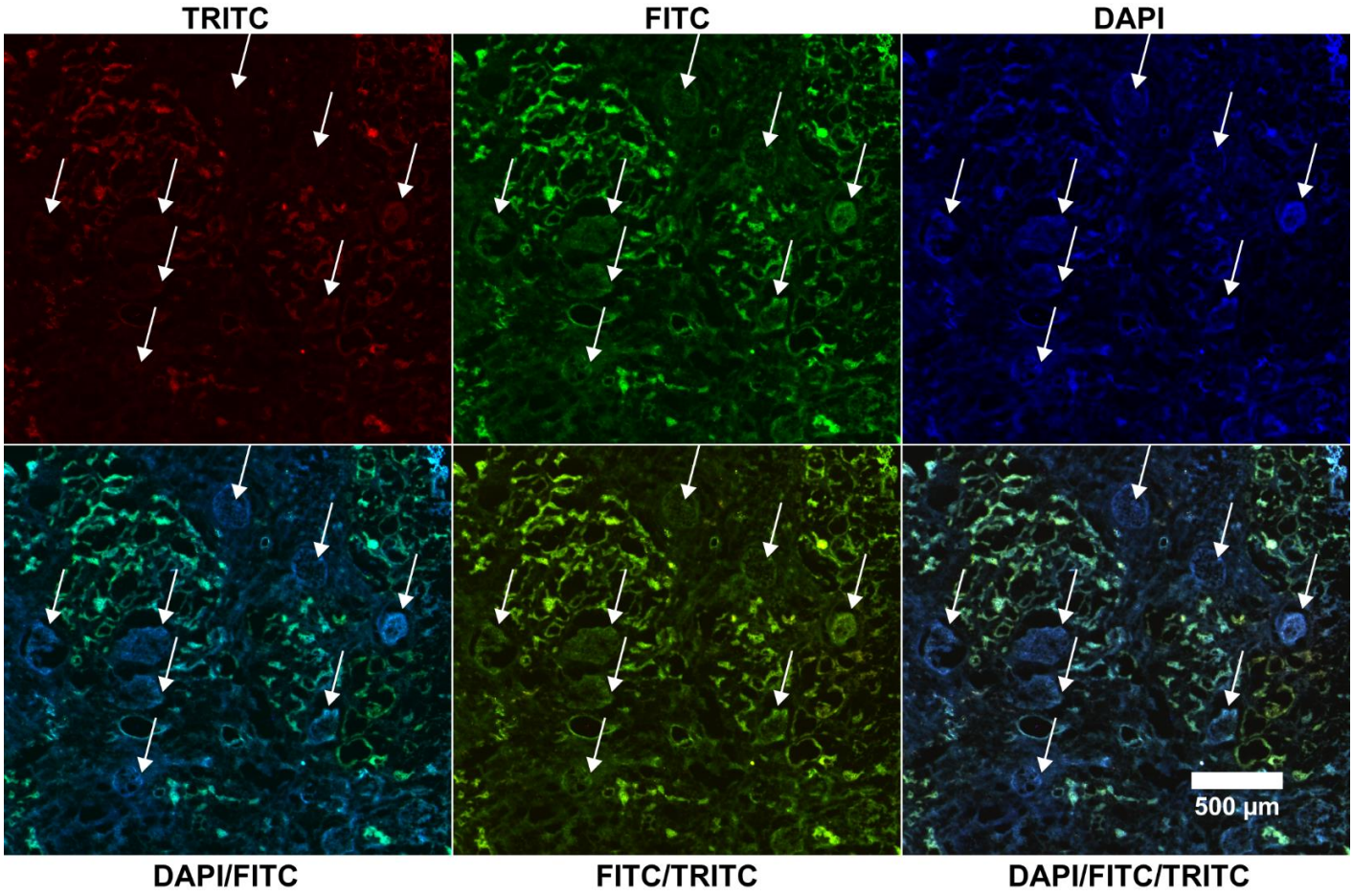
Supplemental Methods

Autofluorescence microscopy for histology-directed workflows

AF microscopy parameters are determined by the experimental workflow. In registration-driven experiments, it suffices to have a single channel AF image for the serial sections to drive registration. The registration software in this workflow only uses a single image channel, affording time saving by a shorter microscopy scan. Generally this single channel is composed from multiple AF channels using the pixel-wise maximum image. For scanning with only one channel for registration, the channel should be selected to be maximally informative of the tissue morphology for registration to work in all areas. An AF signal that is highly localized in the tissue section would not be ideal for global section registration. In our experience, a standard FITC filter has proved most useful for this task as it provides high intensity AF from several endogenous fluorophores but this can be quickly evaluated on a case-by-case scenario. The FITC filter described has a wide band for both excitation (465-495 nm) and emission (515-555 nm). For



Supplemental Methods Figure 1. Comparison of different AF channels for annotation of glomeruli. White arrows indicate most of the glomeruli in the window.

direct AF annotation workflows, generally a longer microscopy acquisition is required, but again, this depends on the annotator and the target ROIs, and the microscope.

