulate ZHP-3/4 through the R-D system, then COs will form on some chromosomes, but there will be no formation on other chromosomes. Second, an R-D

system with exchange and equilibrium between CO foci of ZHP-3 results in the homogenous amounts of protein between COs of different chromosomes. This phenomenon with SCs of drastically variable lengths, and thus drastically different initial amounts of ZHP-3 ortholog, can be observed in electron micrographs of orthologous systems[58]. Third, an R-D system explains the non-homogenous appearance of ZHP-3/4 throughout the SC in pachytene[17], which could represent intermediates of increased ZHP-3/4 or fluctuations in the pattern formation (i.e. enrichment due to the interactions as the pattern is forming), and the observation that RNF212 in mammals concentrates at CO intermediate sites[59]. Fourth, as noted in the introduction, RNAi to partially knockdown SYP-1 can result in multiple COs per pair of homologous chromosomes. In an R-D model, two explanations of this are possible; either small breaks in the SC allow two compartments to form two COs with an individual pattern for each compartment, or dynamics of the ZHPs are changed as such to allow a different pattern of two COs per pair of homologs to form. Fifth, the length of the SC affects CO number. A longer SC with the same interactions and dynamics could establish a pattern with additional COs. Sixth, in the instance that ZHP-1/2 are not present, ZHP-3/4 are unable to concentrate at CO sites, and the recruitment of COSA-1 to multiple sites along the SC. Finally, an R-D system explains the ability of CO designation to occur within polycomplex, as ZHPs within a polycomplex compartment could still interact and establish a single site to enrich CO proteins. While our model is compatible with these findings, and previous CO mechanism models are not, it should be noted that they may still contribute to CO patterning in conjunction with an R-D system of ZHPs. By integrating multiple systems, CO designation could be the robust system that we see, with almost always one CO per homologs in *C. elegans*.

Similarly to the potential of mechanism models involving tension within DNA or others contributing to CO patterning, it's clear that COSA-1, a cyclin-like protein, and any of its acting partners work as a tandem system to reinforce and then designate COs. This is based on the dependence of both ZHP-3 localization to CO sites on COSA-1[40] and the dependence for patterning of COSA-1 foci upon ZHP-3/4[17]. Also, both COSA-1[60] and ZHP-3[17] have been shown to be indispensable to CO intermediate formation. This co-dependence of factors indicates that they work in parallel to designate COs, and are not part of the same system as ZHP-3/4 ultimately determine where COSA-1 sites form. Redundant systems that work to both select and reinforce COs are likely an effective means to increase the robustness of the systems. For example, small perturbations in interactions or concentrations of ZHPs could be stabilized by a normal system of COSA-1. Notably, in contrast to ZHPs, COSA-1 was relatively stable with no dynamic behavior indicated by FRAP experiments. This difference reflects diversity in the redundant systems which reinforce and then form COs.

Although these and the previous experiments of others suggest an R-D system of CO designation, multiple lines of investigation would reinforce the applicability of the model. Biochemistry to determine the activities of ZHPs would help to clarify and remove unknowns from the system. Overexpression of specific ZHPs would be of interest because although this study increases ZHPs within the SC, it is impossible to exclude

pleiotropic effects of the mutations needed to achieve this. While the effects of decreasing ZHP-3 concentration are shown here, the increased concentrations would have predictable effects of either increasing (ZHP-3/4) or decreasing (ZHP-1/2) CO number. Unfortunately, at this time no widespread system for reliable overexpression of meiosis proteins in *C. elegans* is known, and with the demonstrated dosage dependence of ZHP-3, there are likely unknown mechanisms to regulate cellular concentration that would be difficult to circumvent. Finally, a mathematical model could test some of the conclusions and potential of an R-D system to pattern COs.

In conclusion, in this study we have quantified the dynamics of many known meiosis, SC, and CO proteins. We have identified the ZHPs as client molecules in the SC, unique in their combination of dynamics and relevance to COs. Additionally, we have confirmed the dosage dependence of ZHPs in *C. elegans*. These findings suggest a new model of CO designation which synthesizes these findings along with known features of the SC and CO formation. Future experimental work with additional manipulations to CO designation will further evaluate the utility of this model.

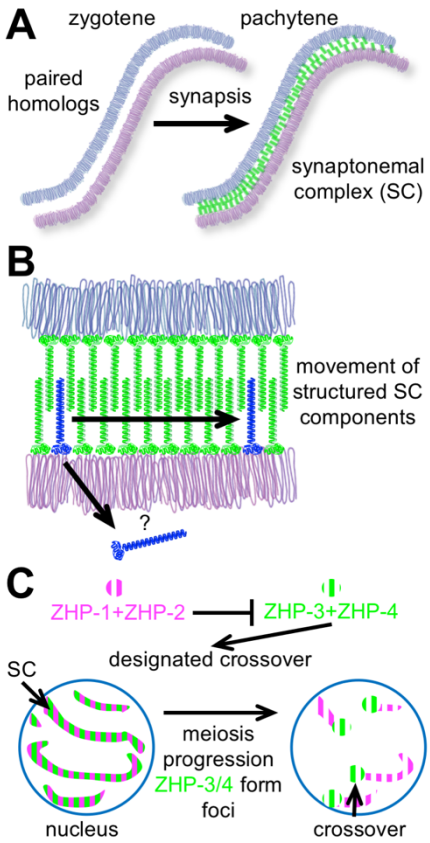

Figure 1: The synaptonemal complex (SC) has liquid crystalline properties, and ZHP proteins localize in CO designation. (**A**) Representation of the synaptonemal complex forming between homologs. The SC is a compartment with ordered subunits between these chromosomes. (**B**) Depiction of SYP proteins organized within the SC. As shown, each subunit has mobility within the SC. Whether SC components can move out of the

SC is unknown. Note: this simplified model does not demonstrate the diversity of SC components, or include axis proteins. (**C**) The interactions and spatial localization of ZHP are shown. With these interactions and resulting localizations, ZHPs regulate CO formation.

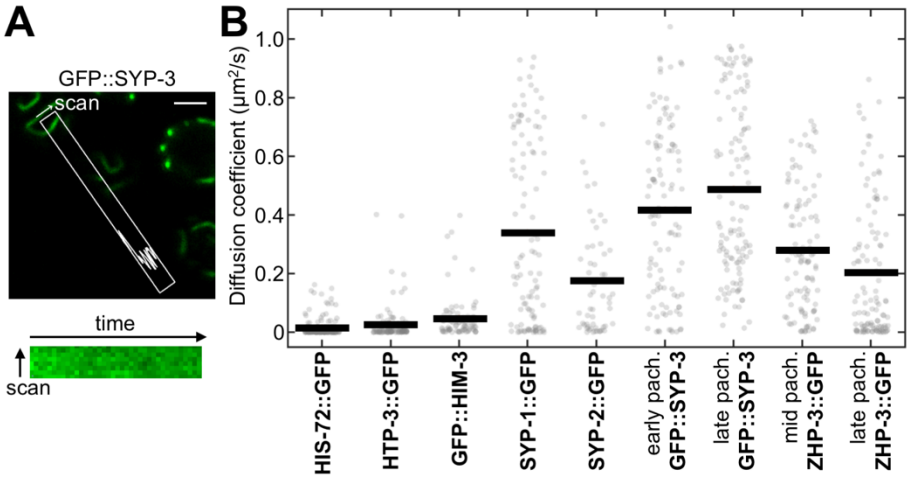

Figure 2: FCS line-scans are used to estimate diffusion coefficients of proteins in *C. elegans* gonads. (**A**) An example line-scan selection of GFP::SYP-3 for an SC within a late pachytene cell of a live immobilized *C. elegans* gonad. Image acquisition scans along this selection to obtain intensity values over time. The scale bar shown is 2 μm. (**B**) Diffusion coefficients obtained by fitting to the ACF of single voxels. Proteins are listed with fluorescent tags in their column. Each grey point is an obtained value, and the black bar is the average value for each protein. N numbers and statistical comparisons are available within **Table S1**. Analogous analyses for average ACF of entire line-scans are **Fig. 2S** and **Table S3**.

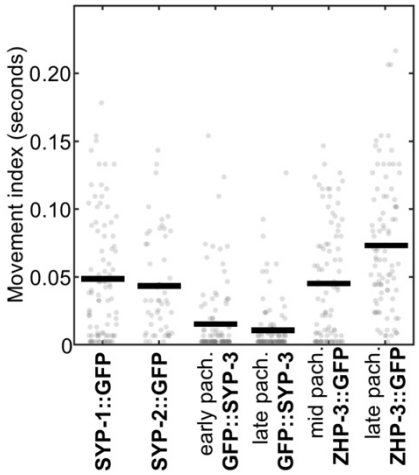

Figure 3: FCS line-scans are used to estimate the movement of proteins within the SC. Line-scan data analyzed is the same as in **Fig. 2**. Time for proteins to move 0.135 μm is estimated as the "movement index" by calculating the pCF and finding the correlation time at its maximum between a pair of voxels. Proteins are listed with fluorescent tags in

their column. Each grey point is a movement index obtained from a pair of voxels, and the black bar is the mean value. N numbers and statistical comparisons are available within **Table S2**. Analogous analyses for average pCF of all voxel pairs along the scan are **Fig. 3S** and **Table S4**.

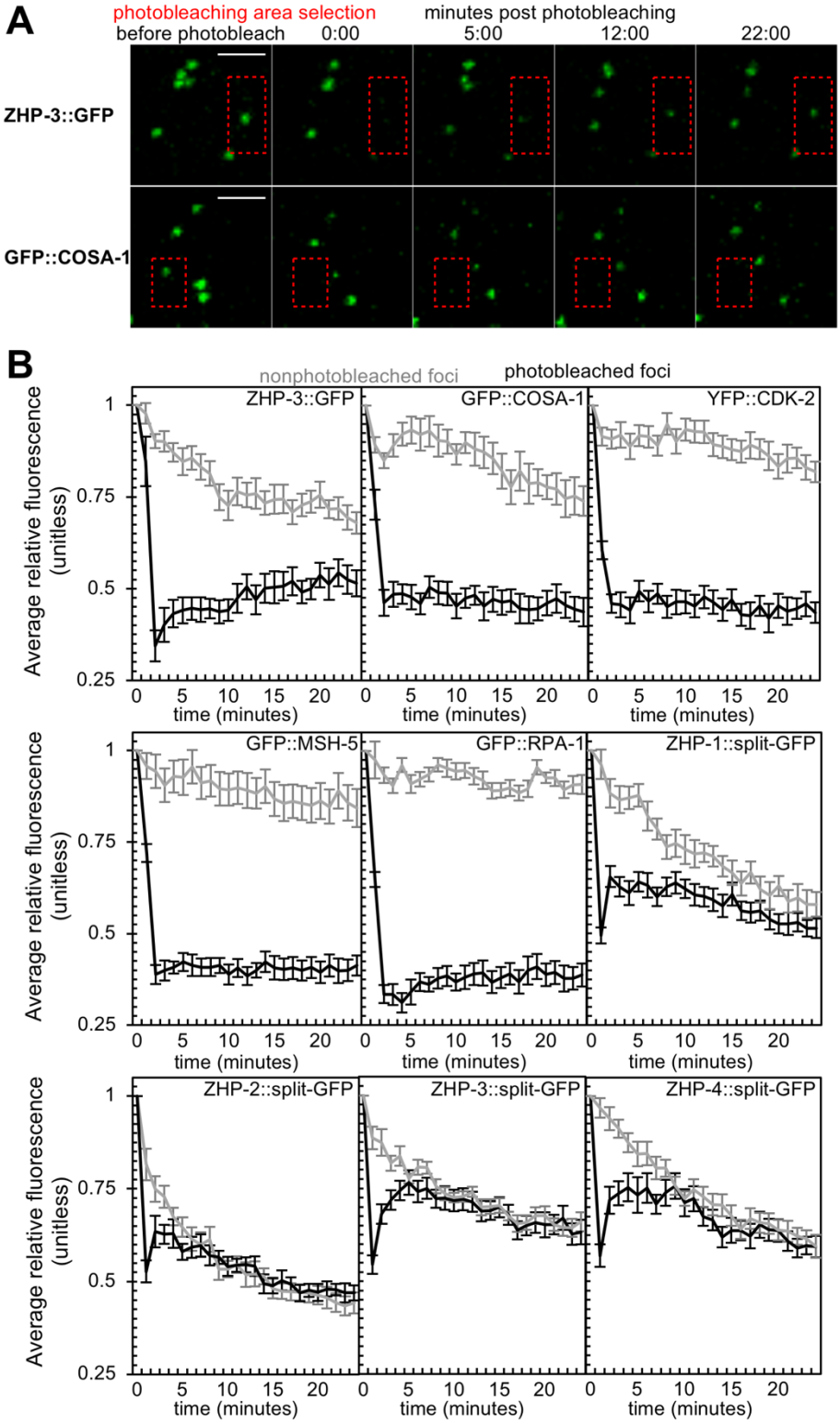

Figure 4: FRAP analysis shows dynamics of puncta forming CO proteins in live *C. elegans* gonads. (**A**) Example maximum projections of FRAP performed on ZHP-3 and COSA-1, showing recovery of ZHP-3 and no recovery of COSA-1 fluorescence. The scale bar shown is 2 μm. (**B**) The average relative fluorescence calculated for photobleached foci (or short-arms) fluorescence. Nonphotobleached foci (or short-arms) are a control for photobleaching due to ongoing imaging. Lines are the average value of the relative fluorescence, and error bars are standard error of the relative fluorescence. N numbers for foci number are as follows. ZHP-3::GFP photobleached, n=17. ZHP-3::GFP nonphotobleached, n=18. GFP::COSA-1 photobleached, n=21. GFP::COSA-1 nonphotobleached, n=19. YFP::CDK-2 photobleached, n=18. YFP::CDK-2 nonphotobleached, n=20. GFP::MSH-5 photobleached, n=20. GFP::MSH-5 nonphotobleached, n=19. GFP::RPA-1 photobleached, n=19. GFP::RPA-1 nonphotobleached, n=20. ZHP-1::split-GFP photobleached, n=15. ZHP-1::split-GFP nonphotobleached, n=20. ZHP-2::split-GFP photobleached, n=13. ZHP-2::split-GFP nonphotobleached, n=13. ZHP-3::split-GFP photobleached, n=15. ZHP-3::split-GFP nonphotobleached, n=19. ZHP-4::split-GFP photobleached, n=14. ZHP-4::split-GFP nonphotobleached, n=15.

