



RESEARCH ARTICLE

# Okapi-EM: A napari plugin for processing and analyzing cryogenic serial focused ion beam/scanning electron microscopy images

Luis M. A. Perdigão<sup>1</sup> , Elaine M. L. Ho<sup>1</sup>, Zhiyuan C. Cheng<sup>1,2</sup>, Neville B.-Y. Yee<sup>1</sup> , Thomas Glen<sup>3</sup>, Liang Wu<sup>3,4</sup>, Michael Grange<sup>3,4</sup>, Maud Dumoux<sup>3</sup>, Mark Basham<sup>1,5</sup> and Michele C. Darrow<sup>1</sup>

<sup>1</sup>Artificial Intelligence and Informatics, The Rosalind Franklin Institute, Didcot, UK

<sup>2</sup>School of Chemistry, University of Edinburgh, Edinburgh, UK

<sup>3</sup>Structural Biology, The Rosalind Franklin Institute, Didcot, UK

<sup>4</sup>Division of Structural Biology, Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK

<sup>5</sup>Diamond Light Source, Didcot, UK

**Corresponding author:** Luis M. A. Perdigão; Email: [luis.perdigao@rfi.ac.uk](mailto:luis.perdigao@rfi.ac.uk)

**Received:** 14 December 2022; **Revised:** 09 March 2023; **Accepted:** 13 March 2023

**Keywords:** Cryogenic; image processing; napari; serial FIB/SEM; software; 3D imaging

## Abstract

An emergent volume electron microscopy technique called cryogenic serial plasma focused ion beam milling scanning electron microscopy (pFIB/SEM) can decipher complex biological structures by building a three-dimensional picture of biological samples at mesoscale resolution. This is achieved by collecting consecutive SEM images after successive rounds of FIB milling that expose a new surface after each milling step. Due to instrumental limitations, some image processing is necessary before 3D visualization and analysis of the data is possible. SEM images are affected by noise, drift, and charging effects, that can make precise 3D reconstruction of biological features difficult. This article presents Okapi-EM, an open-source napari plugin developed to process and analyze cryogenic serial pFIB/SEM images. Okapi-EM enables automated image registration of slices, evaluation of image quality metrics specific to pFIB-SEM imaging, and mitigation of charging artifacts. Implementation of Okapi-EM within the napari framework ensures that the tools are both user- and developer-friendly, through provision of a graphical user interface and access to Python programming.

## Impact Statement

Cryogenic serial plasma focused ion beam milling scanning electron microscopy is an emerging microscopy technique that is used to visualize 3D structures of biological features at mesoscale resolutions. This technique requires common postprocessing of data such as alignment and charge mitigation to enable robust segmentation and analysis. In addition, approaches are needed to quantify data quality to enable an assessment of features and tune data acquisition parameters to enable optimal image acquisition. This article presents Okapi-EM, a combination of software tools designed to facilitate these important initial steps in assessing and processing images from these experiments. These tools have been assembled as a plugin for a popular 3D biological image visualizer called napari, making their usage user-friendly and readily accessible.

## 1. Introduction

Recent advances in cryo-electron microscopy hardware have seen an emergence of instruments which combine microscopy techniques and milling instruments within the same system. Dual-beam focused ion

beam scanning electron microscopy (FIB/SEM), maintained at cryogenic temperatures, provides a workflow to acquire volumetric SEM images of a range of biological samples in their near-native state at nanometer-resolution<sup>(1–3)</sup>. This technique builds volumetric representation of the specimen by cyclic FIB-milling (to remove the freshly imaged surface) and SEM imaging of the specimen. Several SEM images (typically hundreds) are obtained, corresponding to decreasing heights of the specimen, which can be computationally stacked to obtain a volumetric representation of the sample being measured.

Serial FIB/SEM has historically been used to capture images of fixed, stained, and resin-embedded samples providing cellular and subcellular imaging in 3D that can be used for the reconstruction and analysis of biological features. However, fixation, staining, dehydration, and resin embedding can introduce artifacts<sup>(4–8)</sup>. With the development of cryogenically stable microscope stages and reduced rates of ice formation, it is possible to image cryogenically prepared samples providing structural information of these samples in near-native states. Cryogenic serial pFIB/SEM has been done using modified room temperature plasma ion beam milling microscopes which were designed for physical science applications and more recently has been further enabled through the development of fit-for-purpose commercial options, some of which offer multiple plasma generation gases<sup>(1,9–14)</sup>.

As with many experimental techniques which produce 3D data, it is often desirable to annotate biological features and to visualize the structure in three dimensions (3D), but important preprocessing steps are needed before data is suitable as input for these segmentation tasks. For instance, small translational movements between images within the stack caused by stage and/or sample movement are often observed as misalignment between images; while compensatory functions exist within the instrument it is not always possible to correct on-the-fly. Therefore, these must be compensated for, otherwise volumetric segmentation<sup>(15,16)</sup> and computational counting tools like connected components algorithms will fail or struggle to succeed. Additionally, SEM images of biological samples often contain artifacts caused by charging around insulating substances such as lipids<sup>(3,12)</sup>. Automatic and semi-automatic segmentation tools require aligned datasets and data that can be effectively normalized to remove any strong features generated by sample charging. Another common issue observed during pFIB/SEM imaging is the creation of curtaining artifacts during the milling step which are then visible as streaks in the milling direction during SEM imaging. Finally, having quantitative tools that assess the quality of the data under certain imaging (e.g., optimal focus and voltage) and milling conditions (e.g., focusing, curtaining, and milling accuracy) will assist in the generation of data that best mitigate these factors, resulting in optimal further processing and optimization of future data acquisition strategies<sup>(15,16)</sup>. These necessary presegmentation tasks can be time-consuming and often require use of multiple pieces of software or bespoke code. Okapi-EM provides a selection of tools which address some of these needs in a single software package found within napari<sup>(17)</sup>. In this first release of Okapi-EM, there are three tools available:

1. Stack Alignment. This tool provides the user with appropriate transformation options for the alignment of stacks of slices.
2. Charge mitigation (Chafer). This tool requires presegmented “charge centers,” then applies filters to mitigate the charging artifacts found nearby.
3. Resolution estimation (Quoll). This tool requires microscope calibration and provides a measure of the mean resolution and standard deviation for individual slices.

Similar collection of tools for processing serial-FIB/SEM data are also available in *bmiptools*<sup>(18)</sup> software package, which was developed independently of Okapi-EM. However, this is not a napari plugin, therefore requiring users knowledge of programming in order to use it. Unfortunately, at the time of writing this article (*bmiptools* version 1.0.1), execution of its charge artifact removing filter was attempted in our image stack data set (see below), but it resulted in no difference and no suppression.

Example data used throughout this manuscript are available on EMPIAR (see Data Availability) and their sample preparation and pFIB/SEM data collection have been described in Dumoux *et al.*<sup>(1)</sup> and are briefly detailed here. Biological samples were vitrified by plunge freezing (yeast) or by high-pressure

freezing (mouse brain). The microscope used is a dual beam FIB/SEM “Helios G4 Hydra” equipped with an Aquilos II type cryo-stage. SEM electron beam voltage and current were 1 kV and 12.5 pA for the yeast and 1.1 kV and 6.25 pA for the brain. Quoted SEM image angles are relative to sample plane. Argon was used for milling, and the accelerating voltage used was 20 kV.

Typically, plasma-FIB (pFIB) uses ionized gases as milling source, and traditional FIB often refers to ion beam using liquid metal sources, such as gallium. We highlight here that Okapi-EM was designed to process pFIB-SEM images, as it has been thoroughly tested with this data. Nevertheless, Okapi-EM may prove to be equally useful to process data generated by other ion beam source, though usage of Okapi-EM with this type of data is untested.