For human kidney glomeruli, DAPI, FITC and TRITC channels were used (bottom right panel in Supplemental Methods Figure 1), but examination of different color combinations in Sup. Met. Figure 1 allows us to see DAPI/FITC may suffice alone and that the TRITC channel adds limited detail.

Intensity based computational image registration for AF images

The basics of image registration along with the software and parameters used to achieve image registration in throughout this work are described in this section.

The *elastix* registration toolbox (http://elastix.isi.uu.nl/) was used for all registration. *elastix* is a general purpose medical image registration toolbox built on top of the Insight Segmentation and Registration Toolkit (ITK) that registers images using image intensity patterns. Recently, the *elastix* team has developed *SimpleElastix* (https://simpleelastix.github.io/), a complete wrapping of the *elastix* and ITK components in many popular scripting languages such as Python and R. Within the software built for this work, python was used to manage registration inputs and outputs for *SimpleElastix*.

Fundamentals of Image registration

Image registration consists of finding a transformation of the coordinates of one image to closely match the coordinates of a second based on shared information (image intensity in the work, although other models can use shape or other information). The image to be deformed is called the *source image* I_S (sometimes also moving image) while the image to which I_S is being deformed is called the *target image* I_T (sometimes also called fixed or template image). Thus in computational registration one searches for a coordinate transform $\boldsymbol{T}$ which we can apply to I_T such that $I_S(\boldsymbol{T})$ is spatially aligned with I_T. Within *elastix,* the registration task is achieved through iterative optimization of the transformation parameters based on a cost function C which measures image similarity at each iteration. In this way, the registration progressively moves towards improved alignment at each iteration. In this work, microscopy image registration is also performed over multiple image resolutions. Details for the key registration components follows.

The cost function C which drives the iterative optimization of registration parameters consists of a similarity measure between images. In this work mutual information[34] was primarily used. Mutual information is a statistical measure of intensity similarity between the $I_S(\boldsymbol{T})$ and I_T. This measure is general and has had a large impact in multimodal image registration, although we found it to work equally well in mono-modal registration. We also found more mono-modal metrics like mean squares and correlation coefficient worked in most cases, but often did not always converge in the critical early low resolution image registration.

Linear transformations were done with either a rigid body or an affine model. Rigid body transformations allow for rotation and translation but no scaling or shearing. Rigid transformations are useful to obtain a 'safe' transformation without shearing or scaling when image resolution is precisely known, aspect ratio does not change, and gross morphology is not substantially changed, as expected in serial tissue sections. Affine transformations allow deformations with more degrees of freedom: scaling, shearing, rotation and translation. Affine transformations are used when images were taken with two different sensors with different image resolutions but this is not strictly necessary as registration in

elastix occurs in physical space (microns) rather than pixel space. Linear transformation models are practically required in serial section registration to initialize non-linear transformations. Here, these transformations 'roughly' align the tissue sections prior to local warping. As the localized warping is a much more complex model, it can be more error prone, especially when tasked with the major warps that an initial linear transformation would find. To put simply, accurate registration is not likely to be achieved if large warps are necessary to account things like gross morphological alignment.

The non-linear transformations described in this work used exclusively a B-spline transformation model. This model uses an evenly spaced control point grid between which B-spline curves are fit. The density of control points on image is user specified in physical spacing. We found 100 μm to be sufficient for serial sections scanned at ~ 1 μm resolution. A higher number of control points requires a greater number of parameters to be optimized and we found that going too fine often resulted in unrealistic transformations such as swirls. These would not be well estimated with the linear transformation, whereas non-linear transformations can be used for small corrections from these sample preparations.