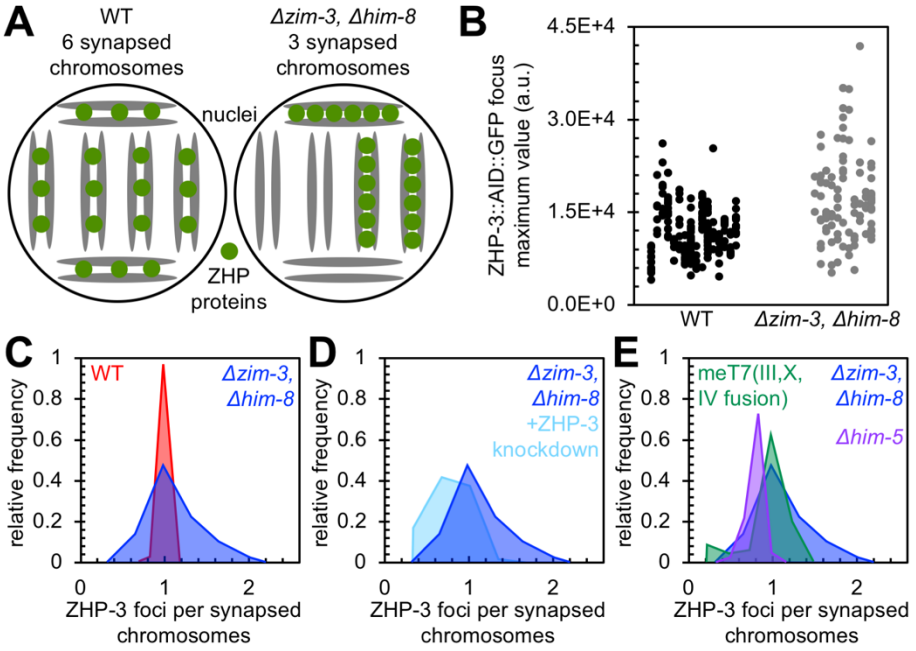

Figure 5: Altering the concentration of ZHP proteins with mutants and auxin-inducible degradation demonstrates their concentration dependence. (**A**) A depiction of how ZHP protein concentration is expected to be affected by altering chromosome pairing. The *Δzim-3, Δhim-8* cell has the same cellular concentration but fewer compartments (SCs). (**B**) Maximum pixel values of ZHP-3 foci for sum projections of wildtype (WT) and *Δzim-3, Δhim-8* mutants. In each row are foci from the gonad closer to the coverslip of a single worm. Data from each worm type come from four different sample mountings. (**C**) Plot showing the relative frequency of ZHP-3 foci per synapsed chromosomes for late pachytene and diplotene nuclei. Each possible value has been connected by a line and the area filled beneath. For the WT background n=72 nuclei, and for *Δzim-3, Δhim-8* background n=95 nuclei. (**D**) Plot as in C for the same *Δzim-3, Δhim-8* data, as well as

the *Δzim-3, Δhim-8* background with a ZHP-3 partial depletion using AID. For the *Δzim-3, Δhim-8* with partial AID depletion n=52 nuclei. (**D**) Plot as in C for the same *Δzim-3, Δhim-8* data, as well as a *Δhim-5* background and a meT7 background. For the *Δhim-5* background n=42 nuclei, and for the meT7 background n=46 nuclei.

REFERENCES

1 Storlazzi, A., Xu, L., Cao, L. & Kleckner, N. Crossover and noncrossover recombination during meiosis: timing and pathway relationships. *Proceedings of the National Academy of Sciences* **92**, 8512-8516, doi:10.1073/pnas.92.18.8512 (1995).

2 Saito, T. T. & Colaiácovo, M. P. Regulation of Crossover Frequency and Distribution during Meiotic Recombination. *Cold Spring Harb Symp Quant Biol* **82**, 223-234, doi:10.1101/sqb.2017.82.034132 (2017).

3 Shinohara, M., Oh, S. D., Hunter, N. & Shinohara, A. Crossover assurance and crossover interference are distinctly regulated by the ZMM proteins during yeast meiosis. *Nature Genetics* **40**, 299-309, doi:10.1038/ng.83 (2008).

4 King, J. S. & Mortimer, R. K. A polymerization model of chiasma interference and corresponding computer simulation. *Genetics* **126**, 1127-1138 (1990).

5 Kaback, D. B., Barber, D., Mahon, J., Lamb, J. & You, J. Chromosome size-dependent control of meiotic reciprocal recombination in Saccharomyces cerevisiae: the role of crossover interference. *Genetics* **152**, 1475-1486 (1999).

6 Fujitani, Y., Mori, S. & Kobayashi, I. A Reaction-Diffusion Model for Interference in Meiotic Crossing Over. *Genetics* **161**, 365 (2002).

7 Hultén, M. A. On the origin of crossover interference: A chromosome oscillatory movement (COM) model. *Molecular Cytogenetics* **4**, 10, doi:10.1186/1755-8166-4-10 (2011).

8 Alleva, B. & Smolikove, S. Moving and stopping: Regulation of chromosome movement to promote meiotic chromosome pairing and synapsis. *Nucleus* **8**, 613-624, doi:10.1080/19491034.2017.1358329 (2017).

9 Wang, S., Zickler, D., Kleckner, N. & Zhang, L. Meiotic crossover patterns: obligatory crossover, interference and homeostasis in a single process. *Cell Cycle* **14**, 305-314, doi:10.4161/15384101.2014.991185 (2015).

10 Schild-Prüfert, K. *et al.* Organization of the synaptonemal complex during meiosis in Caenorhabditis elegans. *Genetics* **189**, 411-421, doi:10.1534/genetics.111.132431 (2011).

11 Moses, M. J. SYNAPTINEMAL COMPLEX. *Annual Review of Genetics* **2**, 363-412, doi:10.1146/annurev.ge.02.120168.002051 (1968).

12 Rog, O., Kohler, S. & Dernburg, A. F. The synaptonemal complex has liquid crystalline properties and spatially regulates meiotic recombination factors. *Elife* **6**, doi:10.7554/eLife.21455 (2017).

13 Libuda, D. E., Uzawa, S., Meyer, B. J. & Villeneuve, A. M. Meiotic chromosome structures constrain and respond to designation of crossover sites. *Nature* **502**, 703-706, doi:10.1038/nature12577 (2013).

14 Codina-Pascual, M. *et al.* Crossover frequency and synaptonemal complex length: their variability and effects on human male meiosis. *Molecular Human Reproduction* **12**, 123-133, doi:10.1093/molehr/gal007 (2006).
15 Lynn, A. *et al.* Covariation of Synaptonemal Complex Length and Mammalian Meiotic Exchange Rates. *Science* **296**, 2222, doi:10.1126/science.1071220 (2002).
16 Kleckner, N., Storlazzi, A. & Zickler, D. Coordinate variation in meiotic pachytene SC length and total crossover/chiasma frequency under conditions of constant DNA length. *Trends in Genetics* **19**, 623-628, doi:https://doi.org/10.1016/j.tig.2003.09.004 (2003).
17 Zhang, L., Kohler, S., Rillo-Bohn, R. & Dernburg, A. F. A compartmentalized signaling network mediates crossover control in meiosis. *Elife* **7**, doi:10.7554/eLife.30789 (2018).
18 Cromie, G. & Smith, G. R. Meiotic Recombination in Schizosaccharomyces pombe: A Paradigm for Genetic and Molecular Analysis. *Genome Dyn Stab* **3**, 195-195, doi:10.1007/7050_2007_025 (2008).
19 Fung, J. C., Rockmill, B., Odell, M. & Roeder, G. S. Imposition of Crossover Interference through the Nonrandom Distribution of Synapsis Initiation Complexes. *Cell* **116**, 795-802, doi:https://doi.org/10.1016/S0092-8674(04)00249-1 (2004).
20 Egel, R. Synaptonemal complex and crossing-over: structural support or interference? *Heredity (Edinb)* **41**, 233-237, doi:10.1038/hdy.1978.92 (1978).
21 Wachsmuth, M., Waldeck, W. & Langowski, J. Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J Mol Biol* **298**, 677-689, doi:10.1006/jmbi.2000.3692 (2000).
22 Qian, H., Elson, E. L. & Frieden, C. Studies on the structure of actin gels using time correlation spectroscopy of fluorescent beads. *Biophysical journal* **63**, 1000-1010, doi:10.1016/S0006-3495(92)81686-7 (1992).
23 Digman, M. A. & Gratton, E. Imaging barriers to diffusion by pair correlation functions. *Biophys J* **97**, 665-673, doi:10.1016/j.bpj.2009.04.048 (2009).
24 Vukojevic, V. *et al.* Study of molecular events in cells by fluorescence correlation spectroscopy. *Cell Mol Life Sci* **62**, 535-550, doi:10.1007/s00018-004-4305-7 (2005).
25 Hefel, A. & Smolikove, S. Tissue-Specific Split sfGFP System for Streamlined Expression of GFP Tagged Proteins in the Caenorhabditis elegans Germline. *G3 (Bethesda)* **9**, 1933-1943, doi:10.1534/g3.119.400162 (2019).
26 Frøkjær-Jensen, C. *et al.* Targeted gene deletions in C. elegans using transposon excision. *Nature Methods* **7**, 451-453, doi:10.1038/nmeth.1454 (2010).
27 Harper, N. C. *et al.* Pairing centers recruit a Polo-like kinase to orchestrate meiotic chromosome dynamics in C. elegans. *Developmental cell* **21**, 934-947, doi:10.1016/j.devcel.2011.09.001 (2011).
28 Guo, L. *et al.* Molecular Diffusion Measurement in Lipid Bilayers over Wide Concentration Ranges: A Comparative Study. *ChemPhysChem* **9**, 721-728, doi:10.1002/cphc.200700611 (2008).

29 Ries, J., Chiantia, S. & Schwille, P. Accurate determination of membrane dynamics with line-scan FCS. *Biophys J* **96**, 1999-2008, doi:10.1016/j.bpj.2008.12.3888 (2009).
30 Berger, S. *et al.* Long-term C. elegans immobilization enables high resolution developmental studies in vivo. *Lab on a Chip* **18**, 1359-1368, doi:10.1039/C7LC01185G (2018).
31 Hinde, E. *et al.* The impact of mitotic versus interphase chromatin architecture on the molecular flow of EGFP by pair correlation analysis. *Biophys J* **100**, 1829-1836, doi:10.1016/j.bpj.2011.02.024 (2011).
32 Hinde, E. & Cardarelli, F. Measuring the flow of molecules in cells. *Biophys Rev* **3**, 119-119, doi:10.1007/s12551-011-0051-x (2011).
33 Görisch, S. M., Lichter, P. & Rippe, K. Mobility of multi-subunit complexes in the nucleus: accessibility and dynamics of chromatin subcompartments. *Histochemistry and Cell Biology* **123**, 217-228, doi:10.1007/s00418-005-0752-y (2005).
34 Vazquez, J., Belmont, A. S. & Sedat, J. W. Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. *Curr Biol* **11**, 1227-1239 (2001).
35 Miné-Hattab, J. & Rothstein, R. Increased chromosome mobility facilitates homology search during recombination. *Nature Cell Biology* **14**, 510-517, doi:10.1038/ncb2472 (2012).
36 Hajjoul, H., Kocanova, S., Lassadi, I., Bystricky, K. & Bancaud, A. Lab-on-Chip for fast 3D particle tracking in living cells. *Lab Chip* **9**, 3054-3058, doi:10.1039/b909016a (2009).
37 Syrjanen, J. L. *et al.* Single-molecule observation of DNA compaction by meiotic protein SYCP3. *Elife* **6**, doi:10.7554/eLife.22582 (2017).
38 Pattabiraman, D., Roelens, B., Woglar, A. & Villeneuve, A. M. Meiotic recombination modulates the structure and dynamics of the synaptonemal complex during C. elegans meiosis. *PLoS Genet* **13**, e1006670, doi:10.1371/journal.pgen.1006670 (2017).
39 Clark, N. M. *et al.* Tracking transcription factor mobility and interaction in Arabidopsis roots with fluorescence correlation spectroscopy. *eLife* **5**, e14770, doi:10.7554/eLife.14770 (2016).
40 Yokoo, R. *et al.* COSA-1 reveals robust homeostasis and separable licensing and reinforcement steps governing meiotic crossovers. *Cell* **149**, 75-87, doi:10.1016/j.cell.2012.01.052 (2012).
41 Vulprecht, J. *et al.* STIL is required for centriole duplication in human cells. *Journal of Cell Science* **125**, 1353, doi:10.1242/jcs.104109 (2012).
42 Mahjoub, M. R., Xie, Z. & Stearns, T. Cep120 is asymmetrically localized to the daughter centriole and is essential for centriole assembly. *The Journal of cell biology* **191**, 331-346, doi:10.1083/jcb.201003009 (2010).
43 Wijnker, E. *et al.* The Cdk1/Cdk2 homolog CDKA;1 controls the recombination landscape in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **116**, 12534, doi:10.1073/pnas.1820753116 (2019).
44 Zetka, M. Homologue Pairing, Recombination and Segregation in Caenorhabditis elegans. *Genome Dynamics* **5**, 43-55, doi:10.1159/000166618 (2009).