## 2. Implementation

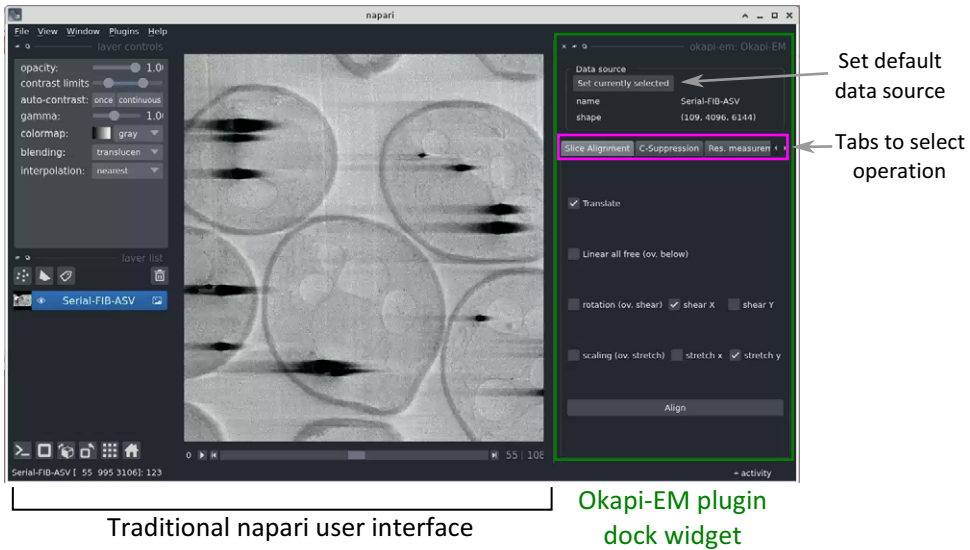
Napari is an open-source, user-friendly, data visualization application that runs in Python programming language<sup>(17)</sup>. It supports three-dimensional visualization of different data types (i.e., image and labels) interactively. Its annotation features are useful in biological image analysis and processing. Furthermore, its functionality can be extended with support for plugins, allowing continuous development of image processing methods alongside of experimental microscopy advances. Developing within the Python ecosystem also facilitates deep learning approaches with easy access to modern machine learning packages such as scikit-learn<sup>(19)</sup> and PyTorch<sup>(20)</sup>. The plugin installation uses the well-established package management system (PyPI), which also enables plugin installation chains, where a single plugin can install many other plugins or packages as needed to create a *bundle*. Bundling several plugins in this way means that a whole toolbox can be created for specific data processing workflows or research themes (e.g., devbio-napari<sup>(21)</sup>, or napari-assistant<sup>(22)</sup>, both of which were tailored for biological image processing). Napari supports and opens a range of image formats through the *skimage* library, including multi-page tiff. It also hosts a variety of other image processing and file format plugins so Okapi-EM tools can be integrated into and combined with other workflows. Okapi-EM can be installed through napari’s plugin search and installer engine. Otherwise, it can be downloaded freely at <https://github.com/rosalindfranklininstitute/okapi-em>, or installed from Python package index (PyPI). Okapi-EM was developed for minimum python version 3.7 and installations of other dependency packages using PyPI or napari’s plugin installation engine are automatic.

Currently, Okapi-EM incorporates three major image processing tools which are organized in the user interface as separate tabs: stack alignment, charge suppression, and resolution measurement (Figure 1). Each tab contains user interface elements to adjust settings and execute data analysis.

## 3. Stack Alignment

In FIB-SEM, samples are milled in preparation for imaging. During this process, a number of issues can cause drift or misalignment in the image stack, such as the movement of the sample due to mechanical stress, small temperature variations and slight, iterative, stage movements going from milling to imaging positions that accumulate over time<sup>(23)</sup>. Inaccurate placement of the milling area by the operator or software may also lead to the observation of shearing during subsequent imaging<sup>(24)</sup>. These misalignments may further be amplified by factors including charging effects and instabilities caused by external disturbance<sup>(24)</sup>. As a result, an important element in the alignment of the resultant SEM images is the shearing or skewing along the direction perpendicular to the ion beam (or other layer removal option) trajectory<sup>(23)</sup>, which is particularly significant along the slow scan direction. Before any subsequent visualization or analysis tasks, aligning the image stack is a crucial step.

While there have been efforts at improving image acquisition hardware and developing real-time correction<sup>(25–27)</sup>, these methods are limited by the need to ensure that the sample is not overexposed during data collection. Therefore, they do not provide fine alignment correction and require the use of subsequent alignment software<sup>(24)</sup>. Several software packages/plugins are available for this task, such as the closed-source and pay-for-service Amira by ThermoFisher<sup>(28)</sup> and open-source options such as



**Figure 1.** View of the napari application highlighting the Okapi-EM plugin (green rectangle at right), and the currently available tools (pink rectangle at right). Each Okapi-EM plugin has its own tab with appropriate options for its use displayed. See *Supplementary Figure S1* for a detailed view and description of the options available in each plugin.

ImageJ plugins Linear Stack Alignment with SIFT<sup>(29)</sup> and TrakEM2<sup>(30)</sup>. Although these software can in general align the images so that there is smooth transition between the slices, a major drawback is that they do not consider the physical process by which the images were acquired, thus they can introduce distortions that are not present in the sample. During alignment, if incorrect transformation parameters are chosen, because slices are aligned to their neighbors, this can cause a cascade of transformations that can ultimately distort the shape of the stack<sup>(31)</sup>. Notably, the ImageJ plugins Linear Stack Alignment with SIFT and TrakEM2 do not offer options to perform alignment with shearing instead of rotation in particular directions, which is what images obtained by scanning line-by-line need to be compensated for.

### 3.1 Alignment method

We have developed an improved image alignment method that aims to provide users with a variety of transformation options to appropriately describe image acquisition processes and to reduce the likelihood of introducing nonphysical distortions during alignment of 3D stacks. It employs scale-invariant feature transform, or SIFT, a widely used algorithm for the detection of local landmarks<sup>(32)</sup>. Since SIFT tends to include some false matches that could affect the alignment outcome, filtering is necessary to remove these outliers. Users can choose between percentile and RANSAC filters to achieve this. The former option is faster but less robust, while the latter is more computationally costly, but more reliable in cases where matches are scarce, or exhibiting misalignment. Users will then define a transformation that would best describe the physical distortion of the images depending on the image acquisition technique (Table 1) by adding or removing components of the transformation matrix, including translation, rotation, shearing, uniform scaling and stretching or shrinking, or selecting affine transformation which allows all six degrees of freedom. A description of each option is shown in *Supplementary Figure S2*. For instance, if the option shearing along the  $y$ -axis is enabled, and no scaling or shearing in  $x$ -axis are chosen, then the transformation matrix and translation vector to be obtained is:

$$\vec{p}' = M_{\text{she},x} \cdot \vec{p} + \vec{t} \quad (1)$$

**Table 1.** Types of transformations that are commonly used to correct distortions from different volumetric electron microscopy techniques.

	Translation	Rotation	Shear	Scale	Stretch	Affine
Block-face imaging, laser sectioning (e.g., FIB/SEM, cryo FIB/SEM)	Yes	No	In scanning direction	No	In scanning direction	No
Block-face imaging, mechanical sectioning (e.g., SBF SEM)	Yes	No	In scanning and sectioning directions	No	In scanning direction	No
Imaging of thin serial sections (e.g., ssTEM) <sup>a</sup>	Yes	Yes	In sectioning direction	No	No	Yes
Tomographic imaging of thin serial sections (e.g., electron tomography) <sup>b</sup>	Yes	Yes	In scanning direction	No	No	Yes

<sup>a</sup>Deformations between subsequent serial section images.

<sup>b</sup>Deformation between projection images of the same thin section taken at different tilt angles.

or:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_t \\ y_t \end{bmatrix}, \quad (2)$$

where  $\vec{p} = (x, y)$  is a feature point on the source image and  $\vec{p}' = (x', y')$  is its location after shearing with  $M_{she,x}$  being the linear transform for shearing along  $x$  by the parameter of  $k$ , and translation of  $\vec{t} = (x_t, y_t)$ .