For optimization, stochastic gradient descent (SGD) was employed. SGD is a variant of GD that uses smaller sample batches (random samples of pixels from the images) when calculating the gradient to save computational time. In image registration, smaller batches equate to sampling a number of pixels in the image selected randomly rather than the entire image for calculation of the cost function (image similarity metric). *Klein et al.* proved that robust registration could still be achieved with this approach.[35] This is critical when registering high resolution microscopy data where there m x n pixels and each image has thousands of pixels in each dimension. The SGD optimizer doesn't stop after reaching a minimum but will continue for the set number of iterations. In our work we have found 500-1000 iterations necessary for linear transformations and >1500 and <2500 necessary for nonlinear transformations. This fits with the degrees of freedom of each model. In general, we tried to overshoot the number of iterations as the computation time for a given registration is less than 5 minutes unless the dimensions of the image are very large.

Another critical parameter is the hierarchical resolution scheme used in the registration. In general, we used 10 resolutions, starting with optimization of registration of very down sampled images and progressively increasing in dimension until the last layer is the original image dimension. In linear transformations, this simply means downsampling the images before searching for a transformation, while in non-linear transformations this entails downsampling the image and downsampling the B-spline control point grid. The latter point is especially important, as the downsampled image is better for estimating large warps to begin the registration, then smaller features are progressively registered as the process goes to higher resolution images. In a sense the registration goes from region localized warping to cellular level warping. This allows large then small local features to be accurately registered. Another key element in the multiresolution scheme is optimizer step size (learning rate) which determines the 'nudge' each parameter can receive at each iteration and is set per resolution in physical amounts. In this we have also selected a large to small regime, where early low resolution transformations take larger steps (100s of microns), (i.e. large translations and rotations for rough alignment) and high resolution makes smaller changes (10s of microns) to the transformation parameters assuming earlier models converged well. In linear models this allows us to first find the gross

translations and rotations, for instance where one tissue section occupies only the top-left quadrant of the image, and the second section's image is centered in the image.

Tweaking of parameters is necessary if registration is poor, troubleshooting advice can be found in the *elastix* manual at http://elastix.isi.uu.nl/. Parameter files used in all registrations are available at https://github.com/NHPatterson/regToolboxMSRC/parameter_files.

Supporting Information and Figures

Autofluorescence as an IMS compatible microscopy technique

The lack of sample preparation for AF microscopy makes it ideally suited for integration into IMS experiments where it can yield a microscopy-grade view into the structure of the tissue without introducing non-biological variation. The AF modality is general in that it can measure both fresh frozen and formalin-fixed paraffin-embedded (FFPE) tissue sections and has been used to interrogate numerous biological phenomena.[36] Additionally, AF images can also be acquired from tissues on opaque