45 Koury, E., Harrell, K. & Smolikove, S. Differential RPA-1 and RAD-51 recruitment in vivo throughout the C. elegans germline, as revealed by laser microirradiation. *Nucleic acids research* **46**, 748-764, doi:10.1093/nar/gkx1243 (2018).
46 Köker, T., Fernandez, A. & Pinaud, F. Characterization of Split Fluorescent Protein Variants and Quantitative Analyses of Their Self-Assembly Process. *Scientific Reports* **8**, 5344, doi:10.1038/s41598-018-23625-7 (2018).
47 Reynolds, A. *et al.* RNF212 is a dosage-sensitive regulator of crossing-over during mammalian meiosis. *Nat Genet* **45**, 269-278, doi:10.1038/ng.2541 (2013).
48 Serra, H. *et al.* Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. *Proceedings of the National Academy of Sciences* **115**, 2437, doi:10.1073/pnas.1713071115 (2018).
49 Ziolkowski, P. A. *et al.* Natural variation and dosage of the HEI10 meiotic E3 ligase control Arabidopsis crossover recombination. *Genes & development* **31**, 306-317, doi:10.1101/gad.295501.116 (2017).
50 Zhang, L., Ward, J. D., Cheng, Z. & Dernburg, A. F. The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in C. elegans. *Development* **142**, 4374-4384, doi:10.1242/dev.129635 (2015).
51 Jantsch, V. *et al.* Caenorhabditis elegans prom-1 is required for meiotic prophase progression and homologous chromosome pairing. *Mol Biol Cell* **18**, 4911-4920, doi:10.1091/mbc.e07-03-0243 (2007).
52 Meneely, P. M., McGovern, O. L., Heinis, F. I. & Yanowitz, J. L. Crossover distribution and frequency are regulated by him-5 in Caenorhabditis elegans. *Genetics* **190**, 1251-1266, doi:10.1534/genetics.111.137463 (2012).
53 Hillers, K. J. & Villeneuve, A. M. Chromosome-wide control of meiotic crossing over in C. elegans. *Curr Biol* **13**, 1641-1647, doi:10.1016/j.cub.2003.08.026 (2003).
54 Kondo, S. & Miura, T. Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation. *Science* **329**, 1616, doi:10.1126/science.1179047 (2010).
55 Harmon, T. S., Holehouse, A. S., Rosen, M. K. & Pappu, R. V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife* **6**, doi:10.7554/eLife.30294 (2017).
56 Holliday, R., Lewis, K. R. & Hultén, M. Recombination and Meiosis [and Discussion]. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **277**, 359-370 (1977).
57 White, M. J. D. *Animal cytology and evolution*. (University Press, 1973).
58 Pigozzi, M. I. & Solari, A. J. Recombination nodule mapping and chiasma distribution in spermatocytes of the pigeon, Columba livia. *Genome* **42**, 308-314, doi:10.1139/g98-138 (1999).
59 Holloway, J. K., Sun, X., Yokoo, R., Villeneuve, A. M. & Cohen, P. E. Mammalian CNTD1 is critical for meiotic crossover maturation and deselection of excess precrossover sites. *The Journal of Cell Biology* **205**, 633-641, doi:10.1083/jcb.201401122 (2014).

60 Woglar, A. & Villeneuve, A. M. Dynamic Architecture of DNA Repair Complexes and the Synaptonemal Complex at Sites of Meiotic Recombination. *Cell* **173**, 1678-1691 e1616, doi:10.1016/j.cell.2018.03.066 (2018).

ACKNOWLEDGEMENTS

It should be noted that some of the FCS and FRAP data presented in this work have previously been included in a conference proceedings[46], although with limited strains and replicates. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Additionally, we thank Nicola Silva, Sarit Smolikove, Anthony Hyman, Liangyu Zhang, and Regina Bohn for constructing and sharing strains. Finally, we thank Dernburg lab members for ongoing helpful discussions during this project.

SUPPLEMENTARY INFORMATION

crRNAs/Repair templates

Target sequence	Sequence	Source
ZHP-1 C-terminus	5'tcgtctcaatcgaatcgtgg3'	[1]
ZHP-1 C-terminus-linker-GFP11	5'ttcatcatcatccaactgctcgtctcaatcTaaCAgAggaggatctctgttGAGCTTGGATCCGGATCCGGATCCGGATCCGGATCCTCCTCCGCCGGACGCGATCACATGGTCCTGCATGAGTATGTGAACGCCGCCGGGATCACTTAAcataatctgtaatatagatgaccatag3'	This study
ZHP-2 C-terminus	5'tggatccaatgcaatcattg3'	[1]
ZHP-2 C-terminus-linker-GFP11	5'ttgtttggcgcaggacttgagcatccgtcgccaatCctAcgAcaGGAGCTTGGATCCGGATCCGGATCCGGATCCGGATCCTCCGCCGGACGCGATCACATGGTCCTGCATGAGTATGTGAACGCCGCCGGGATCACTTAAaactccctgtatgtcacctccgtattttta3'	This study
ZHP-3 C-terminus	5'gagattaaaacattaatcgg3'	[1]
ZHP-3 C-terminus-	5'AATGGTCGGAGCTTCATTGGACCCGCCGATGAGCTTGGATCCGGATCCGGATCCGGATCCGGATCCTCCGCCGGACGCGATCACATGGTCCTGCATGAGTA	This study

linker-GFP11	TGTGAACGCCGCCGGGATCACTtaatgttttaatctcgttttttt ctgaatt3’	
ZHP-4 C-terminus ZHP-4 C-terminus-linker-GFP11	5’ gagaaaagcacaaggagcat 3’ 5’atgaagctcaaagagaaaagcacaaggagcatcgTaatagtcaaG AGCTTGGATCCGGATCCGGATCCGGATCCGGATC CTCCGCCGGACGCGATCACATGGTCCTGCATGAG TATGTGAACGCCGCCGGGATCACTTAAatcgatttttttta atatatgtgtcgc3’	[1] This study

Strains used

Strains	**Strain/allele sources**	**Identifier**
zuIs178 [his-72(1kb 5' UTR)::his-72::GFP::his-72 3'UTR + unc-119(+)]; unc-119(ed3) III	[2]	JJ1850
htp-3(tm3655) I; *ieSi6[htp-3p::htp-3::GFP::htp-3 3'UTR + unc-119(+)], unc-119 (ed3) III*	[3]	CA1230
gfp::*msh-5 IV*	[4]	-
meIs8 [pie-1p::GFP::cosa-1 + unc-119(+)] II	[5]	AV630
ddIs30 [pie-1p::YFP::cdk-2 (K03e5.3a)+unc-119(+)] II; unc-119(ed3) III	[6]	TH98
gtIs2368 [pie-1p::GFP::rpa-1 + unc-119(+)]	[7]	-
zhp-1::gfp11 I; *iowSi8*[*pie-1p*::*gfp1-10*::*him-3 3’UTR + unc119(+)] II*; *unc-119(ed3) III*	This study, [8]	-
zhp-2::gfp11 I; *iowSi8*[*pie-1p*::*gfp1-10*::*him-3 3’UTR + unc119(+)] II*; *unc-119(ed3) III*	This study, [8]	-
zhp-3::gfp11 I; *iowSi8*[*pie-1p*::*gfp1-10*::*him-3 3’UTR + unc119(+)] II*; *unc-119(ed3) III*	This study, [8]	-

iowSi8[*pie-1p*::*gfp1-10*::*him-3 3'UTR + unc119(+)] II*; *unc-119(ed3) III; zhp-4::gfp11 V*	[8], this study	-
him-3(ie114[gfp::him-3]) IV	[9]	CA1282
syp-3(ok758) I; ieSi11 [syp-3p::EmeraldGFP::syp-3::syp-3 3'UTR + unc-119(+)] II; unc-119(ed3) III	[10]	CA1218
zhp-3(ie80[zhp-3::AID::GFP]) I; ieSi19 [syp-3p::mRuby::syp-3::syp-3 3'UTR + unc119(+)] II	[1],[11], this study	CA1443
zhp-3(ie80[zhp-3::AID::GFP]) I; ie20[Δzim-3, Δhim-8, unc-119(+)] IV	This study	-
zhp-3(ie80[zhp-3::AID::GFP]) I; dpy-18(e364), unc-3(e151) meT7 (III;X;IV)	This study, [12]	-
zhp-3(ie80[zhp-3::AID::GFP]) I; him-5 (ok1896) V	This study	-
syp-1::gfp	[13]	CA1335
syp-2::gfp	[13]	CA1336
zhp-3(ie80[zhp-3::AID::GFP]) I; ieSi64[gld-1p::TIR1::mRuby::gld-1 3'UTR, unc-119(+)] II; ie20[Δzim-3, Δhim-8, cbunc-119(+)] IV	This study, [14]	-

Supplementary References

1 Zhang, L., Kohler, S., Rillo-Bohn, R. & Dernburg, A. F. A compartmentalized signaling network mediates crossover control in meiosis. *Elife* **7**, doi:10.7554/eLife.30789 (2018).

2 Ooi, S. L., Priess, J. R. & Henikoff, S. Histone H3.3 variant dynamics in the germline of Caenorhabditis elegans. *PLoS genetics* **2**, e97-e97, doi:10.1371/journal.pgen.0020097 (2006).

3 Kim, Y. *et al.* The chromosome axis controls meiotic events through a hierarchical assembly of HORMA domain proteins. *Developmental cell* **31**, 487-502, doi:10.1016/j.devcel.2014.09.013 (2014).

4 Janisiw, E., Dello Stritto, M. R., Jantsch, V. & Silva, N. BRCA1-BARD1 associate with the synaptonemal complex and pro-crossover factors and influence RAD-51 dynamics during Caenorhabditis elegans meiosis. *PLoS Genet* **14**, e1007653, doi:10.1371/journal.pgen.1007653 (2018).

5 Yokoo, R. *et al.* COSA-1 reveals robust homeostasis and separable licensing and reinforcement steps governing meiotic crossovers. *Cell* **149**, 75-87, doi:10.1016/j.cell.2012.01.052 (2012).

6 Cowan, C. R. & Hyman, A. A. Cyclin E–Cdk2 temporally regulates centrosome assembly and establishment of polarity in Caenorhabditis elegans embryos. *Nature Cell Biology* **8**, 1441-1447, doi:10.1038/ncb1511 (2006).
7 Koury, E., Harrell, K. & Smolikove, S. Differential RPA-1 and RAD-51 recruitment in vivo throughout the C. elegans germline, as revealed by laser microirradiation. *Nucleic acids research* **46**, 748-764, doi:10.1093/nar/gkx1243 (2018).
8 Hefel, A. & Smolikove, S. Tissue-Specific Split sfGFP System for Streamlined Expression of GFP Tagged Proteins in the Caenorhabditis elegans Germline. *G3 (Bethesda)* **9**, 1933-1943, doi:10.1534/g3.119.400162 (2019).
9 Woglar, A. *et al.* Quantitative Cytogenetics Reveals the Longitudinal Architecture of Meiotic Chromosome Axes. *bioRxiv*, 724997, doi:10.1101/724997 (2019).
10 Rog, O. & Dernburg, A. F. Direct Visualization Reveals Kinetics of Meiotic Chromosome Synapsis. *Cell reports* **10**, 1639-1645, doi:10.1016/j.celrep.2015.02.032 (2015).
11 Rog, O., Kohler, S. & Dernburg, A. F. The synaptonemal complex has liquid crystalline properties and spatially regulates meiotic recombination factors. *Elife* **6**, doi:10.7554/eLife.21455 (2017).
12 Hillers, K. J. & Villeneuve, A. M. Chromosome-wide control of meiotic crossing over in C. elegans. *Curr Biol* **13**, 1641-1647, doi:10.1016/j.cub.2003.08.026 (2003).
13 Sarov, M. *et al.* A genome-scale resource for in vivo tag-based protein function exploration in C. elegans. *Cell* **150**, 855-866, doi:10.1016/j.cell.2012.08.001 (2012).
14 Zhang, L., Ward, J. D., Cheng, Z. & Dernburg, A. F. The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in C. elegans. *Development* **142**, 4374-4384, doi:10.1242/dev.129635 (2015).

Supplementary Figures and Tables

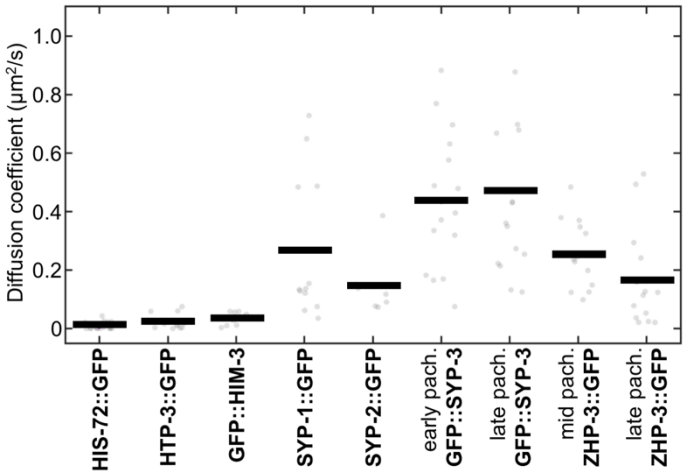

Figure 2S: FCS line-scans are used to estimate diffusion coefficients of proteins with the *average* ACF. Line-scan data analyzed is the same as in **Fig. 2**, however the ACF value for all points was calculated and then averaged before a fit was performed to estimate diffusion coefficients. Proteins are listed with fluorescent tags in their column. Each grey

point is an obtained value from a line-scan, and the black bar is the average value for each protein. N numbers and statistical comparisons are available within **Table S3**.

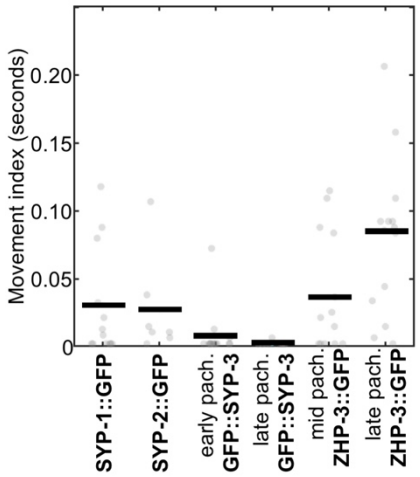

Figure 3S: FCS line-scans estimate the movement of proteins within the SC using *average* pCF. Line-scan data analyzed is the same as in **Fig. 2**. Time for proteins to move 0.135 µm is estimated as the "movement index" by calculating the pCF for each pair of voxels, averaging these pCFs, and finding the correlation time at the maximum value of this average pCF. Proteins are listed with fluorescent tags in their column. Each grey point is a movement index obtained from a line-scan, and the black bar is the mean value. N numbers and statistical comparisons are available within **Table S4**.

			Protein measured							
	Mann Whitney U P-values	*n* =	HTP-3::GFP	GFP::HIM-3	SYP-1::GFP	SYP-2::GFP	early pachytene GFP::SYP-3	late pachytene GFP::SYP-3	middle pachytene ZHP-3::GFP	late pachytene ZHP-3::GFP
Protein measured	HIS-72::GFP	152	9.0160e-04	7.9001e-13	7.3321e-27	1.2204e-16	5.3768e-37	4.7271e-44	3.0514e-34	1.1130e-23
	HTP-3::GFP	104		1.8383e-06	2.1435e-18	1.9542e-11	1.7148e-28	4.8254e-35	7.0655e-26	7.6536e-15
	GFP::HIM-3	80			4.0969e-10	6.2315e-05	9.1870e-21	1.8833e-28	3.8618e-17	9.5612e-06
	SYP-1::GFP	104				0.0089	0.0337	2.2880e-05	0.6474	0.0033
	SYP-2::GFP	56					1.1945e-07	7.0811e-12	8.2855e-04	0.8519
	early pachytene GFP::SYP-3	120						0.0518	6.3686e-04	5.1399e-09
	late pachytene GFP::SYP-3	120							1.2864e-07	1.3604e-14
	middle pachytene ZHP-3::GFP	104								7.1770e-04
	late pachytene ZHP-3::GFP	112								

Table S1: Statistical comparisons and n numbers of data from Fig. 2. Values shown are the P-values of the Mann-Whitney U test.