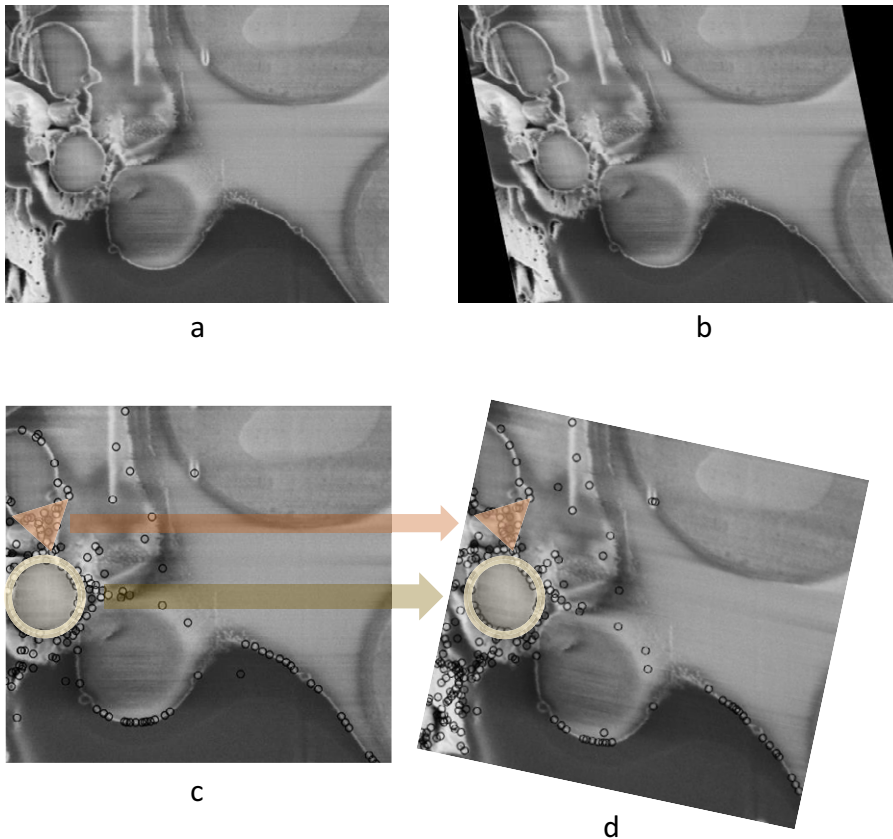
The alignment algorithm starts from the first image slice and calculates the optimized transformation parameters between consecutive slices along the whole stack. Similar features between images are identified and matched. The distance between a pair of feature points is  $d = |\vec{p}' - \vec{p}|$  and, considering all the feature points pairs between images, the sum of the squared distances is minimized by adjusting the parameters  $k, x_t, y_t$ .

### 3.2 Rotational distortion introduced with existing software

As discussed previously, limiting the modes of transformation according to the physical process of image acquisition is crucial. Failing to do so when aligning FIB-SEM image stacks could result in a large angle rotation in the image slices that is not actually present in the sample. To demonstrate this, an artificial shearing transformation was applied to the original image (Figure 2a) acquired using cryogenic serial pFIB-SEM to mimic a shear force distortion (Figure 2b). For this image pair, using SIFT, many matches can be found on the left side of the image, especially in the triangular and circular region (Figure 2c). Using least square or other optimization methods without restraining the transformation mode results in a large angle rotation or an affine transformation that aligns those feature-rich regions well, but fails to align other parts of the image and introduces a nonphysical distortion (Figure 2d).

### 3.3 Results

Okapi-EM stack alignment was applied to a cryogenic serial pFIB/SEM image stack of yeast cells (109 slices). As described in Table 1, translation, shear  $x$ -axis, and stretch along  $y$ -axis were chosen to

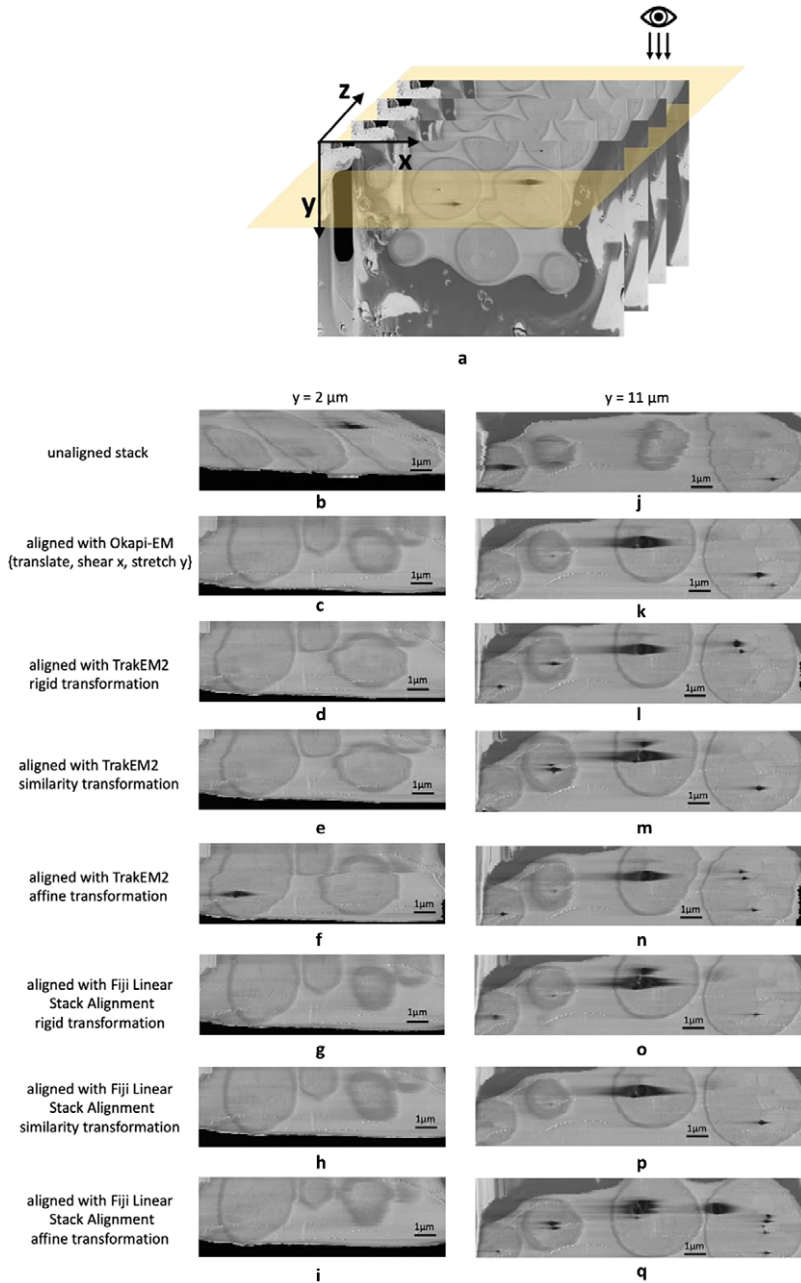


**Figure 2.** *Nonphysical distortion in the alignment process if the modes of transformation are not appropriately limited. (a) Original SEM image of yeast cells. (b) Artificially distorted image after a shearing transformation was applied to the original image. A shear factor of 0.15 was selected for visibility. (c) Feature points found in (b) using SIFT, with two feature-dense regions highlighted with circle and triangle markers. (d) Possible alignment result when the modes of transformation are not restrained, where even though the feature-rich regions are well-aligned, a nonphysical rotation is introduced. Data is FIB-SEM of yeast, available at EMPIAR with ID 11416.*