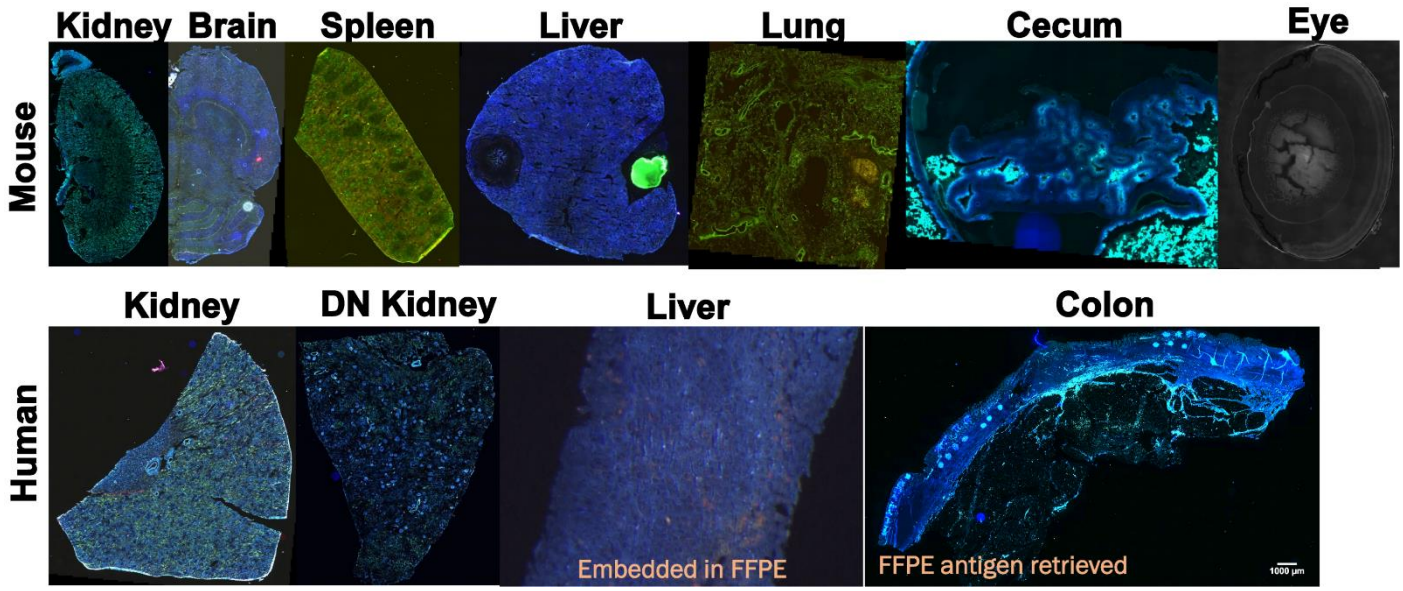


Figure S1. Autofluorescence microscopy images of sections from various tissue types. Thumbnail size AF images of various tissue types. Samples were all fresh frozen, except where noted as FFPE. Note: these images are not matched in terms of scale and are not intended for comparison. Blue = DAPI filter, Green = FITC filter, Red = TRITC filter. See manuscript methods for information on the wavelengths of these fluorescence filters.

surfaces (*e.g.*, metallic slides) if an upright microscope is used, further expanding the application domains in which microscopic histological information can be obtained. AF microscopy of tissue sections has presented rich morphology in all tissue types that we have examined to date (several examples are shown in Figures S1-4). A first critical aspect to determine feasibility of any workflow incorporating AF is to measure whether AF influences or changes analyte composition in the sample. This aspect was examined for lipid analysis specifically by performing IMS measurement on three serial sections of mouse brain, two after AF microscopy and one without AF. The data were subjected to principal components analysis (PCA) to examine underlying correlative trends in the data for AF derived artifacts.

In Figure S5, the first 6 PCs account for ~97% of the data variance are shown for the dataset. They show remarkable similarity across all datasets, suggesting very little to no impact from AF microscopy on the IMS measurements.