Protein measured							
Protein measured	**Mann Whitney U P-values**	*n* =	SYP-2::GFP	early pachytene GFP::SYP-3	late pachytene GFP::SYP-3	middle pachytene ZHP-3::GFP	late pachytene ZHP-3::GFP
	SYP-1::GFP	91	0.9059	1.9120e-11	2.3273e-13	0.5736	2.1056e-04
	SYP-2::GFP	49		3.4699e-10	2.9592e-12	0.6110	3.1591e-04
	early pachytene GFP::SYP-3	105			0.6471	7.0625e-09	2.9914e-22
	late pachytene GFP::SYP-3	105				1.9356e-10	2.0321e-25
	middle pachytene ZHP-3::GFP	91					4.6899e-05
	late pachytene ZHP-3::GFP	98					

Table S2: Statistical comparisons and n numbers of data from Fig. 3. Values shown are the P-values of the Mann-Whitney U test.

Protein measured										
Protein measured	**Mann Whitney U P-values**	*n* =	HTP-3::GFP	GFP::HIM-3	SYP-1::GFP	SYP-2::GFP	early pachytene GFP::SYP-3	late pachytene GFP::SYP-3	middle pachytene ZHP-3::GFP	late pachytene ZHP-3::GFP
	HIS-72::GFP	19	0.0460	0.0014	5.0554e-06	1.3603e-04	8.4281e-07	8.4281e-07	2.3650e-06	1.1370e-05
	HTP-3::GFP	13		0.3364	6.3914e-05	4.8904e-04	9.7668e-06	7.8830e-06	1.6497e-05	6.2370e-04
	GFP::HIM-3	10			4.1921e-04	1.0284e-04	3.5881e-05	3.5881e-05	6.3310e-05	0.0207
	SYP-1::GFP	13				0.4824	0.0481	0.0539	0.5316	0.2268
	SYP-2::GFP	7					0.0048	0.0048	0.0476	0.7940
	early pachytene GFP::SYP-3	15						1	0.0427	0.0018
	late pachytene GFP::SYP-3	15							0.0476	0.0018
	middle pachytene ZHP-3::GFP	13								0.0494
	late pachytene ZHP-3::GFP	14								

Table S3: Statistical comparisons and n numbers of data from Fig. 2S. Values shown are the P-values of the Mann-Whitney U test.

Protein measured							
Protein measured	**Mann Whitney U P-values**	*n* =	SYP-2::GFP	early pachytene GFP::SYP-3	late pachytene GFP::SYP-3	middle pachytene ZHP-3::GFP	late pachytene ZHP-3::GFP
	SYP-1::GFP	13	0.5707	0.0569	0.0047	0.5633	0.0056
	SYP-2::GFP	7		0.0092	2.0626e-04	1	0.0319
	early pachytene GFP::SYP-3	15			0.2441	0.0135	5.9342e-05
	late pachytene GFP::SYP-3	15				0.0135	5.9342e-05
	middle pachytene ZHP-3::GFP	13					0.0246
	late pachytene ZHP-3::GFP	14					

Table S4: Statistical comparisons and n numbers of data from Fig. 3S. Values shown are the P-values of the Mann-Whitney U test.

Chapter 3: CDK-2 is essential for meiotic recombination in *C. elegans*

ABSTRACT

Meiosis is a specialized cell division which incorporates recombination between homologous chromosomes and a decrease in the ploidy of daughter cells. In meiosis, an excess of programmed double-stranded breaks are formed in DNA. Amongst these sites, most will be repaired as non-crossovers and a subset will be repaired as crossovers, with at least one crossover formed per pair of homologous chromosomes. Prior to formation of an established crossover, numerous crossover intermediates form at potential crossover sites with reduced amounts of crossover proteins. While many of the essential and nonessential crossover proteins have been characterized, the enzymatic activities which catalyze the formation of crossovers and crossover intermediates are unknown. Work in multiple orthologous systems has demonstrated the importance of the *C. elegans* ortholog of CDK-2 in forming crossovers. Additionally, the crossover essential protein COSA-1 in *C. elegans* is a cyclin-like protein. Based on these findings, CDK-2 was investigated as a crossover-dependent protein in *C. elegans*. Here, we show that CDK-2 localizes to crossovers, crossover formation depends on its continued activity, and localization of crossover proteins to both crossover intermediates and mature crossovers depends on CDK-2 presence. Together, our characterization implicates CDK-2 as a likely partner of COSA-1 in designating, reinforcing and forming crossovers in *C. elegans*.

INTRODUCTION

In meiosis, cells undergo two rounds of cell division preceded by a single S-phase. This decreases the ploidy of the daughter cells for the production of gametes and enables sexual reproduction. Central to the process of meiosis is the formation of crossovers (COs) during meiosis I. These contacts between chromosomes form chiasmata and are critical for faithful segregation of chromosomes to avoid aneuploidy[1,2].

COs form in prophase I during a stage known as pachytene. This stage follows the pairing of homologous chromosomes and the formation of an excess of double-stranded breaks (DSBs) in DNA. Pachytene is defined by the presence of a proteinaceous structure between chromosomes known as the synaptonemal complex (SC), and a number of highly coordinated events must occur towards forming COs. DSBs are resected by the cooperation of the RTR complex[3] in conjunction with the MRN/X complex[4]. RPA binds the ssDNA, inhibiting it from reannealing[4], and RAD51 protein binds to enable the homology search necessary for DSB repair[4]. DSBs greatly outnumber COs, and so some DSBs will be repaired as COs while the majority will be non-COs[5]. However, during this stage, pro-crossover factors, including the MutSγ complex, will localize to sites that outnumber the final amount of CO sites. Therefore, these early numerous sites are known as CO intermediates.

For the study of these events in CO formation, *C. elegans* has emerged as a powerful model system. Different stages of meiosis can be observed in a single gonad, due to the a temporally organized gonad in hermaphrodites. Additionally, cells in the stage of CO formation, pachytene, are readily visible for study. This has led to a significant degree of characterization for orthologs within *C. elegans* responsible for the previously described steps. However, even with the discovery of a multitude of CO-promoting and CO-marking proteins, the protein activities which designate COs within *C. elegans* is still unclear. In other words, it is unknown what protein activities allow CO intermediates to form into COs.

Numerous discoveries suggest that one such essential activity is produced by a cyclin-dependent kinase (CDK) in *C. elegans*. First, there is the existence of COSA-1, an essential CO factor[6]. This protein is commonly used in the *C. elegans* meiosis to mark the sites of COs, and is known to have a cyclin-like structure, and it is not predicted to have enzymatic activity[6]. CDK proteins rely on a cyclin partner to establish functionality and specificity for a target substrate to phosphorylate. The cyclin-like structure of COSA-1 suggests that a CDK partner may act along with COSA-1 at COs to designate them. COSA-1 is noted to also have essential orthologues throughout metazoa including CNTD1 in mammalian systems[6,7]. Second, in orthologous systems, CDKs are known to have essential roles in CO designation at CO sites[8-10], and other roles at CO intermediates[9]. Specifically, CDK2 has been identified to localize to CO sites in mammals[9]. CDKA, an ortholog of both CDK1 and CDK2, has a function in forming COs in *Arabidopsis*[8]. In this study, we characterize the role of CDK-2 in *C. elegans* CO designation. Using a variety of techniques, we verify that CDK-2 has a role in CO designation, that its activity is necessary in establishing COs, and its relative timing in the process of meiosis.

METHODS AND MATERIALS

Chromosome spreads

To prepare for these experiments, approximately 60 L4 *C. elegans* hermaphrodites were transferred to a fresh NGM plate and allowed to develop for 24 hours at 20°C prior to performing spreads. Chromosome spreads were performed by following previous protocols[11] with a few exceptions. In substitution of BSA, 1× Roche blocking solution in PBST was used. Next, the second PBST wash after secondary antibody incubation contained (1:10000 from 5mg/ml stock 4′,6-diamidino-2-phenylindole DAPI) and was incubated for ten minutes. Finally, ProLong Diamond (Invitrogen) was used to mount samples instead of Vectashield.

Immunofluorescence

Following treatment or control treatment with auxin, auxin/RNAi, or 3-IB-PP1 for 24 to 48 hours from L4, *C. elegans* hermaphrodites were picked to a coverslip with 1× egg buffer (25 mM HEPES pH 7.4, 118 mM NaCl, 48 mM KCl, 2 mM EDTA, 0.5 mM EGTA) containing 0.05% tetramisole and 0.1% Tween-20. These hermaphrodites were cut with a scalpel and fixed in 1% formaldehyde for one minute and transferred to a 1.5 ml

microcentrifuge tube with 1 ml of PBST. Next, the PBST was removed and -20°C methanol was added for five minutes. This methanol wash was followed by three five-minute PBST washes, and then blocking with 1× Roche blocking solution in PBST for 30 minutes. The blocking solution was replaced with fresh blocking solution containing primary antibodies and samples were stained overnight at 4°C. Following this staining, three PBST washes were performed and the tubes were stained with secondary antibodies in PBST for 1.5 to 2 hours. Three more PBST washes were performed with the second containing DAPI (1:10000 from 5mg/ml stock). Finally, the dissected hermaphrodites and gonads were mounted in ProLong Diamond (Invitrogen).

Analog-sensitive CDK-2 allele design and implementation

To develop an analog-sensitive CDK-2, the primary protein structure of CDK-2 (obtained from WormBase[12]) was aligned to the primary sequence of c-SRC from Uniprot[13] and the gatekeeper residue of CDK-2 was identified as phenylalanine-91. This amino acid was mutated to glycine to generate an AS1 allele[14]. An AS2 allele was also generated but found to be insensitive to analog treatments. Sensitivity to five different ATP analogs was tested by first dissolving each analog into ethanol at 2 mg/ml. Next, NGM plates containing one of the five analogs in ethanol were made in 30 mm×15 mm petri dishes at varying analog concentrations, with the analog added while the plates were still liquid (i.e. not hardened from cooling). Overnight LB culture of E. coli (OP50) was centrifuged, had excess LB removed, and was resuspended in minimal LB. Forty microliters of this OP50 resuspension was pipetted onto the analog plates and allowed to dry overnight at 4°C. N2 strain worms were tested for sensitivity at these varied concentrations of analog by picking four L4s onto plates, performing egg counts, and observing whether the eggs produced viable progeny. Once a maximum concentration for each analog was found, which had no impact on N2 viability and reproduction (**Fig. 2S**), *cdk2*^{as1} was tested for sensitivity to the five analogs. The CDK-2 analog-sensitive strain was found to produce no eggs in the presence of 3-IB-PP1, even at significantly lower than maximum concentration, and was established as sensitive to 3-IB-PP1. The five analogs tested and final concentrations in NGM were: 3-IB-PP1 (8 µg/ml), 3-MB-PP1(40 µg/ml), 1-NA-PP1(40 µg/ml), 3-BrB-PP1(40 µg/ml), 1-NM-PP1(40 µg/ml).

Auxin/RNAi treatment

For auxin treatment, L4 hermaphrodite worms were picked onto plates made in the same manner as the analog plates, with the deviation that auxin in ethanol was added to a final concentration of 1 mM. For further depletion of CDK-2, an auxin and RNAi combination was implemented, in which plates contained 0.2 mM auxin, 1 mM IPTG, and 100 mg/ml carbenicillin. These plates were seeded with RNAi bacteria (CDK-2 exon 3; see Supplementary Information) which were grown overnight on the plate at 37°C, and worms were washed in M9 prior to picking onto these plates.

Live worm immobilization

For live imaging, adult hermaphrodites (24-hour post L4, with treatment (3-IB-PP1) time added) were immobilized in PBS with 20% glycerol and 20% PEG2000 on a microscopy slide with a coverslip.

Imaging

Images of fixed samples were collected as z-stacks at intervals of 0.2 µm on a Deltavision Elite microscope (GE), using a 100×, 1.4 N.A. oil-immersion objective. Images were deconvolved within the SoftWoRx package, colorized and projected in ImageJ. For whole gonad images, multiple overlapping images were taken, and later combined using Adobe Photoshop CC 2018.

Live imaging was collected as z-stacks on a Marianas spinning-disc confocal microscope (Intelligent Imaging Innovations, Inc. (3i)) at intervals of 0.25 µm, using a 100×, 1.46 N.A. oil-immersion objective, and with a total z-stack depth of 8 µm. Images shown were deconvolved using Slidebook 6 software, and projected in ImageJ.

Transgenic/mutant worm generation

Previously established protocols for use of the CRISPR/Cas9 system were used to generate mutant alleles and tag genes[15]. Repair template used was either ssDNA provided by Integrated DNA Technologies (IDT), or PCR product for longer inserts. Roller progeny were chosen from micro-injections as described and PCR checked, followed by sequencing to confirm correct sequence for epitope tagging or mutations. PCR primers for size checks and sequencing available upon request. WT worms were then selected from roller progeny.

RESULTS

CDK-2 and COSA-1 colocalize in late pachytene

Enrichment of protein at CO sites substantially implicates a protein in CO designation/formation. Prior to this work, the CDK-2 ortholog of mammalians has been observed to localize to CO sites in prophase I[9], but the localization of CDK-2 in *C. elegans* prophase I has not been observed. To detect the localization of CDK-2 in pachytene and diplotene, when COs are designated and formed, we visualized it along with known CO factors COSA-1 and ZHP-3 using a previously developed “chromosome spread” technique[11]. Results of this (**Fig. 1**) show that CDK-2 colocalizes with COSA-1 from mid/late pachytene, demarcated by the presence of ZHP-3 along the SC, all the way through diplotene, forming bright foci at CO sites, as marked by COSA-1. Of note, CDK-2 localizes to CO sites prior to the enrichment of ZHP-3, which is still heterogeneously present throughout the SC. This localization of CDK-2 indicates it may function in both CO designation and formation throughout pachytene. Additionally, this colocalization of CDK-2 and COSA-1, a cyclin-like protein, implicate them as possible functional partners.