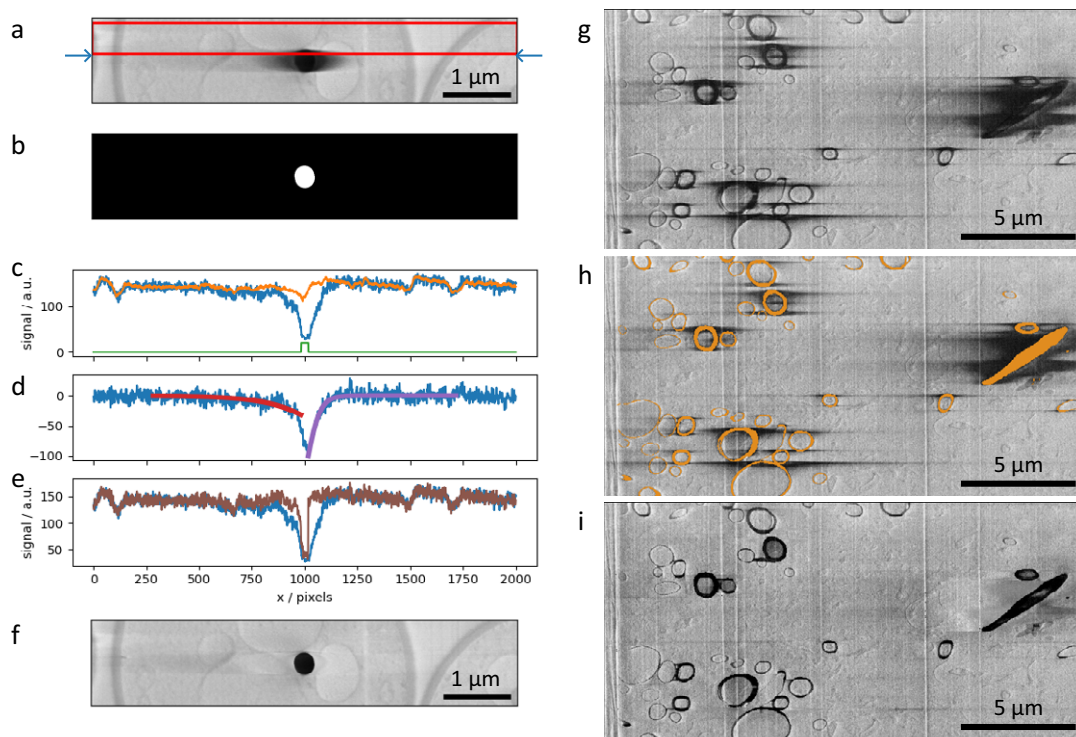
compute an alignment based on the landmarks found using SIFT with RANSAC outlier filtering. Cross-sectional views are used to visualize the alignment result (Figure 3a). To compare outcomes, the same image stack was aligned in Fiji plugins TrakEM2 and Linear Stack Alignment with each transformation option offered as well. Substantial drift and distortion can be observed in the unaligned stack (Figure 3b,j and Supplementary Movie S1). After alignment, cell shapes are restored with mostly smooth outlines (Figure 3c,k and Supplementary Movie S2). While all alignment approaches improved the cross-sectional alignment to some degree, distortion can still be observed in the outline of the cells when using either TrackEM2 or Linear Stack Alignment (Figure 3d-i,l-q), especially at the  $y = 2 \mu\text{m}$  when using TrakEM2 and when affine transformation is selected in Fiji Linear Stack Alignment.

#### 4. Charge Artifact Mitigation

Charging artifacts often appear when insulating materials interact with the electron beam<sup>(1,3)</sup>. This effect is normally minimized by adjusting beam energy or scanning parameters, and its severity depends on the target substance being imaged. Biological samples can present a challenge, with cells often containing



**Figure 3.** Alignment results and comparison between Okapi-EM, TrakEM2, and Fiji Linear Stack Alignment with SIFT (a) Cross-sectional views of SEM image stack of yeast cells ( $109$  slices along  $z$  direction, slice size  $20.7 \times 13.8 \mu\text{m}^2$ ) are obtained. (b–i) Cross-sectional view of a  $11.2 \times 3.7 \mu\text{m}^2$  area at  $y = 2 \mu\text{m}$  of the unaligned stack and the aligned stacks using Okapi-EM alignment with RANSAC and {translate, shear  $x$ , stretch  $y$ } selected, Fiji plugin TrakEM2 with rigid transformation, similarity transformation, or affine transformation selected, respectively, and Fiji plugin Linear Stack Alignment with rigid transformation, similarity transformation, or affine transformation selected respectively. (j–q), Cross-sectional view of a  $15.0 \times 3.7 \mu\text{m}^2$  area at  $y = 11 \mu\text{m}$  of the aligned stacks using the three aforementioned alignment methods and settings respectively. Data is EMPIAR-11416. In both Fiji plugins, rigid transformation allows translation and rotation. Similarity transformation allows translation, rotation, and scaling. Affine transformation is defined as shown in Supplementary Figure S2. Details about the rendering of the cross-section images is detailed in Supplementary Figure S8.



**Figure 4.** (a–f) SEM images and plots illustrating the charge artifact removal algorithm on a lipid droplet from a yeast dataset. Data is EMPIAR-11416. (a) SEM image of a lipid droplet within a yeast cell (image size  $6.75 \times 1.35 \mu\text{m}^2$ ). Arrows indicate a scanning line of interest, with its signal profile in (c–e) as blue line. Red rectangle represents data region where signal was averaged (width set by the *nlinesaverage* parameter) resulting in signal profile in (c) as orange line. (b) Annotation image of the charge center, corresponding to the charging artifact in Figure (a). Green line in plot (c) is the line profile of the annotation along the row of interest (arbitrary units). (d) Line profiles describing the background corrected signal (blue line) and the optimized functions  $f_{\text{left}}$  and  $f_{\text{right}}$  (red and purple) respectively. (e) Line profiles of the uncorrected signal (blue line) and corrected signal (brown line). (g–i) SEM images illustrating charge artifact suppression on myelinated sheaths found in mouse brain (image size  $20.7 \times 10.3 \mu\text{m}^2$ ). Data is EMPIAR-11415. (g) Original SEM image displaying copious amounts of charging. (h) SEM image with overlaid segmentation of charging centers and extending to complete rings of myelin sheaths. (i) Result after applying filter with default parameters.

different compartments with distinct electrical conduction properties that cannot be completely balanced through adjustments in acquisition parameters. Inevitably, images will show evidence of charging, causing artifacts which manifest in the SEM images collected in the form of elongated dark regions in the scanning direction that extend asymmetrically beyond the charging feature itself (Figure 4a,g). The practical outcome of this artifact is the partial obscuring of biological features nearby to insulating materials. This makes both manual and automated downstream data processing and analysis (e.g., segmentation or visualization) more difficult.

A previous attempt to mitigate charging in SEM images presented by Spohner *et al.*<sup>(33)</sup> utilizes the python scikit-image dilation/morphology function<sup>(33)</sup>. This method separates the charging tail signals from the SEM images and then partially subtracts them from the original image. In our hands, this method did not work with the charging artifact tails in our datasets (Supplementary Figure S6), possibly indicating it is specific to the datasets it was developed for or to the instrument used to collect them.



Okapi-EM contains a charging artifact suppression tool<sup>(34)</sup> which is designed to restore the image contrast within and around charging artifacts, while retaining the charge centers themselves. As the artifact appears elongated along the direction of scanning (laterally) and is therefore influenced by the rastering nature of the scanning, it immediately suggests that to reverse this effect in the images, a row-by-row filter that uses information of the surrounding areas, in particular along the direction perpendicular to scanning, should be used to subtract the charging effect. With Okapi-EM's chafer, the restoration method operates sequentially row-by-row in down and up passes, and the estimation of the charging artifact tails is done by fitting with a smoothing function, which is later subtracted.

For this filter to work a semantic segmentation of the charging centers must be provided as a separate layer in napari. Within the plugin user interface, there is a field where user sets this layer as being the annotation of the charging artifacts (Supplementary Figure S1B). This segmentation can be done manually or by using shallow or deep-learning<sup>(35)</sup> predictors using either tools within napari (as demonstrated here) or tools available elsewhere<sup>(28,36)</sup> and then loaded in napari. In our experience, simple thresholding methods to label either charging centers or full artifacts do not work well due to the presence of other features of similar intensity. Correct annotation of the charging centers is crucial as it provides both an inverted mask of charging locations for correction and an indication of where the charging artifact is relative to the charging center, hence being useful for choosing initial values during optimization of the functions used (see below).

The filtering scheme works in the following way. First, it takes the previous rows (Figure 4a, red box) and averages perpendicular to the scanning direction (vertical here). This is used as an approximation of the signal without the charging effect. The difference between the current row (Figure 4a, blue arrows) and the previous-row average is assumed to be an estimation of the effect of the charging (Figure 4c–e, blue line). Simply subtracting this estimation of charging effects from the current row of data gives noisy results from row-to-row. Instead, the tails of the charging effect signal (artifact that is outside of the charging center) can smoothed or modeled by curves, and when subtracting this from the original signal, it gives results that are more consistent from scanning row to row. Throughout this data processing tasks, the data within the charging artifact region is not considered in the calculations; the charging artifact label in Figure 4b is used as a negative mask and also to direct the direction of the function to optimize in either side (see below).