COSA-1 localization depends on the presence of CDK-2

If CDK-2 is functioning in establishing COs as suggested by its localization, then its absence should suspend COs. Additionally, CDK-2’s possible functional partner, COSA-1, should be affected by its absence. To see if this was the case, we used the auxin-inducible degradation (AID) system to selectively degrade epitope-tagged CDK-2 within

the gonad of hermaphrodite *C. elegans*. CDK-2 is an essential cell cycle regulator[16], and so we elected to use this approach to only degrade CDK-2 in the gonad, while maintaining viable animals. Gonads of treated animals had significant degradation of CDK-2, demonstrated by a lack of increased CDK-2 expression in diplotene and diakinesis nuclei (**Fig. 1S**). Upon degradation of CDK-2 (**Fig. 2**), there were two main observable effects. First, there are fewer cells within the treated gonad. This likely results from a decrease in mitosis efficiency. Germline cells mitotically replicate within the *C. elegans* hermaphrodite gonad prior to entrance into meiosis. CDK-2 is known to drive the progression of this mitosis, and so its degradation here leads to less efficient mitosis and fewer cells. Despite this disruption of mitosis, the onset and presence of meiosis is largely unaffected. This is observable in normal timing of axis assembly as an indicator of entrance into meiosis, marked by HTP-3[17]. The second, and more substantial effect upon CDK-2 depletion in prophase I is in CO designation. In the untreated gonad, cells still form the normal number of COs, marked by COSA-1 foci, with six per cell. In contrast, the CDK-2 depleted cells form no COSA-1 sites, and therefore no crossovers. **Fig. 2** has been scaled to show any potential faint COSA-1 signal, but it is apparent that there is no appreciable signal from COSA-1. This means that COSA-1 concentration at CO sites, a condition of CO formation, requires the presence of CDK-2.

COSA-1 localization depends on CDK-2 activity

Kinases are known to sometimes serve scaffolding functions[18]. Because of this potential role for CDK-2, we could not determine from our CDK-2 depletion whether its presence or kinase activity was required in forming COs. To experimentally make this distinction, we pursued inhibition of CDK-2 kinase activity through an analog-sensitive allele of CDK-2, similar to the analog-sensitive allele of PLK-1 previously used in *C. elegans* embryos[19]. The analog-sensitive allele functions as depicted in **Fig. 3A**. Normally, CDK-2 has an ATP pocket with dimensions to only accommodate ATP, and is able to exclude other larger ATP analogs. However, when phenylalanine-91, known as the “gatekeeper” is mutated to the less bulky amino acid glycine, larger ATP analogs which are non-hydrolyzable can bind within the ATP pocket. In this way, non-hydrolyzable ATP analogs can act as a competitive inhibitor to ATP, to specifically inhibit a single kinase. It is important to note that this competitive inhibitor competes with native ATP, so it’s possible for the kinase to maintain some activity.

We made the aforementioned mutation and found that our generated *cdk-2*^{as1} allele was sensitive to the treatment of 3-IB-PP1, an ATP analog, administered to adult hermaphrodites through NGM plates, as treatment eliminated the generation of progeny. Control worms treated with the same concentration of 3-IB-PP1 still had progeny and have six crossovers per cell (**Fig. 2S**). We tested the necessity of CDK-2 activity in establishing crossovers by using COSA-1 as a marker with and without treatment on the *cdk-2*^{as1} allele (**Fig. 3B**). When untreated, *cdk-2*^{as1} still forms six COs as shown by the six COSA-1 sites per nucleus. In contrast, when treated with 3-IB-PP1, many cells had no COSA-1 sites, while others still had one or two. This establishes the requirement of CDK-2 kinase activity and not solely its presence in establishing CO sites. As noted, 3-IB-PP1 is a competitive inhibitor, so in cases where some COs form

(COSA-1 foci) in spite of treatment, these are likely due to residual low activity of inhibited CDK-2^{as1}. We also note that as in CDK-2 depletion (**Fig. 2**), there are fewer cells in the gonad, likely for the same inhibition of pre-meiotic germline proliferation described above.

Low CDK-2 activity is sufficient to establish infrequent COs

We reasoned that the COSA-1 sites seen during CDK-2^{as1} inhibition were due to residual activity of CDK-2. To examine whether these COSA-1 sites were able to mature into COs, even in the presence of CDK-2 inhibition, we counted DAPI staining bodies in diakinesis under inhibited and non-inhibited CDK-2 conditions. As performed in other studies[15], six DAPI staining bodies reflects six COs formed to support six structures known as bivalents. Contrastingly univalents, that have no CO formed and unassociated homologs, increase the number of DAPI staining bodies. We counted the bivalents/univalents in the most mature oocyte of the gonad arm, referred to as the -1 oocyte, of the *cdk-2^{as1}* allele (**Fig. 4A**) for 48-hour post L4s. Counting of univalents and bivalents in this oocyte is standard due to its maximum separation and space between univalent/bivalents to facilitate their segmentation and count.

Without inhibition, *cdk-2^{as1}* was proficient in forming COs (**Fig. 4A**) with six bivalents. The exceptions of five bivalents are likely due to the close proximity of bivalents making their distinction for counting not possible. With 3-IB-PP1 inhibition of CDK-2^{as1} there were major effects. The majority of nuclei either had non-condensed chromosomes (**Fig. 4A**), or no DNA present in the -1 oocyte, likely due to apoptosis. In both of these cases, the number of DAPI staining bodies was counted as zero. In other oocytes with inhibition treatment, the univalent number were increased, but not always twelve. These oocytes with less than twelve univalent, and thus DAPI staining bodies, indicate the presence of COs. Even with inhibition which yields most cells apoptotic or without condensed DNA, there were still some COs formed. This result emphasizes that the CO step performed by CDK-2 is robust and can function with minimal enzymatic activity.

To again confirm that 3-IB-PP1 treatment has no effect on non-analog-sensitive alleles, we treated N2/Bristol with the analog for 48 hours (**Fig. 4B**). We found that as in the untreated *cdk-2^{as1}*, cells had a normal number of crossovers. A few oocytes appeared to have fewer than six bivalents, but again this was due to close association of two bivalents. Since the 48-hour treated CDK-2^{as1} had effects on DNA beyond CO formation, we repeated the experiment with a decreased treatment, 24-hour treated post L4s (**Fig. 4B**). With this treatment, there was a range of DAPI staining body number from six to twelve. Again, this indicated that with treatment there was still a presence of COs, as not all cells had twelve univalents. This presence of COs suggests that COSA-1 foci seen in **Fig. 3B** were likely real COs. Interpreting this in terms of CDK-2 activity, again even with low activity in the presence of a competitive inhibitor, CDK-2 can establish some COs.

Continued CDK-2 activity is needed to maintain CO site formation

While we have demonstrated the necessity of CDK-2 in establishing COs, it was uncertain whether the activity of CDK-2 was switch-like and dispensable after bright COSA-1 foci formation, or if ongoing kinase activity was necessary to maintain CO sites. To examine if there is an ongoing requirement of CDK-2 activity throughout meiosis, we performed variable time inhibition on *cdk-2*^{as1}, observing possible effects on COSA-1 sites (i.e. CO sites). We took 24-hour post-L4 hermaphrodite *cdk-2*^{as1} and performed 3-IB-PP1 treatment for two, four, and eight hours (**Fig. 5**).

First, we imaged GFP::COSA-1 foci in 24-hour post-L4 worms without treatment and found that COSA-1 progresses from a state of substantial nucleoplasmic signal with numerous dim foci to six bright foci and no nucleoplasmic signal. Notably, these experiments were carried out in live worms, because the observable nucleoplasmic signal of GFP::COSA-1 is eliminated in standard *C. elegans* gonad immunofluorescence (**Fig. 2**). Upon CDK-2 inhibition for two hours, this normal appearance of GFP::COSA-1 becomes disrupted. In earlier pachytene, there are still abundant dim foci and significant nucleoplasmic signal, but the bright foci have drastically changed. In the marked cell (**Fig. 5**), one nucleus has decreased to only have two COSA-1 foci, instead of the standard six. Additional cells display both bright and dim foci within the same nucleus. After four hours some cells display seemingly increased nucleoplasmic signals. This may reflect an inability of COSA-1 to sequester to CO sites while being stably localized in the nucleus by the co-occurrence of CDK-2.. Upon eight-hour treatment, the effects were consistent with the two shorter treatments, although with greater effects. All cells with bright COSA-1 foci have an abnormal number (<six) and appreciable heterogeneity in the brightness foci. Together, these results demonstrate the necessity of ongoing CDK-2 activity to pattern COs. Without CDK-2's activity, there is a rapid decrease in bright COSA-1 site number, heterogeneity in COSA-1 brightness, and many cells with a surplus of dim foci. All of these effects on the CO essential factor COSA-1 reflect an inability to designate, maintain designation, and eventually form COs.

CDK-2 activity is needed to establish CO intermediates

In early and middle pachytene, numerous CO intermediates, which outnumber COs, are formed. These intermediates include many CO factors at lower levels than the later established COs. To examine whether CDK-2 was needed to establish CO intermediates, we depleted CDK-2 in the presence of tagged MSH-5, a component of the MutSγ complex which marks CO intermediates, to see if there was any observable effect (**Fig. 6A**). In addition to using the AID system, we also elected to incorporate RNAi of CDK-2 in an attempt to deplete CDK-2 to the highest degree possible (**Fig. 3S**).

In the presence of CDK-2, there are detectable foci of MSH-5 throughout pachytene. These sites outnumber the six sites of COs and therefore are intermediates, sites thought to have the potential to form COs. Upon depletion of CDK-2, there are major effects on MSH-5 localization. First, bright foci are absent from cells, but a few foci form later in pachytene. For example, the center cell in the inset has a dim focus (**Fig. 6A**). We interpret these remaining foci as a result of residual CDK-2 activity. This is similar to the remaining COSA-1 sites in the presence of analog-sensitive treatment and again

reflects that CDK-2 can sometimes fulfill a functional requirement with limited functional units. Also of interest, the majority of MSH-5 signal has moved into the nucleoplasm, and the image has been scaled to appreciate this significant signal. Together, these results show CDK-2 is essential in forming and stabilizing CO intermediates.

A COSA-1 mutation mimics CDK-2 depletion for CO intermediate formation

Since we believe that COSA-1 and CDK-2 are acting together, we tested whether a COSA-1 mutant would recapitulate the same phenotype with regards to MSH-5 as the CDK-2 depletion. Previous studies have shown COSA-1 to be present at CO intermediates[11], as demonstrated by the dim COSA-1 foci which number greater than six in the zero hour control of **Fig. 5**. It was also concluded that COSA-1 was not needed to form MSH-5 foci, based on antibody staining[11] of MSH-5.

We constructed a COSA-1 null mutant with a mutation at a likely CHK-2 phosphorylation site. CHK-2 is a "master regulator" with essential activity for the commencement of meiosis which decreases through its progression[20]. By mutating this site (threonine 166) to be phosphomimetic (aspartic acid), we hoped to disturb COSA-1's regulation. We also again made use of GFP::MSH-5 instead of MSH-5 antibody. This reporter did not necessitate the use of an antibody and allowed us to exclude any possibility off target signal being misinterpretted as MSH-5 foci. With this mutation and visualization strategy, there is a substantial impact on MSH-5 localization, and it replicates the results of the CDK-2 depletion. MSH-5 foci are largely absent with an occasional focus, indicating a loss of COs (**Fig. 6B**). The majority of MSH-5 signal has again shifted into the nucleoplasm. These effects of mutating COSA-1 are consistent with the depletion of CDK-2, meaning they have the same effect on CO intermediate formation.

CDK-2 is not detectable at CO intermediates

We have shown that CDK-2 has a role in establishing CO intermediates (**Fig. 6A**), indicating it exists at CO intermediates and is detectable at CO sites (**Fig. 1**). However, we had not evaluated if it was detectible at CO intermediates. We again performed chromosome spreads. MSH-5 was used as a marker for CO intermediates, and CDK-2 localization was observed (**Fig. 7**). While MSH-5 sites are visible earlier in meiosis, CDK-2 foci are not visible at the same locations, and remain so until six CO sites form bright foci of CDK-2. Although CDK-2 is essential in establishing CO intermediates, it must be in relatively low abundance at those sites, but is sufficient to establish the sites. This agrees with the appearance of some MSH-5 foci in the CDK-2 depletion and COSA-1 mutant because in both cases remaining protein complex was able to stabilize occasional MSH-5 sites.

DISCUSSION

CO designation depends on the coordination of proteins to select a DSB site, and enable it to become a CO. Prior to this study, many of the enzymatic activities which determine a CO in *C. elegans* were unknown. CDK-2 is a cyclin-dependent kinase with

major roles in *C. elegans* cell cycles[21]. However, here we have demonstrated a number of its roles in CO formation in meiosis, revealing one of the protein activities responsible for designating COs. CDK-2's presence is essential in establishing CO intermediates and its presence and activity are essential in forming mature COs.

As a cyclin-dependent kinase, CDK-2 depends on a cyclin partner to confer the target substrate for phosphorylation. Our results suggest that the cyclin-like protein, COSA-1, is the cyclin partner of CDK-2. COSA-1 and CDK-2 colocalize at CO sites, COSA-1 localization is dependent on CDK-2 activity and presence, and CO intermediates (MSH-5 foci) are similarly dependent on both COSA-1 and CDK-2. Finally, shown elsewhere in Chapter II of this thesis work, we demonstrated that both CDK-2 and COSA-1 are not dynamic within foci.

This similarity in dynamics is consistent with them as kinase and cyclin partners. Interestingly, the stable association of CDK-2/COSA-1 with CO sites is maintained by ongoing CDK-2 activity, as demonstrated by the decreased number of COSA-1 foci with brief CDK-2^{as1} inhibition (2 hours). This ongoing need for CDK-2 activity opens the possibility that its phosphorylation target may actively be dephosphorylated by a phosphatase, or that the phosphorylation is not stable over the course of meiosis. While all these results implicate COSA-1 as the cyclin partner of CDK-2, it does not exclude the possibility that it has other partners. However, no other known CDKs or cyclins are known to have these commonalities, dependencies, or functions in CO formation we have shown here. Further, CDK-2 orthologs are known to have many substrate targets[22]. As such, it may also utilize other cyclins for altering its target or even depend on switching of cyclins to alter its targets as in the progression of mitosis.

In this study, CO intermediates are shown to be dependent on both CDK-2 and COSA-1. This reveals a complicated relationship between CDK-2/COSA-1 and two RING domain-containing proteins known as ZHP-3/4[15]. These proteins are also essential for CO formation. They have an interesting series of localizations in the SC, and in Chapter 2 of this dissertation, as well as previous studies, were postulated to function with two of their paralogs (ZHP-1/2) in a reaction-diffusion system[23]. ZHP-3/4 have been shown to influence the localization of COSA-1 such that when they remain throughout the SC, multiple sites of COSA-1 form. Based on this result, it would seem that ZHP-3/4 determine COSA-1 localization and thus COs. However, COSA-1 mutants have been shown to not localize ZHP-3/4 to CO sites, as ZHP-3/4 remain throughout the SC[15]. In advancing these findings, we have found that COSA-1, and now CDK-2, impact CO intermediate formation, which will determine the ability of ZHP-3/4 to concentrate at CO sites. So, while CDK-2 is needed to establish CO intermediates and COs, concurrently so are ZHP-3/4. It seems that these two systems of proteins reinforce one another to select and then designate a CO, with both of them reinforcing the outcome. Since we show that relatively low functional units of CDK-2 are necessary to establish both COs and CO intermediates, having tandem systems to select and establish COs may provide necessary robustness to the system which may be sensitive without it.

A number of investigative directions would help to inform CDK-2's relation to the ZHPs. Biochemistry could be used to determine the relevant targets of CDK-2 in CO designation, and the complex formation of CDK-2 with COSA-1 would help to inform its role. *C. elegans* may not be the best model system to perform this due to a relatively low abundance of CDK-2 in CO foci versus nucleoplasmic CDK-2 in diakinesis. Another informative result towards how CDK-2 functionally establishes COs from multiple CO intermediates would be whether additional CDK-2/COSA-1 activity through experimental manipulations is sufficient to form additional COs or shift CO designation timing. Overall CDK-2 has proven to be an essential CO factor, and its continued study will assist in unwinding the regulation of COs.

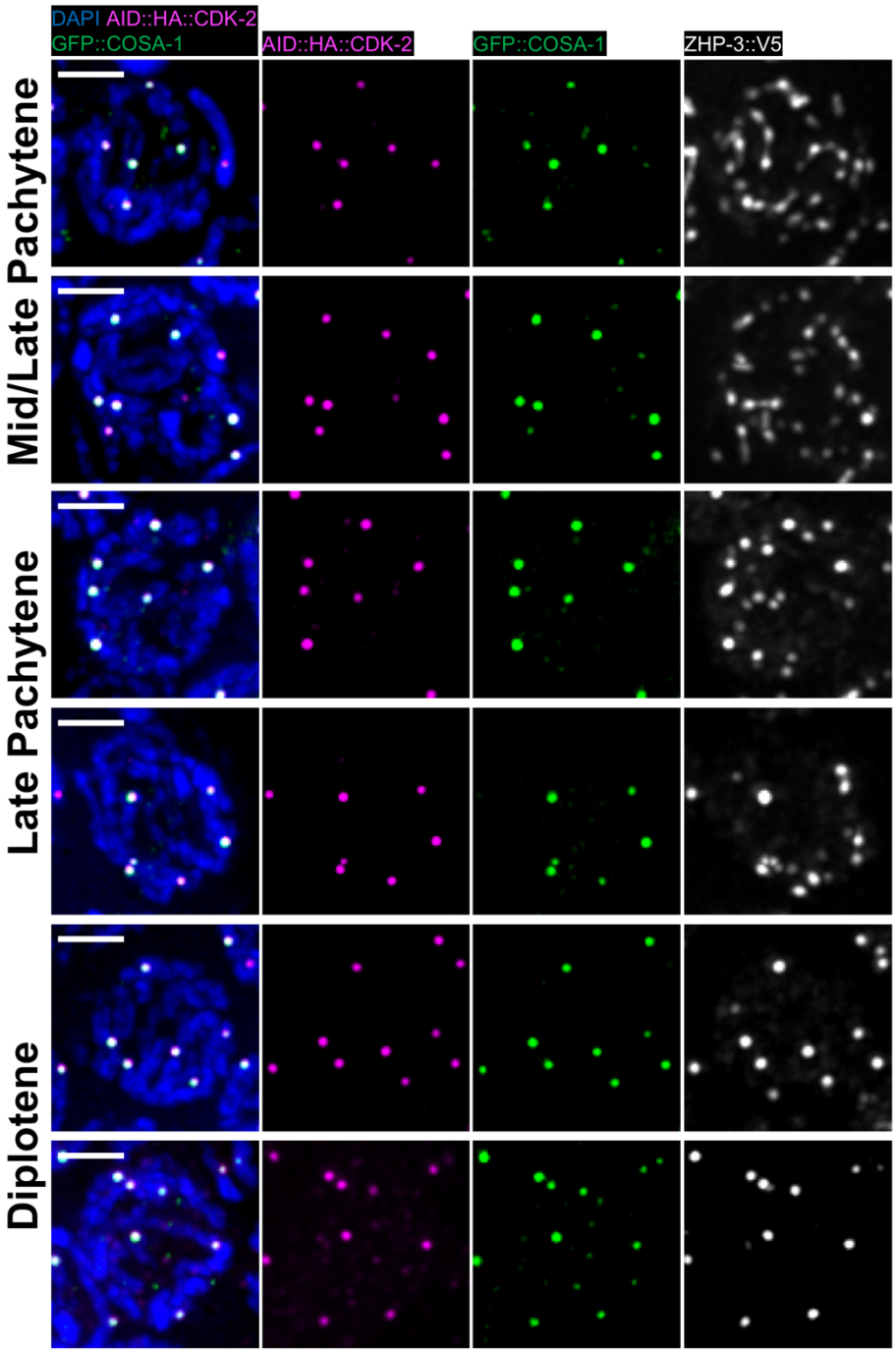

Figure 1: Chromosome spreads show colocalization of CDK-2 and COSA-1 from middle pachytene through diplotene. Maximum projections of nuclei show CDK-2 (magenta) and COSA-1 (green). ZHP-3 (grey) is used as a meiosis progression marker. Scale bars are 2 μm.

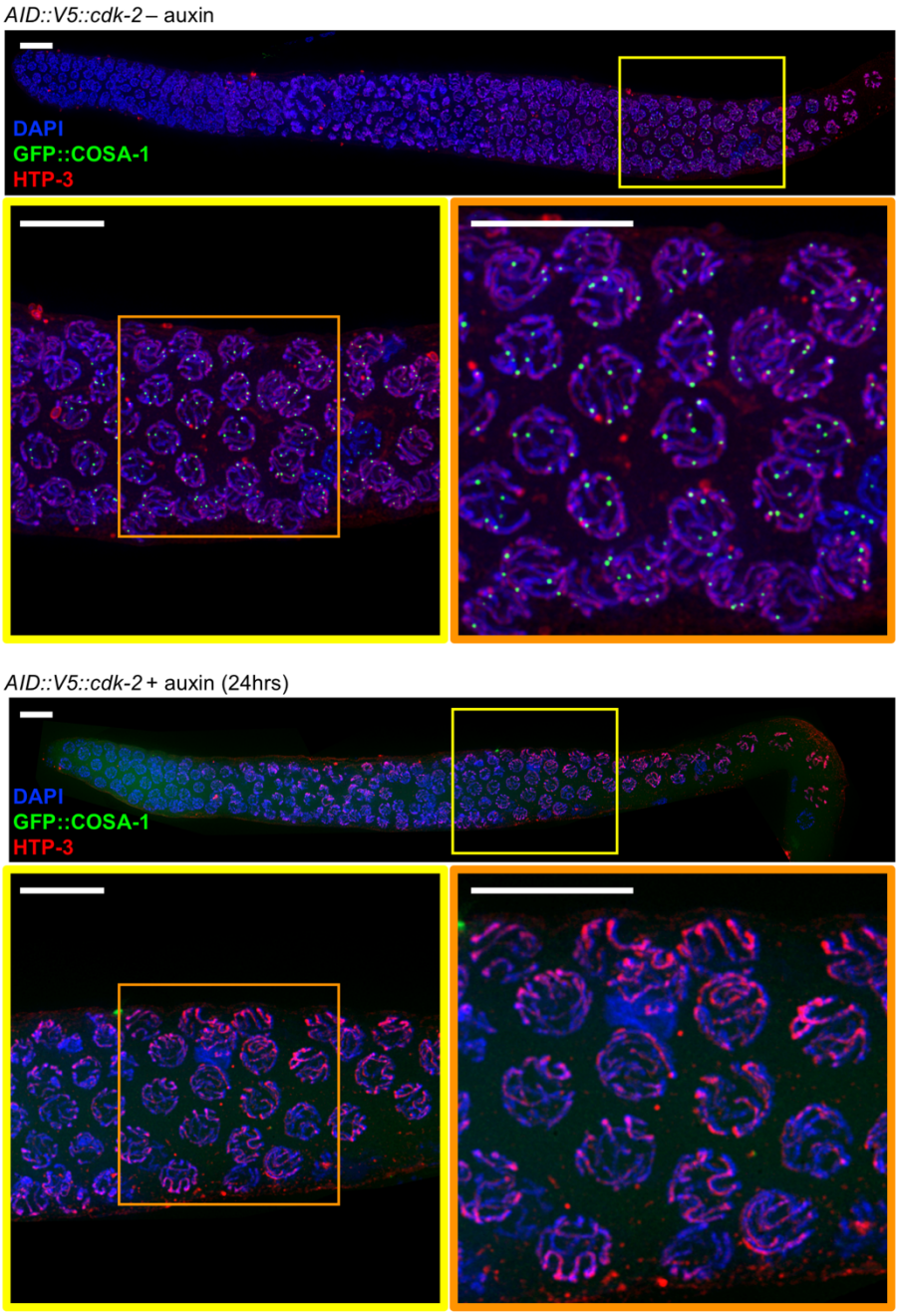

Figure 2: Appearance of late COSA-1 foci depend on CDK-2. Representative maximum-intensity projections of gonads from hermaphrodites at 24 hours post-L4. Depletion of degron-tagged CDK-2 was induced by treatment with auxin for 24 hours, starting from the L4 stage. DAPI (blue) and COSA-1 (green) mark DNA and designated CO sites, respectively. HTP-3 (red) marks the chromosome axes. For images from auxin-treated hermaphrodites, images showing green (COSA-1) fluorescence were scaled much higher to confirm an absence of foci. Insets are framed in colors that indicate the corresponding boxed areas in the lower-magnification images. Scale bars are all 10 μm.

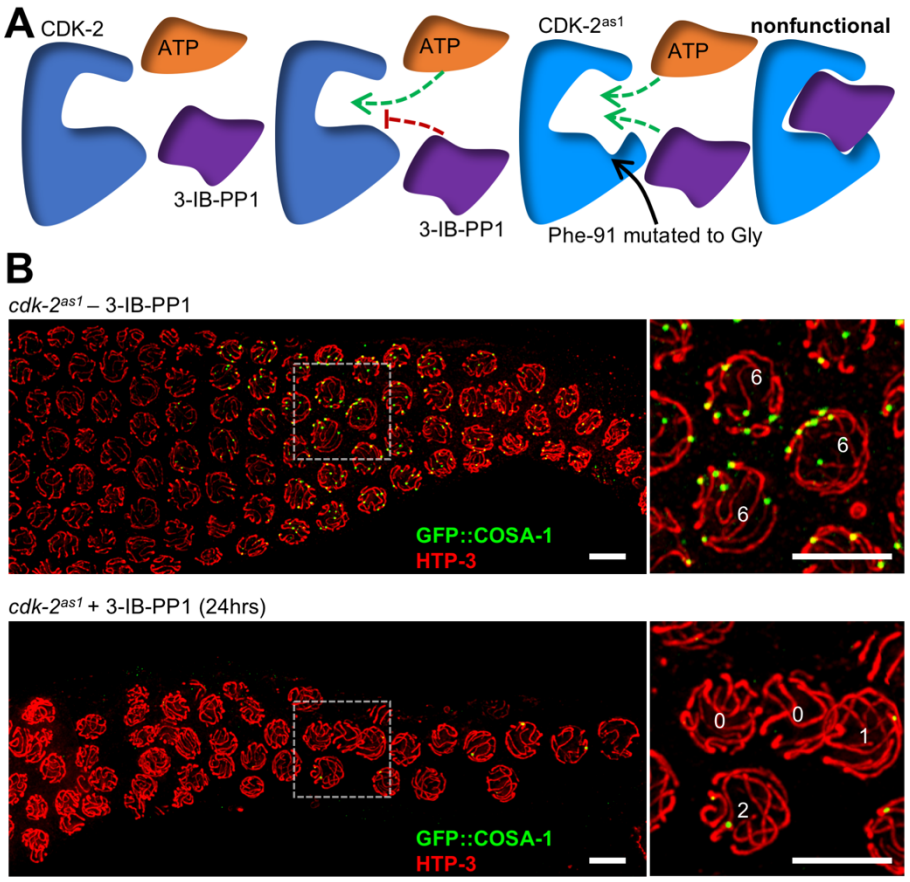

Figure 3: An analog-sensitive CDK-2 allele shows that CDK-2 activity is needed to establish COSA-1 foci. (**A**) A representation of the design and use of an analog-sensitive CDK-2. (**B**) Immunofluorescence maximum projection images of 24-hour post-L4 gonads in CDK-2 analog-sensitive background (*cdk-2*^{as1}). Analog treatment (3-IB-PP1) was for 24hrs. HTP-3 (red) is shown as a nucleus and meiosis marker. COSA-1 (green) is shown as a CO marker. Over insets, the number of COSA-1 sites are counted for the example nuclei. Analog inhibition of CDK-2 results in near-zero COSA-1 sites, with a wildtype six without inhibition. Scale bars (including insets) are 5 μm.