We have trialed several functions that may fit and optimize best to the charging artifact tail signals. Exponential ( $Ae^{-kx}$ ) and Gaussian functions ( $Ae^{-x^2/\sigma^2}$ ) can fit reasonably well to these tails but overall give poor results when reconstructing images, in particular at the function optimization step the Gaussian filter often fails to converge to suitable parameters (see Supplementary Figure S5). Some of the reasons why it fails to converge is that the signal is often flat in the rows above the charging artifacts, while other times, there are other biological features nearby that interfere with the function fitting, leading to unrealistic curve fitting parameters. Instead, we found that a shifted logistic sigmoid function (or Fermi-Dirac distribution type) works very well for this task. This function appears like a smoothed step function and is widely used in machine learning algorithms as an activation function, having the advantage that it “saturates” on either side of the curve, which is more characteristic of the charging artifact tails observed. Because the tails themselves were asymmetrical, the functions used to mitigate them were different for the left side and right side of the charging artifacts, and given by:

$$f_{\text{left}} = A \left( \frac{1}{e^{\frac{(x-x_0)}{\sigma}} + 1} - 1 \right), f_{\text{right}} = A \left( -\frac{1}{e^{\frac{(x-x_0)}{\sigma}} + 1} \right) \quad (3)$$

with  $A$ ,  $x_0$ , and  $\sigma$  being parameters to be fitted, and functions  $f_{\text{left}}$  fitted on the left of the labeled artifact, and  $f_{\text{right}}$  fitted on the right side.

After optimizing  $f_{\text{left}}$  and  $f_{\text{right}}$  (Figure 4d, red and purple curves, respectively) these functions are subtracted from the row signal (excluding masked regions), resulting in a charge-mitigated signal that still represents the charging object (Figure 4e, brown). Running this process row-by-row in up-down and down-up passes results in a filtered image (Figure 4f), which is a significant improvement compared to

Figure 4a, suppressing the main artifacts, while recovering some of the washed-out data previously hidden below.

In addition to SEM images of lipid droplets found in yeast cells, a second dataset featuring myelinated sheaths from mouse brains was used to test this process (Figures 4g–i). The manual segmentation of charging centers (Figure 4h) included both the charging centers and whole organelles, despite many regions not displaying charging artifact tails (including these noncharging regions has no negative impact, see Supplementary Figure S7). The filtered image (Figure 4i) demonstrates a substantial visual improvement to the image quality. We note that in regions near significant charging effects, although the contrast can be matched to the remaining image, there are no recoverable biological features present. This is particularly noticeable around the elongated diagonal biological feature on the right side of the image.

The main advantage of mitigating charging artifacts is the improved contrast of organelles near the charge centers. As such, both manual and automatic segmentation tools are expected to perform better when charging artifacts are absent, as these are notoriously sensitive to contrast changes which could lead to poor visualization and missing or misidentified organelles.

To better understand the effects of our charge mitigation scheme, we have compared the background in the vicinity of charging centers in corrected images (e.g., Figure 4f or i) to their uncorrected counterparts (e.g., Figure 4a or g), taking care to exclude regions that have been labeled as charging centers from the calculation. Through filtering, we expect to “uncover” biological features that were previously obscured by the charging tails and as such would expect a decrease in the standard deviation of the signal intensity around the charging center. The standard deviation of the signal intensity of the results shown in Figure 4f is 60% lower than the signal intensity in Figure 4a. Similarly, the standard deviation signal in Figure 4i (excluding labeled regions) is 58% less than in Figure 4g, suggesting the regions previously obscured by the charging artifacts are now in a more comparable range of signal intensity to the regions throughout the rest of the dataset. Through visual inspection, it is clear that while some information is recovered, at close proximity to the charging centers information is lost due to the presence of the charging artifact.

## 5. Resolution Estimation Using One-Image FRC

Measuring image quality enables users to evaluate the best imaging protocols for the requirements of their research question. Image resolution is a helpful measure of image quality as it can be directly related to physical dimensions, so it is intuitive to interpret. Resolution can be defined as “a maximum spatial frequency at which the information content can be considered reliable”<sup>(37)</sup>.

One method of determining image resolution is via a two-image Fourier Ring Correlation (FRC)<sup>(38)</sup>, which has been implemented in fluorescence microscopy. It is both a measure of resolution and consistency between images within a dataset. In this method, two independently acquired images of the same field-of-view are compared in the frequency domain to find the highest frequency where the images can be said to be similar. This is performed by determining the frequency where cross-correlation drops below a given threshold, often  $1/7 = 0.143$ , as commonly employed in cryo-electron microscopy<sup>(39,40)</sup>. The requirement for two independently acquired images is challenging to fulfill as it cannot be done retrospectively, and repeated acquisition could introduce image artifacts (i.e., beam damage), affecting the resolution.

Koho *et al.* proposed a one-image FRC calculation based on subsampling a single image to produce pairs of images<sup>(41)</sup>. This method arranges alternate pixels of an image in a checkerboard pattern to create subimages from a subset of pixels and then calculates the FRC between the subimages. However, in reality, the calculated FRC from these subimages is not equal to the value determined from two separate experimental images, but it can be calibrated (see below). A calibration function is then applied to match the one-image FRC from subimages to the gold-standard two-image FRC for the specific microscope and imaging conditions used.

The method of Koho *et al.* has been adapted for serial FIB/SEM imaging in a software tool called Quoll<sup>(42)</sup>. Quoll is an open-source, user-friendly tool, and library to calculate the local resolution of single images with the one-image FRC. This method can be applied to any imaging modality once the calibration

has been performed. The 2D image is first split into tiles, and the one-image FRC is calculated on each, returning a map of the spatial variation of resolution across the image, and a plot of its distribution. This process can be done on multiple 2D images within a 3D stack to understand resolution throughout the dataset. If large regions of featureless background or artifacts are present, it could skew the output as measurements would be taken on areas without appropriate levels of information present; these can be excluded through masking prior to assessment.

### 5.1 Calibration

The FRC curve obtained from the subsampling method of the one-image FRC is shifted from the gold-standard two-image FRC, where the same details in the one-image FRC curve are shifted to lower frequencies than they should be. As a result, resolution values reported by the one-image FRC are higher than the gold-standard two-image FRC, so calibration of the one-image FRC is required to match the resolution obtained by both methods. The rationale for this calibration is explained in further detail by Koho *et al.* in their original implementation of the one-image FRC resolution measurement<sup>(41)</sup>. This is an instrument-specific calibration as it is affected by the noise generation of the instrument. The calibration dataset requires two repeated images of the same field of view taken at several pixel sizes/magnifications. Ideally, there should be no significant, artificial changes between the images such as artifacts or image shifts, as these could affect the two-image FRC that is calculated between them. If necessary, image registration can be used to correct shifts and choosing specimens which do not suffer from specimen degradation through repeated imaging is helpful (i.e., inorganic options such as gold or polystyrene beads).

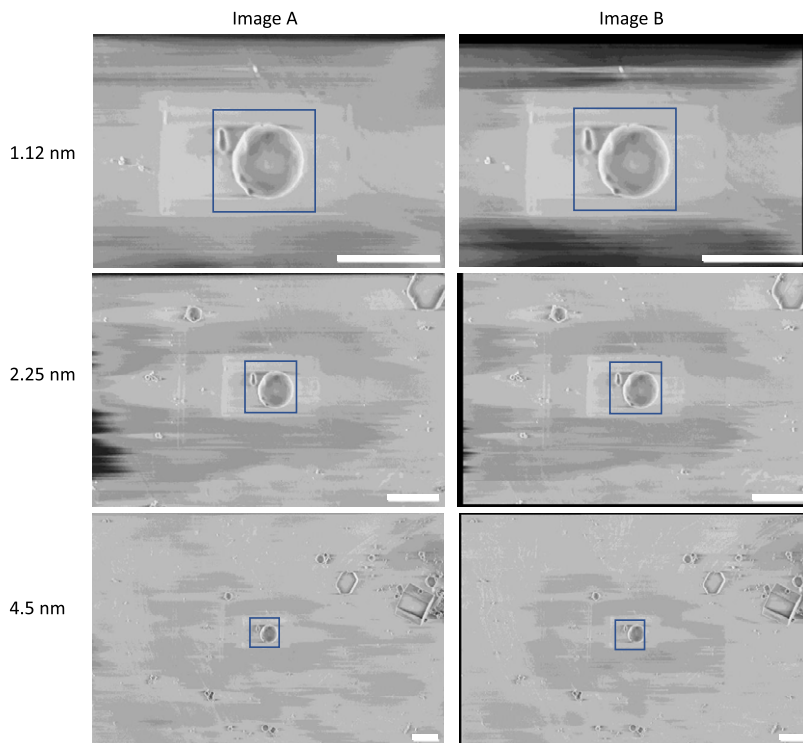
In our recent cryogenic pFIB/SEM study<sup>(1)</sup>, calibration was performed using images of polystyrene beads of 1  $\mu\text{m}$  diameter (Abvigen, Newark, NJ) in cryogenic conditions. The images show the same field-of-view at 1.12, 2.25, and 4.5 nm pixel size. The final dataset consisted of six images, with a pair of images taken at each pixel size and SEM angle (angle between the SEM and the specimen face). The linear stack alignment with SIFT plugin in Fiji 2.3.0/1.53q was used to register each pair of images to each other, to ensure the exact same field-of-view was considered for both images<sup>(43)</sup>. The images were cropped to cover the same physical field-of-view for all pixel sizes,  $2000 \times 2000 \text{ nm}^2$  for the  $52^\circ$  images and  $1600 \times 1600 \text{ nm}^2$  for the  $90^\circ$  images (Figure 5). The number of pixels in each pair of images was different due to the varying pixel size.

The FRC curve (cross-correlation vs. frequency) was calculated from each pair of images, and normalized to a scale of 0–1, where 1 was the maximum frequency calculated (Figure 6a). This normalization enabled direct comparison of FRC curves between images. The normalized frequency at which the cross-correlation fell below 0.143<sup>(39)</sup> was taken as the reference resolution ( $r_{\text{ref}}$ ). The one-image FRC resolution was also calculated following the checkerboard sampling method of Koho *et al.*, this was the  $r_{\text{col}}$ .

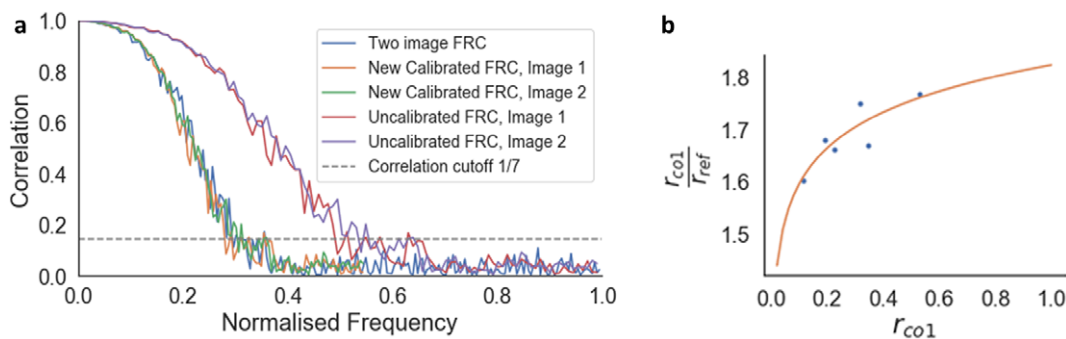
A calibration function was fitted to the plot of  $r_{\text{col}}/r_{\text{ref}}$  against  $r_{\text{col}}$  (Figure 6b). This calibration function shifted the one-image FRC curve to match the two-image FRC, so that the resolution from both methods were comparable. This calibration function is instrument-specific, so it can be reused for any images taken from that instrument, provided that the image acquisition parameters are within the range of the calibration dataset.

### 5.2 Application to cryogenic serial pFIB/SEM data

Quoll was used in a recent publication to develop cryogenic serial pFIB/SEM imaging of biological specimens<sup>(1)</sup>. Here, Quoll resolution measurement was demonstrated on five different biological specimens (*R. rubrum*, HeLa cells, Vero cells, mouse brain, *Saccharomyces cerevisiae*). The measurements were used to show that imaging at  $90^\circ$  SEM angle produced better results than  $52^\circ$ , to determine the depth of field of the instrument, and to show that there was no degradation of image quality through serial sectioning of the specimen, even at depths of  $\approx 25 \mu\text{m}$ . These performance evaluations would not have



**Figure 5.** Calibration dataset of polystyrene beads used to calculate the gold-standard two-image FRC. These images were taken at  $90^\circ$  SEM angle. A region-of-interest (blue box) covering  $2 \times 2 \mu\text{m}^2$  was used for the calibration, where the average one-image FRC and two-image FRC were calculated for these regions at all pixel sizes. Scale bars represent  $2 \mu\text{m}$ .

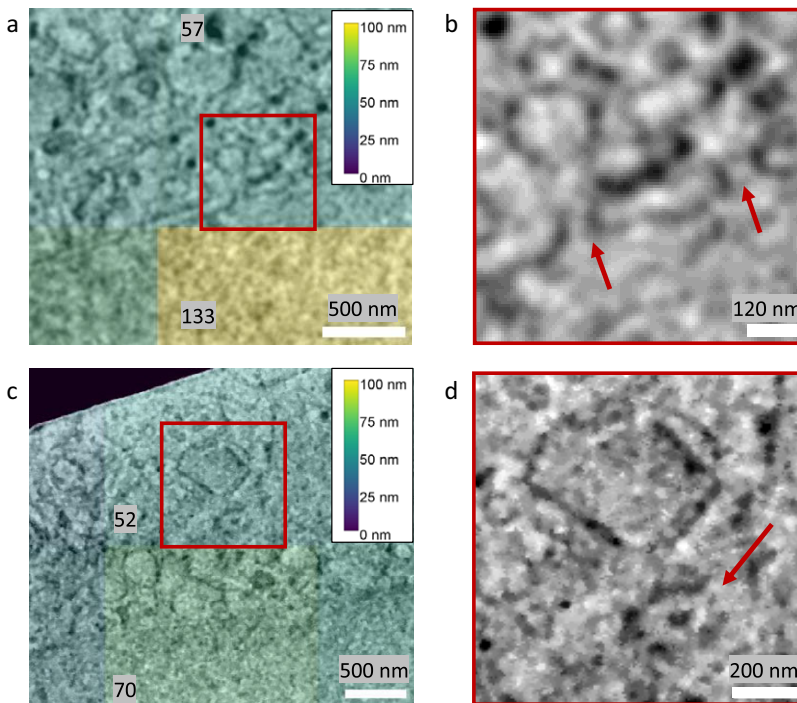


**Figure 6.** Calibration of the one-image FRC measurement to the gold-standard two-image FRC. (a) Two-image FRC curve and one-image FRC curves before and after calibration for the image pair at  $4.5 \text{ nm}$  pixel size at  $52^\circ$  SEM angle. (b)  $r_{\text{co1}}/r_{\text{ref}}$  versus  $r_{\text{co1}}$  scatter plot (blue dots) and calibration curve  $r_{\text{co1}}/r_{\text{ref}} = 2.067 + 0.099 \log(0.085r_{\text{co1}})$ , where  $r_{\text{co1}}$  is the average resolution measured from the one-image FRC for the image pair (Figure a, normalized frequency value at the location the curve crosses the correlation cutoff line of  $1/7$ ), and  $r_{\text{ref}}$  is the resolution from the gold-standard two-image FRC measured from the image pair. Calibration shifted the uncalibrated one-image FRC curves along the frequency axis to match the two-image FRC curves, ensuring that the resolution measurement for the one-image FRC matches the gold standard.

been as accurate and exhaustive without the quantitative measurements from image resolution estimation with Quoll.

The image resolution reported by Quoll was validated against biological structures with known physical sizes. The local image resolution was calculated for a cryogenic serial pFIB/SEM image of HeLa cells on tiles measuring  $256 \times 256$  pixels, which corresponded to a physical field-of-view of  $1.73 \times 1.73 \mu\text{m}^2$ . These images contained nuclear pore complexes and centrosomes, which are approximately 120 and 200 nm in diameter, respectively<sup>(44,45)</sup>. These structures were clearly resolvable in the images, and resolution in the tiles containing these structures surpassed the known diameters of these structures, validating the resolution measurements (Figure 7). FRC measurements are carried out on tiles within image slices and the reported measures provide a plot of the mean and standard deviation of resolution across the slice, as well as a heatmap for visualization. It is important to emphasize that these values are not the highest resolution visible within the slice, but instead a representation of the mean resolution of each tile. The goal of this method is to provide a quality metric for the raw data as a whole.