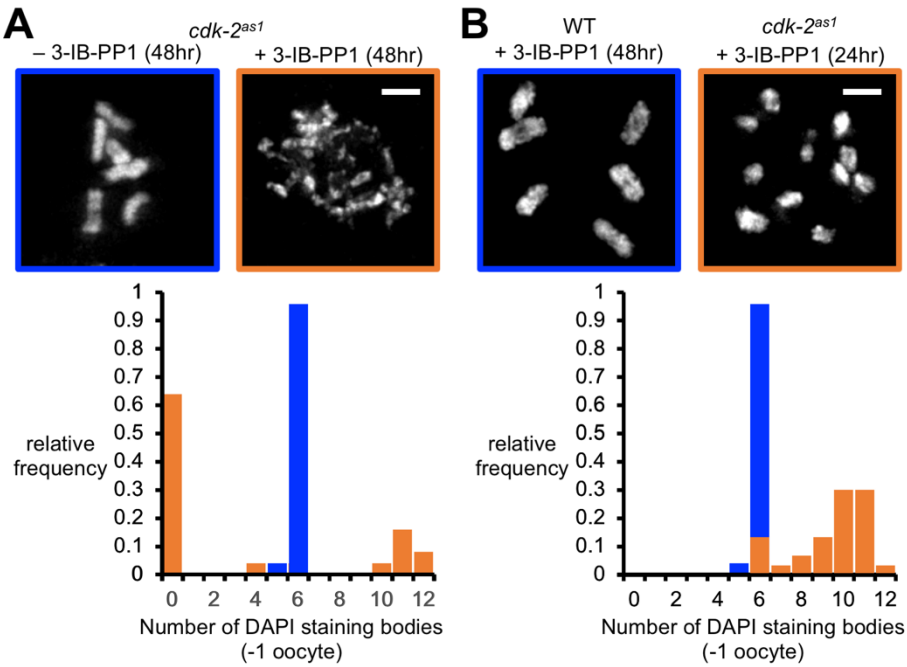

Figure 4: Inhibition of CDK-2 activity impairs crossover formation. (**A**) Examples of " -1" oocytes from *cdk-2*^{as1} hermaphrodites at 48 hours post-L4, in the absence and presence of 48-hour ATP analog (3-IB-PP1) treatment. Gonads were dissected and stained with DAPI, and are shown as maximum-intensity projection images. Quantification of DAPI-staining bodies is shown below the images. In animals treated with the inhibitor, many -1 oocytes contained no DNA or uncondensed DNA, as shown in A. Both cases were scored as zero DAPI staining bodies. n=25 for each condition. Scale bar is 2 μm. (**B**) Examples and plots of wildtype *cdk-2* (48-hour post-L4; 48-hour analog treatment) and *cdk-2*^{as1} (48-hour post-L4; 24-hour analog treatment) shown as in **A**. Scale bar is 2 μm.

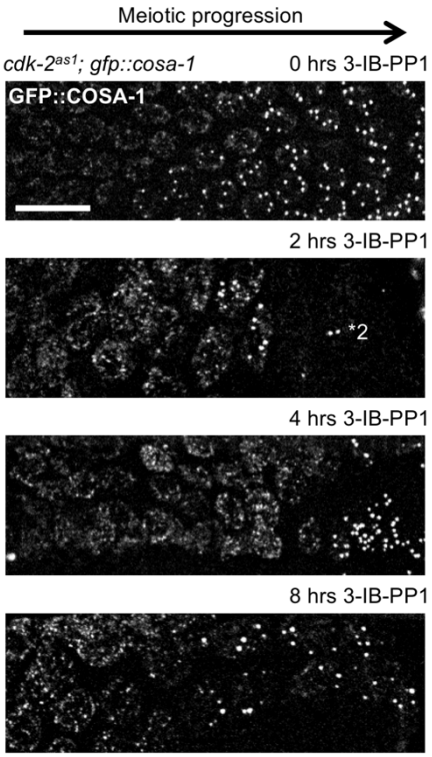

Figure 5: A time series of analog treatment shows that ongoing CDK-2 activity is needed to maintain COSA-1 sites. Representative maximum projections of live gonads (24-hour post-L4 + treatment time). Live gonads are used to show nucleoplasmic and CO associated COSA-1 (grey). An image of a 2-hour treatment, one nucleus with only two COSA-1 foci, is marked. Scale bar is 10 μm.

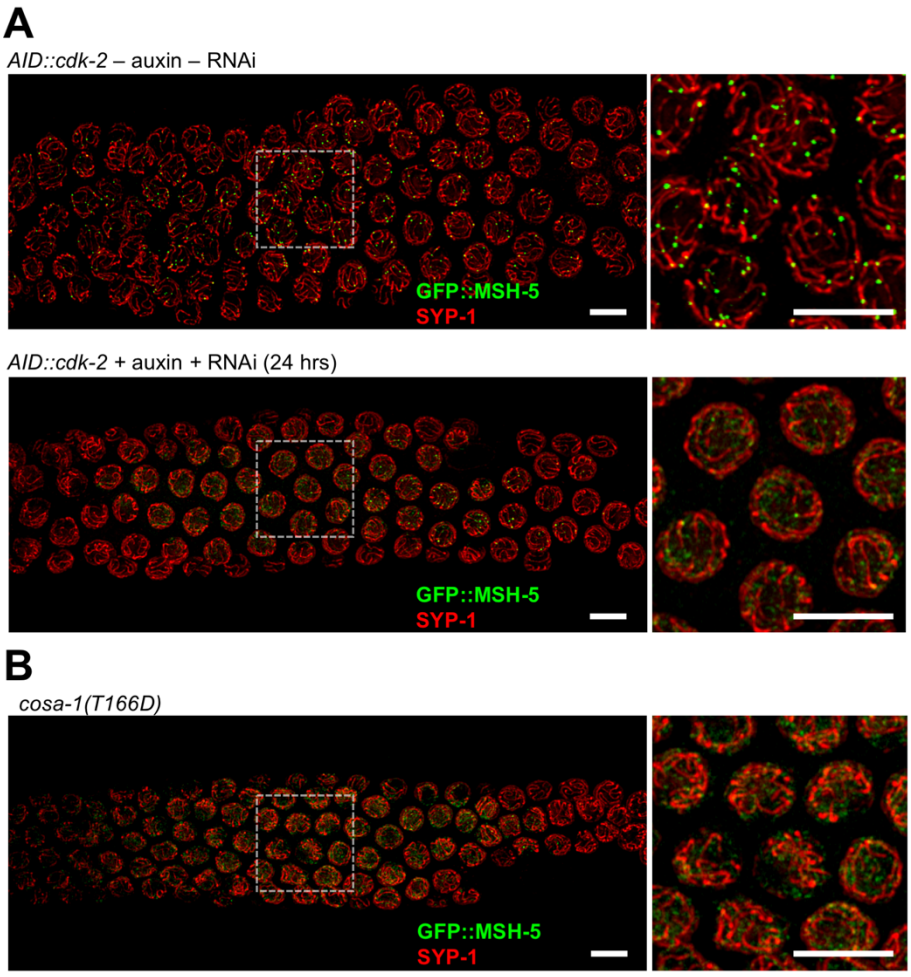

Figure 6: CDK-2 and COSA-1 similarly influence the formation of MSH-5 foci. (**A**) Maximum projections for representative gonads of 24-hour post-L4s. SYP-1 (red) is shown as a synapsis marker. MSH-5 (green) is shown as a CO intermediate marker. Auxin and RNAi depletion of CDK-2 performed for 24 hours. Scale bars (including insets) are 5 μm. (**B**) A representative gonad of *cosa-1(T166D,)* shown as in A. Scale bars (including insets), are 5 μm.

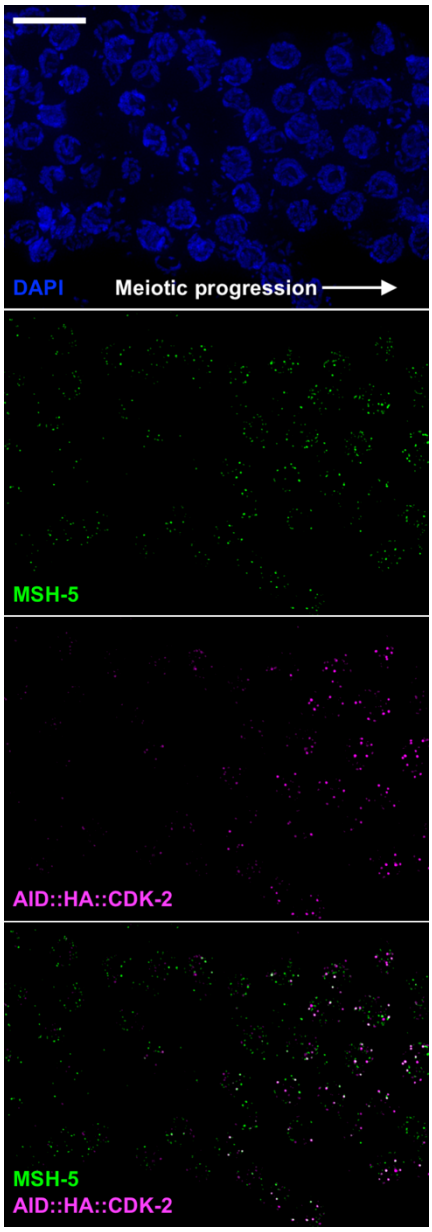

Figure 7: CDK-2 is not present in significant concentrations at CO intermediates. A maximum projection of a representative chromosomal spread is shown. DAPI (blue) marks DNA, MSH-5 (green) marks CO intermediates, and CDK-2 (magenta) is shown. Scale bar is 10 μm.

REFERENCES

1 Wang, S., Zickler, D., Kleckner, N. & Zhang, L. Meiotic crossover patterns: obligatory crossover, interference and homeostasis in a single process. *Cell Cycle* **14**, 305-314, doi:10.4161/15384101.2014.991185 (2015).

2 MacLennan, M., Crichton, J. H., Playfoot, C. J. & Adams, I. R. Oocyte development, meiosis and aneuploidy. *Seminars in Cell & Developmental Biology* **45**, 68-76, doi:https://doi.org/10.1016/j.semcdb.2015.10.005 (2015).

3 Jagut, M. *et al.* Separable Roles for a Caenorhabditis elegans RMI1 Homolog in Promoting and Antagonizing Meiotic Crossovers Ensure Faithful Chromosome Inheritance. *PLOS Biology* **14**, e1002412, doi:10.1371/journal.pbio.1002412 (2016).
4 Koury, E., Harrell, K. & Smolikove, S. Differential RPA-1 and RAD-51 recruitment in vivo throughout the C. elegans germline, as revealed by laser microirradiation. *Nucleic acids research* **46**, 748-764, doi:10.1093/nar/gkx1243 (2018).
5 Storlazzi, A., Xu, L., Cao, L. & Kleckner, N. Crossover and noncrossover recombination during meiosis: timing and pathway relationships. *Proceedings of the National Academy of Sciences* **92**, 8512-8516, doi:10.1073/pnas.92.18.8512 (1995).
6 Yokoo, R. *et al.* COSA-1 reveals robust homeostasis and separable licensing and reinforcement steps governing meiotic crossovers. *Cell* **149**, 75-87, doi:10.1016/j.cell.2012.01.052 (2012).
7 Holloway, J. K., Sun, X., Yokoo, R., Villeneuve, A. M. & Cohen, P. E. Mammalian CNTD1 is critical for meiotic crossover maturation and deselection of excess precrossover sites. *The Journal of Cell Biology* **205**, 633-641, doi:10.1083/jcb.201401122 (2014).
8 Wijnker, E. *et al.* The Cdk1/Cdk2 homolog CDKA;1 controls the recombination landscape in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **116**, 12534, doi:10.1073/pnas.1820753116 (2019).
9 Ashley, T., Walpita, D. & de Rooij, D. G. Localization of two mammalian cyclin dependent kinases during mammalian meiosis. *Journal of Cell Science* **114**, 685 (2001).
10 Viera, A. *et al.* CDK2 is required for proper homologous pairing, recombination and sex-body formation during male mouse meiosis. *Journal of Cell Science* **122**, 2149, doi:10.1242/jcs.046706 (2009).
11 Woglar, A. & Villeneuve, A. M. Dynamic Architecture of DNA Repair Complexes and the Synaptonemal Complex at Sites of Meiotic Recombination. *Cell* **173**, 1678-1691 e1616, doi:10.1016/j.cell.2018.03.066 (2018).
12 Harris, T. W. *et al.* WormBase: a comprehensive resource for nematode research. *Nucleic Acids Research* **38**, D463-D467, doi:10.1093/nar/gkp952 (2009).
13 The UniProt, C. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research* **47**, D506-D515, doi:10.1093/nar/gky1049 (2018).
14 Umaña, A. C., Iwahori, S. & Kalejta, R. F. Direct Substrate Identification with an Analog Sensitive (AS) Viral Cyclin-Dependent Kinase (v-Cdk). *ACS Chemical Biology* **13**, 189-199, doi:10.1021/acschembio.7b00972 (2018).
15 Zhang, L., Kohler, S., Rillo-Bohn, R. & Dernburg, A. F. A compartmentalized signaling network mediates crossover control in meiosis. *Elife* **7**, doi:10.7554/eLife.30789 (2018).
16 Fox, P. M. *et al.* Cyclin E and CDK-2 regulate proliferative cell fate and cell cycle progression in the *C. elegans* germline. *Development* **138**, 2223, doi:10.1242/dev.059535 (2011).

17 Goodyer, W. *et al.* HTP-3 Links DSB Formation with Homolog Pairing and Crossing Over during C. elegans Meiosis. *Developmental Cell* **14**, 263-274, doi:https://doi.org/10.1016/j.devcel.2007.11.016 (2008).
18 Rauch, J., Volinsky, N., Romano, D. & Kolch, W. The secret life of kinases: functions beyond catalysis. *Cell Communication and Signaling* **9**, 23, doi:10.1186/1478-811X-9-23 (2011).
19 Wueseke, O. *et al.* Polo-like kinase phosphorylation determines Caenorhabditis elegans centrosome size and density by biasing SPD-5 toward an assembly-competent conformation. *Biology open* **5**, 1431-1440, doi:10.1242/bio.020990 (2016).
20 Kim, Y., Kostow, N. & Dernburg, A. F. The Chromosome Axis Mediates Feedback Control of CHK-2 to Ensure Crossover Formation in C. elegans. *Developmental cell* **35**, 247-261, doi:10.1016/j.devcel.2015.09.021 (2015).
21 Kipreos, E. T. & van den Heuvel, S. Developmental Control of the Cell Cycle: Insights from *Caenorhabditis elegans*. *Genetics* **211**, 797, doi:10.1534/genetics.118.301643 (2019).
22 Chi, Y. *et al.* Identification of CDK2 substrates in human cell lysates. *Genome Biol* **9**, R149-R149, doi:10.1186/gb-2008-9-10-r149 (2008).
23 Kondo, S. & Miura, T. Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation. *Science* **329**, 1616, doi:10.1126/science.1179047 (2010).

ACKNOWLEDGEMENTS

We thank Nicola Silva, Liangyu Zhang, and Joseph Robinson for constructing and sharing strains. Finally, we thank Dernburg lab members for ongoing helpful discussions during this project.