We found that tile size does not affect the overall resolution distribution of the image. The local image resolution was calculated on cryogenic pFIB/SEM images of *S. cerevisiae* and *R. rubrum*, at 3.37 and 1.94 nm pixel sizes, respectively. The images were sampled with tile sizes of  $128 \times 128$ ,  $256 \times 256$ , and



**Figure 7.** Validation of Quoll resolution measurements compared to physical features of known sizes in HeLa cells. Images were taken with 6.745 nm pixel size at  $52^\circ$  SEM angle. Nuclear pore complexes (a,b) of 120 nm diameter were clearly resolved in the images, as indicated by the arrows in (b), and the resolution in this region was estimated as better than the resolvable features. Similarly, in (c) and (d), centrosomes of 200 nm diameter could be resolved from the images (indicated by the arrow in d), and resolution was better than the size of the centrosomes. (a,c) A resolution heatmap is overlaid onto regions of the image showing nuclear pore complexes and centrosomes respectively, where the colors represent the resolution values of that local region and numbers are the local resolution values in nm calculated on the respective rectangular regions. (b,d) are the zoomed-in regions indicated in (a) and (c) with a red square, with arrows indicating the position of the nuclear pore complexes and centrosomes, respectively. Data is EMPIAR-11419.

**Table 2.** Summary statistics of resolution distribution measured from different tile sizes.

Tile size	Number of tiles	Mean resolution (nm)	Standard deviation of resolution (nm)	Median resolution (nm)
<i>S. cerevisiae</i> , pixel size 3.3724 nm				
128	256	45.2	9.68	45.3
256	64	47.5	10.4	46.6
512	16	44.0	6.96	45.3
<i>R. rubrum</i> , pixel size 1.94171 nm				
128	64	20.9	5.31	19.7
256	16	21.1	5.31	20.6
512	4	20.3	1.67	20.3

Note. Tile size was not found to affect the overall resolution distribution of the image. The resolution distribution median was found to be equal for all tile sizes ( $p > .05$ ) by the Kruskal–Wallis H-test. All measurements are rounded to three significant figures.

512 × 512 pixels. The difference in minimum and maximum median resolutions for all tile sizes was within 0.38–0.48 pixels. The Kruskal–Wallis H-test was applied with the null hypothesis that the population median of all groups was equal<sup>(46,47)</sup>. The null hypothesis could not be rejected ( $p > 0.05$ ) so the median resolution for all three tile sizes was considered equal (Table 2).

## 6. Discussion

Cryogenic serial pFIB/SEM imaging provides exciting opportunities for *in situ* structural biology, though as a relatively new method of imaging, requires development of appropriate computational tools both for assessment of the method and processing of the data to enable biologically relevant qualitative and quantitative outcomes. Okapi-EM has been developed to begin this process. It includes three plugins to align serial SEM stacks, mitigate charging artifacts and to assess the resolution of SEM data. In future, Okapi-EM will also include a separate plugin to measure and mitigate curtaining artifacts.

Many of the approaches developed here for cryogenic serial pFIB/SEM are likely also applicable to nonplasma-based and/or room temperature SBF/SEM or FIB/SEM or serial TEM (serial section TEM or array tomography) with little or no modifications needed to the method or implementation. If modifications are needed, we are happy to adapt Okapi-EM to meet these needs and encourage feedback from users and developers through contacting the corresponding author here or via our GitHub page.

Okapi-EM will continue to be developed into a more automated, quantitative workflow for data processing. For serial FIB/SEM, alignment of 3D stacks is generally one of the first data processing steps and its outcome can have an impact on all downstream processing and analysis. It is therefore important to apply the minimum amount of transformations (i.e., change the data as little as possible), but do enough to ensure the data is interpretable. Finding this balance is currently a manual process. In the future, development of image alignment quality metrics will enable this to shift to an automated process.

Currently, the next step in data processing for serial FIB/SEM data is mitigation of artifacts such as charging. As described above, our implementation relies on the prior identification of charging centers through either manual or machine learning-based segmentation approaches. Our focus in this area will be on the development of a more automated approach to charge center identification and segmentation. Additionally, we would like to extend our charge mitigation approach to data which display charging in visually different ways. We are aware of charging which is characterized by artifact tails that are bright white surrounding black charging centers or a mixture of white and black charging artifact tails again surrounding black charging centers. Further testing of the filters suggested here, and others are necessary to identify or adapt mitigation strategies for these data types. However, this testing has been hindered by lack of access to publicly archived example data (e.g., in EMPIAR).

Finally, once the data has been aligned and artifacts mitigated, the next steps are visualization, assessment, and segmentation. The FRC resolution measurement is a first step toward quantitative assessment of the output data though others related to “segmentability” (i.e., contrast, information content, presence/absence of a feature of interest) could be developed.

It would also be valid to use image quality assessments at the beginning of the process—data collection—guiding users in optimizing their imaging protocols, where acquisition settings can be determined based on the specified quality metrics and the researcher’s specific question(s). This is especially helpful for new users who may not have the experience to quickly determine the optimum combination of acquisition settings, so image resolution and other quality measurements at the microscope could guide the user to quantitatively optimize their protocols, removing some of the superstitious/anecdotal approaches that are sometimes found in research.

This thought process can then be extended to automated microscopy methods, where acquisition settings (such as electron beam energy) are adjusted during serial imaging to obtain good quality images throughout the session even as the sample features change (e.g., automatically decreasing charging or curtaining artifacts as they appear). This would enable higher quality, longer image acquisition sessions as the user is not required to manually adjust the imaging settings throughout the session. Real-time measurements and parallelized computational approaches will be required for these assessments to quickly identify and react to issues during imaging.

Image quality information can also be saved as metadata with each acquired image or stack of images, which is useful for indexing the images for future reuse. For example, future users could search for images of specific specimen types with a minimum resolution to answer their research question, or methods developers could search for images with low resolution to explore how their methods could improve existing images.

## 7. Conclusion

Volume EM (room temperature or cryogenic) provides exciting opportunities for visualization and quantification of cellular and tissue components, though in many cases, the outcomes of these studies are hampered by the manual nature of data processing and analysis. We have provided a bundle of plugins within the napari data visualization package to speed up the process and ease the burden on researchers using serial SEM or TEM imaging techniques. These tools allow for the alignment of 3D stacks without the introduction of unnecessary transformations, the mitigation of charging artifacts caused during scanning imaging techniques and the assessment of data quality through one-image resolution measurement. These tools are in regular use and development, and we welcome feedback and contributions as they are extended to new data types, imaging modalities, and purposes.

**Authorship contribution.** Data acquisition: M.D. and T.G. for the yeast data, T.G. and M.G. for mouse brain data, M.D. for the beads FRC calibration dataset, and L.W. for the HeLa FRC validation dataset. Manuscript: L.M.A.P., E.M.L.H., and Z.C.C. wrote the manuscript in consultation with M.B. and M.C.D. All authors reviewed the manuscript prior to submission. Project management: M.B. and M.C.D.; Software development: L.M.A.P. for Okapi-EM graphical user interface plus bundling and chafer filter, E.M.L.H. for Fourier ring correlation software (Quoll), Z.C.C. for SIFT stack alignment, and N.B.-Y.Y. for general discussions and ongoing development of additional tools.

**Competing interest.** M.C.D. is an employee of SPT Labtech. The remaining authors declare that they have no competing interests.

**Funding statement.** The Rosalind Franklin Institute is funded by UK Research and Innovation through the Engineering and Physical Sciences Research Council. Funding was also provided by the Wellcome Trust through the Electrifying Life Science grant 220526/Z/20/Z. L.W. is supported by a Sir Henry Dale Fellowship, jointly funded by the Wellcome Trust and the Royal Society (218579/Z/19/Z).

**Data availability statement.** Cryo-serial-FIB/SEM sample data used for alignment and filtering is available on EMPIAR (<https://www.ebi.ac.uk/empiar/>) alongside Dumoux *et al.* <sup>(1)</sup>. *Saccharomyces cerevisiae* (yeast) data set is EMPIAR-11416, HeLa cells is EMPIAR-11419 and mouse brain is EMPIAR-11415.

**Supplementary materials.** The supplementary material for this article can be found at <http://doi.org/10.1017/S2633903X23000119>.