SUPPLEMENTARY INFORMATION

Antibodies used in IF

Designation	Source or reference	Identifier(s)	Final concentration used
Goat polyclonal anti-SYP-1	[1]	-	1:500
Chicken polyclonal anti-HTP-3	[2]	-	1:500
Mouse monoclonal anti-GFP	Millipore Sigma	Cat#11814460001; RRID:AB_390913	1:500
Mouse monoclonal anti-HA, clone 2–2.2.14	Thermo Fisher Scientific	Cat#26183; RRID:AB_10978021	1:500

Rabbit polyclonal anti-V5	Millipore Sigma	Cat#V8137; RRID:AB_261889	1:500
Mouse monoclonal anti-V5	Thermo Fisher Scientific	Cat#R960-25; RRID:AB_2556564	1:500
Rabbit polyclonal anti MSH-5	ModENCODE/SDIX	Cat#SDQ2376	1:5,000
Rabbit polyclonal anti-GFP11	MyBiosource, Inc.	Cat#MBS8565385	1:1000
Secondary antibodies: AF488, Cy3, and Cy5 labeled	Jackson ImmunoResearch	-	1:1000

crRNAs/Repair templates

Target sequence	**Sequence**	**Source**
CDK-2 gatekeeper CDK-2^{as1} allele	5'tccagACAGCAAACTCTACA3' 5'aaatttaaaacaattaattaattaatcaatttccagACAGTAAGCTTTATATGGTAGGAGAATTCATCGACCGAGATCTGAAGAATCTTCTGGAAATGCTTGAGCC3'	This study
COSA-1 Threonine166 COSA-1 T166D mutation	5'ATTGCAtcTAGAATaCgaAg3' 5'GGAAGAAAAGAGAATGGGAAAATCTTGAATCGAATATGGAACGACAGATTCCGcTtcGtATTCTAgaTGCAATTCAGATTAGCAGCAAATTTCACAGTTATCATGATgtatacgcggggatttttcgaatttcaacgatgctc3'	Liangyu Zhang (Dernburg lab, UC Berkeley)
CDK-2 N-terminus AID V5 CDK-2	5'aagcttacCTGGAGCGATAT3' 5'cgtcattctttgagaaactctttaccgttaaatccttttgaaaatcttaattatcttccagAAAACAACGGGATATGggagctggatcaCCTAAAGATCCAGCCAAACCTCCGGCCAAGGCACAAGTTGTGGGATGGCCACCGGTGAGATCATACCGGAAGAACGTGATGGTTTCCTGCCAAAAATCAAGCGGTGGCCCGGAGGCGGCAGCGTTCGTGAAGgg	Joseph Robinson (Garriga lab, UC Berkeley)

	agccggatctGGAAAGCCAATTCCAAACCCACTTCTT GGACTCGACTCCACCggagctggatcaACAACTGATA TCGCTCCAGgtaagctttttccgagattttcatttttcaatatttaata aaaaataaataattaaaataatttcccatttcagAAAGAGACCTG CA3’	
CDK-2 N-terminus AID HA CDK-2	5’aagcttacCTGGAGCGATAT3’ 5’cgtcattctttgagaaactctttaccgttaaatccttttgaaaatcttaatt atcttccagaaaacaacgggatATGGGAGCTGGATCACC TAAAGATCCAGCCAAACCTCCGGCCAAGGCACA AGTTGTGGGATGGCCACCGGTGAGATCATACCG GAAGAACGTGATGGTTTCCTGCCAAAAATCAAG CGGTGGCCCGGAGGCGGCAGCGTTCGTGAAGG GAGCCGGATCTTATCCTTATGACGTACCTGACTA CGCAGGAGCTGGATCAACAACtGATATCGCTCCA Ggtaagctttttccgagattttcatttttcaatatttaataaaaaataaata attaaaataatttcccatttcagAAAGAGACCTGCA3’	Joseph Robinson (Garriga lab, UC Berkeley)

Strains used

Strains	**Source**	**Identifier**
cdk-2^{as1}I; meIs8 [pie-1p::GFP::cosa-1 + unc-119(+)] II	This study, [3]	-
AID::V5::cdk-2 I ; meIs8 [pie-1p::GFP::cosa-1 + unc-119(+)] II; ieSi38[sun-1p::TIR1::mRuby::sun-1 3'UTR, unc-119(+)] IV	This study, [4]	CA1451
AID::HA::cdk-2 I ; meIs8 [pie-1p::GFP::cosa-1 + unc-119(+)] II; ieSi38[sun-1p::TIR1::mRuby::sun-1 3'UTR, unc-119(+)] IV	This study, [4]	CA1452
AID::V5::cdk-2 I; ieSi64[gld-1p::TIR1::mRuby::gld-1 3'UTR, unc-119(+)] II; gfp::msh-5 IV	This study, [4], [5]	-
gfp11::rmh-1, AID::HA::cdk-2 I; iowSi8[pie-1p::gfp1-10::him-3 3’UTR + unc119(+)]	This study, [6], [4]	-

II; *ieSi38[sun-1p::TIR1::mRuby::sun-1 3'UTR, unc-119(+)] IV*		
3×FLAG::cosa-1^{T166D} III; gfp::*msh-5 IV*	This study, [5]	-
Bristol wild-type isolate	Caenorhabditis Genetics Center	N2

RNAi plasmid

To create the CDK-2 RNAi plasmid the below sequence of CDK-2 exon 3 was amplified and inserted into the L4440 RNAi vector[7] with Spe1 and Kpn1 digestions:
ACAGCAAACTCTACATGGTATTCGAATTTATCGATCGAGATCTGAAGAATCTTCTGG
AAATGCTTGAGCCAACGAATAGTGTGCTTCCGCCAAACTATGTGAAGTCATTCATG
TGGCAACTTCTATCAGCTCTATCGTATTGTCATCTCCGACGGATTGTTCACCGTGAT
CTGAAGCCTCAGAATATTCTGGTTTCAGATTCTGGAGTCATCAAAATTGCGGATTTC
GGATTGGCAAG

Supplementary References

1 Harper, N. C. *et al.* Pairing centers recruit a Polo-like kinase to orchestrate meiotic chromosome dynamics in C. elegans. *Developmental cell* **21**, 934-947, doi:10.1016/j.devcel.2011.09.001 (2011).
2 MacQueen, A. J. *et al.* Chromosome sites play dual roles to establish homologous synapsis during meiosis in C. elegans. *Cell* **123**, 1037-1050, doi:10.1016/j.cell.2005.09.034 (2005).
3 Yokoo, R. *et al.* COSA-1 reveals robust homeostasis and separable licensing and reinforcement steps governing meiotic crossovers. *Cell* **149**, 75-87, doi:10.1016/j.cell.2012.01.052 (2012).
4 Zhang, L., Ward, J. D., Cheng, Z. & Dernburg, A. F. The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in C. elegans. *Development* **142**, 4374-4384, doi:10.1242/dev.129635 (2015).
5 Janisiw, E., Dello Stritto, M. R., Jantsch, V. & Silva, N. BRCA1-BARD1 associate with the synaptonemal complex and pro-crossover factors and influence RAD-51 dynamics during Caenorhabditis elegans meiosis. *PLoS Genet* **14**, e1007653, doi:10.1371/journal.pgen.1007653 (2018).
6 Hefel, A. & Smolikove, S. Tissue-Specific Split sfGFP System for Streamlined Expression of GFP Tagged Proteins in the Caenorhabditis elegans Germline. *G3 (Bethesda)* **9**, 1933-1943, doi:10.1534/g3.119.400162 (2019).

7 Timmons, L., Court, D. L. & Fire, A. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. *Gene* **263**, 103-112, doi:https://doi.org/10.1016/S0378-1119(00)00579-5 (2001).

Supplementary Figures

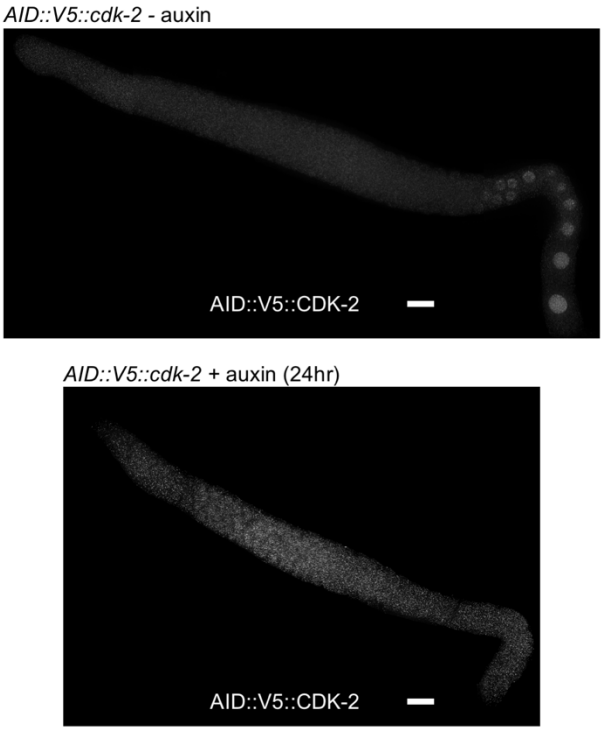

Figure 1S: Degron-tagged CDK-2 is efficiently degraded upon auxin addition. Maximum projections of 24-hour post-L4 gonads with tagged CDK-2 and transgenic TIR1 have CDK-2 present without auxin treatment. In the top, untreated gonad, CDK-2 (grey) signal can be seen to rise in the nuclei of cells in late pachytene through diplotene, while substantial background staining is visible. In contrast, AID tagged CDK-2 treated for auxin-induced degradation have no visible increase in diplotene of CDK-2. The visible signal is the same background signal present in the control and likely does not represent CDK-2 signal. The scale bars are 15 μm.

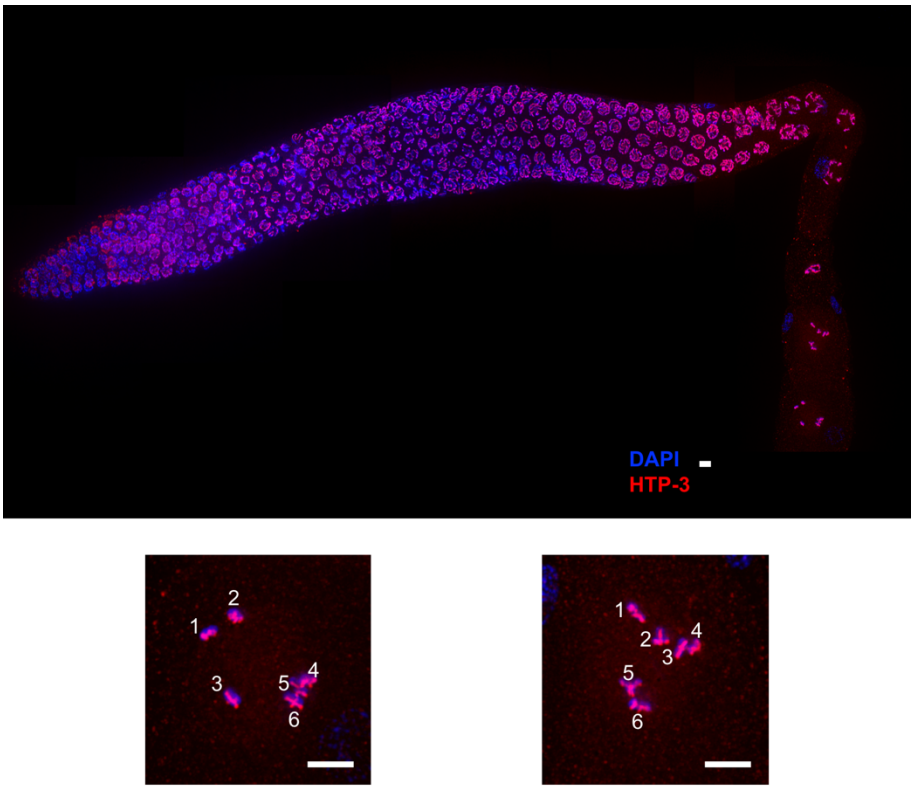

Figure 2S: In the absence of inhibitor, the analog-sensitive *cdk-2*^{as1} allele does not impair meiosis. A maximum projection of a representative 24-hour post-L4 gonad with HTP-3 (red) as a CO marker and DAPI (blue) as a DNA marker is shown. Cells progress normally through meiosis. Enlarged images below show that the -1 oocyte and -2 oocyte both form six DAPI staining bodies with normal chiasmata. These six DAPI staining bodies, along with the six COSA-1 sites per cell in **Fig. 3,** indicate that *cdk-2*^{as1} has normal meiosis and CO formation. Scale bars shown are 5 μm.

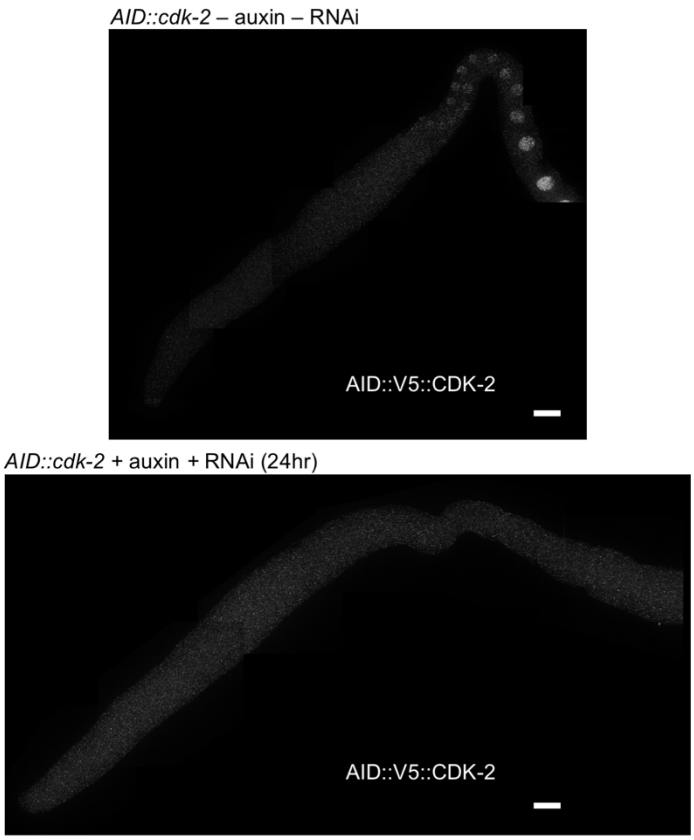

Figure 3S: RNAi and auxin treatment of AID tagged CDK-2 degrades the protein. Images are presented in the same manner as in **Fig. 1S**, and depletion results in the loss of CDK-2 signal in diplotene nuclei. The scale bars are 15 μm.