## References

- Dumoux M, Glen T, Ho EML, *et al.* (2022) Cryo-plasma FIB/SEM volume imaging of biological specimens, p. 2022.09.21.508877. <https://www.biorxiv.org/content/10.1101/2022.09.21.508877v1> (accessed October 25, 2022).
- Vidavsky N, Akiva A, Kaplan-Ashiri I, *et al.* (2016) Cryo-FIB-SEM serial milling and block face imaging: large volume structural analysis of biological tissues preserved close to their native state. *J Struct Biol* **196**, 487–495.
- Spehner D, Steyer AM, Bertinetti L, *et al.* (2020) Cryo-FIB-SEM as a promising tool for localizing proteins in 3D. *J Struct Biol* **211**, 107528.
- Thompson RF, Walker M, Siebert CA, *et al.* (2016) An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods* **100**, 3–15.
- Kilpatrick PK, Miller WG & Talmon Y (1985) Staining and drying-induced artifacts in electron microscopy of surfactant dispersions. II. Change in phase behavior produced by variation in pH modifiers, stain, and concentration. *J Colloid Interface Sci* **107**, 146–158.
- Kushida H (1962) A study of cellular swelling and shrinkage during fixation, dehydration and embedding in various standard media. *J Electron Microsc* **11**, 135–138.
- Mollenhauer HH (1993) Artifacts caused by dehydration and epoxy embedding in transmission electron microscopy. *Microsc Res Tech* **26**, 496–512.
- Stokes DJ & Hayles MF (2009) Methodologies for the preparation of soft materials using cryoFIB SEM, p. 73780G. <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.821834> (accessed January 16, 2023).
- Zhu Y, Sun D, Schertel A, *et al.* (2021) Serial cryoFIB/SEM reveals cytoarchitectural disruptions in Leigh syndrome patient cells. *Structure* **29**, 82–87.e3.
- Scher N, Rechav K, Paul-Gilloteaux P, *et al.* (2021) *In situ* fiducial markers for 3D correlative cryo-fluorescence and FIB-SEM imaging. *iScience* **24**, 102714.
- Schertel A, Snaidero N, Han H-M, *et al.* (2013) Cryo FIB-SEM: volume imaging of cellular ultrastructure in native frozen specimens. *J Struct Biol* **184**, 355–360.
- Mendonça L, Howe A, Gilchrist JB, *et al.* (2021) Correlative multi-scale cryo-imaging unveils SARS-CoV-2 assembly and egress. *Nat Commun* **12**, 4629.
- Wu G-H, Mitchell PG, Galaz-Montoya JG, *et al.* (2020) Multi-scale 3D cryo-correlative microscopy for vitrified cells. *Structure* **28**, 1231–1237.e3.
- Berger C, Dumoux M, Glen T, *et al.* (2022) Plasma FIB milling for the determination of structures in situ, p. 2022.08.01.502333. <https://www.biorxiv.org/content/10.1101/2022.08.01.502333v1> (accessed December 5, 2022).
- Pennington A, King ONF, Tun WM, *et al.* (2022) SuRVoS 2: accelerating annotation and segmentation for large volumetric bioimage workflows across modalities and scales. *Front Cell Dev Biol* **10**, 842342. <https://www.frontiersin.org/article/10.3389/fcell.2022.842342> (accessed April 1, 2022).
- Darrow MC, Luengo I, Basham M, *et al.* (2017) Volume segmentation and analysis of biological materials using SuRVoS (super-region volume segmentation) workbench *J Vis Exp* **126**, e56162.
- Sofroniew N, Lambert T, Evans K, *et al.* napari/napari: 0.4.14. <https://zenodo.org/record/3555620> (accessed February 15, 2022).
- Curcuraci L, Weinkamer R & Bertinetti L (2022) bmitools: biomaterials image processing tools. *J Open Source Softw* **7**, 4859.
- Pedregosa F, Varoquaux G, Gramfort A, *et al.* (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* **12**, 2825–2830.
- Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A, Kopf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S (2019) PyTorch: An Imperative Style, High-Performance Deep Learning Library In H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox and R. Garnett, ed. *Advances in Neural Information Processing Systems*, Curran Associates, Inc. [https://proceedings.neurips.cc/paper\\_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf) (accessed December 14, 2022).
- Haase R napari hub | plugin | devbio-napari by Robert Haase. <https://www.napari-hub.org/plugins/devbio-napari> (accessed November 2, 2022).
- Haase R & Savill R napari hub | plugin | napari-assistant by Robert Haase, Ryan Savill. <https://www.napari-hub.org/plugins/napari-assistant> (accessed December 5, 2022).
- Stephensen HJT, Darkner S & Spurring J (2020) Restoring drifted electron microscope volumes using synaptic vesicles at sub-pixel accuracy. *Commun Biol* **3**, 1–7.
- Yuan H, de Moortèle BV & Epicier T (2021) Accurate post-mortem alignment for focused ion beam and scanning electron microscopy (FIB-SEM) tomography. *Ultramicroscopy* **228**, 113265.
- Dahmen T, Engstler M, Pauly C, *et al.* (2016) Feature adaptive sampling for scanning electron microscopy. *Sci Rep* **6**, 25350.



26. Holzer L, Indutnyi F, Gasser PH, *et al.* (2004) Three-dimensional analysis of porous BaTiO<sub>3</sub> ceramics using FIB nanotomography *J Microsc* **216**, 84–95.
27. Schaffner M, Wagner J, Schaffner B, *et al.* (2007) Automated three-dimensional X-ray analysis using a dual-beam FIB. *Ultramicroscopy* **107**, 587–597.
28. Amira Software | Life Science Research – UK. <https://www.thermofisher.com/uk/en/home/electron-microscopy/products/software-em-3d-vis/amira-software.html> (accessed November 1, 2022).
29. Linear Stack Alignment with SIFT <https://imagej.github.io/plugins/linear-stack-alignment-with-sift> (accessed December 13, 2022).
30. Cardona A, Saalfeld S, Schindelin J, *et al.* (2012) TrakEM2 software for neural circuit reconstruction. *PLoS One* **7**, e38011.
31. Hennies J, Lleti JMS, Schieber NL, *et al.* (2020) AMST: alignment to median smoothed template for focused ion beam scanning electron microscopy image stacks *Sci Rep* **10**, 2004.
32. Vourvoulakis J, Kalomiros J & Lygouras J (2017) FPGA accelerator for real-time SIFT matching with RANSAC support *Microprocess Microsyst* **49**, 105–116.
33. Robinson K & Whelan PF (2004) Efficient morphological reconstruction: a downhill filter. *Pattern Recognit Lett* **25**, 1759–1767.
34. Perdigão L Rosalindfranklininstitute/chafer <https://github.com/rosalindfranklininstitute/chafer> (accessed December 8, 2022).
35. Ronneberger O, Fischer P & Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234–241. Cham: Springer.
36. Kremer JR, Mastrorade DN & McIntosh JR (1996) Computer visualization of three-dimensional image data using IMOD *J Struct Biol* **116**, 71–76.
37. Penczek PA (2010) Resolution measures in molecular electron microscopy. *Methods Enzymol* **482**, 73–100. <https://linkinghub.elsevier.com/retrieve/pii/S0076687910820038> (accessed October 25, 2022).
38. Saxton WO & Baumeister W (1982) The correlation averaging of a regularly arranged bacterial cell envelope protein. *J Microsc* **127**, 127–138.
39. Rosenthal PB & Henderson R (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* **333**, 721–745.
40. van Heel M & Schatz M (2005) Fourier shell correlation threshold criteria. *J Struct Biol* **151**, 250–262.
41. Koho S, Tortarolo G, Castello M, *et al.* (2019) Fourier ring correlation simplifies image restoration in fluorescence microscopy. *Nat Commun* **10**, 3103.
42. Ho EML, quoll. <https://github.com/rosalindfranklininstitute/quoll> (accessed December 8, 2022).
43. Schindelin J, Arganda-Carreras I, Frise E, *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682.
44. Lin DH & Hoelz A (2019) The structure of the nuclear pore complex (an update) *Annu Rev Biochem* **88**, 725–783.
45. Alieva IB & Uzbekov RE (2016) Where are the limits of the centrosome? *BioArchitecture* **6**, 47–52.
46. Kruskal WH & Wallis WA (1952) Use of ranks in one-criterion variance analysis. *J Am Stat Assoc* **47**, 583–621.
47. scipy.stats.kruskal — SciPy v1.9.3 Manual <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kruskal.html> (accessed November 1, 2022).

**Cite this article:** Perdigão L. M. A, Ho E. M. L, Cheng Z. C, Yee N. B. -Y, Glen T, Wu L, Grange M, Dumoux M, Basham M and Darrow M. C (2023). Okapi-EM: A napari plugin for processing and analyzing cryogenic serial focused ion beam/scanning electron microscopy images. *Biological Imaging*, 3: e9. doi:<https://doi.org/10.1017/S2633903X23000119>