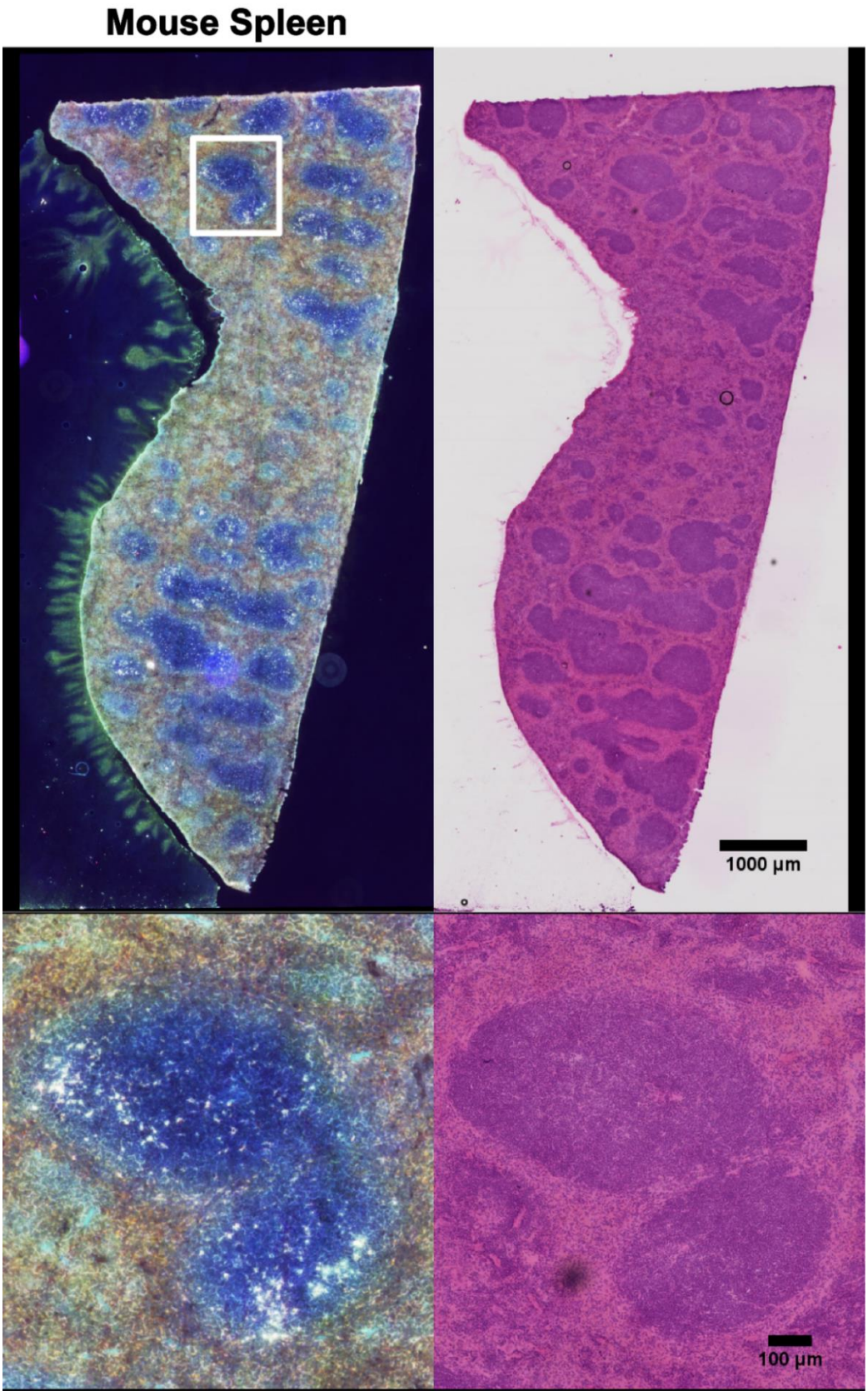


Figure S2. Mouse Spleen registered AF and H&E images. High magnification shows an area of white pulp.

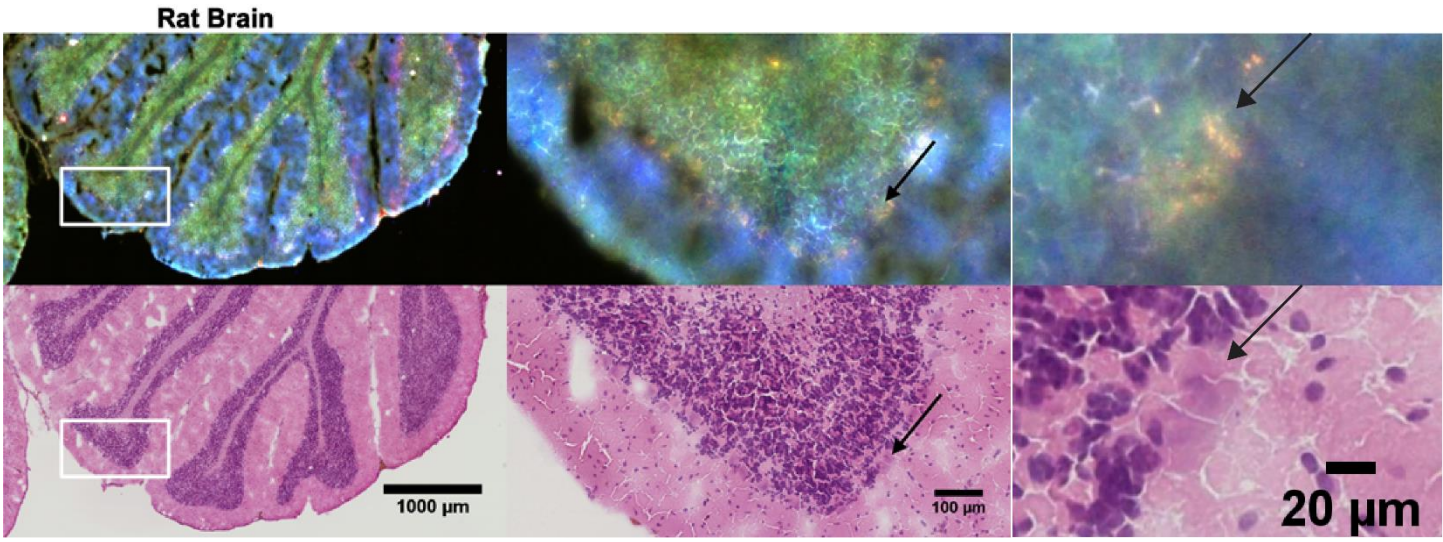


Figure S3. Rat brain registered AF and H&E. Mid magnification shows an area of cerebellum, arrows indicate a Purkinje cell. High magnification shows a further zoom on the Purkinje cell body.

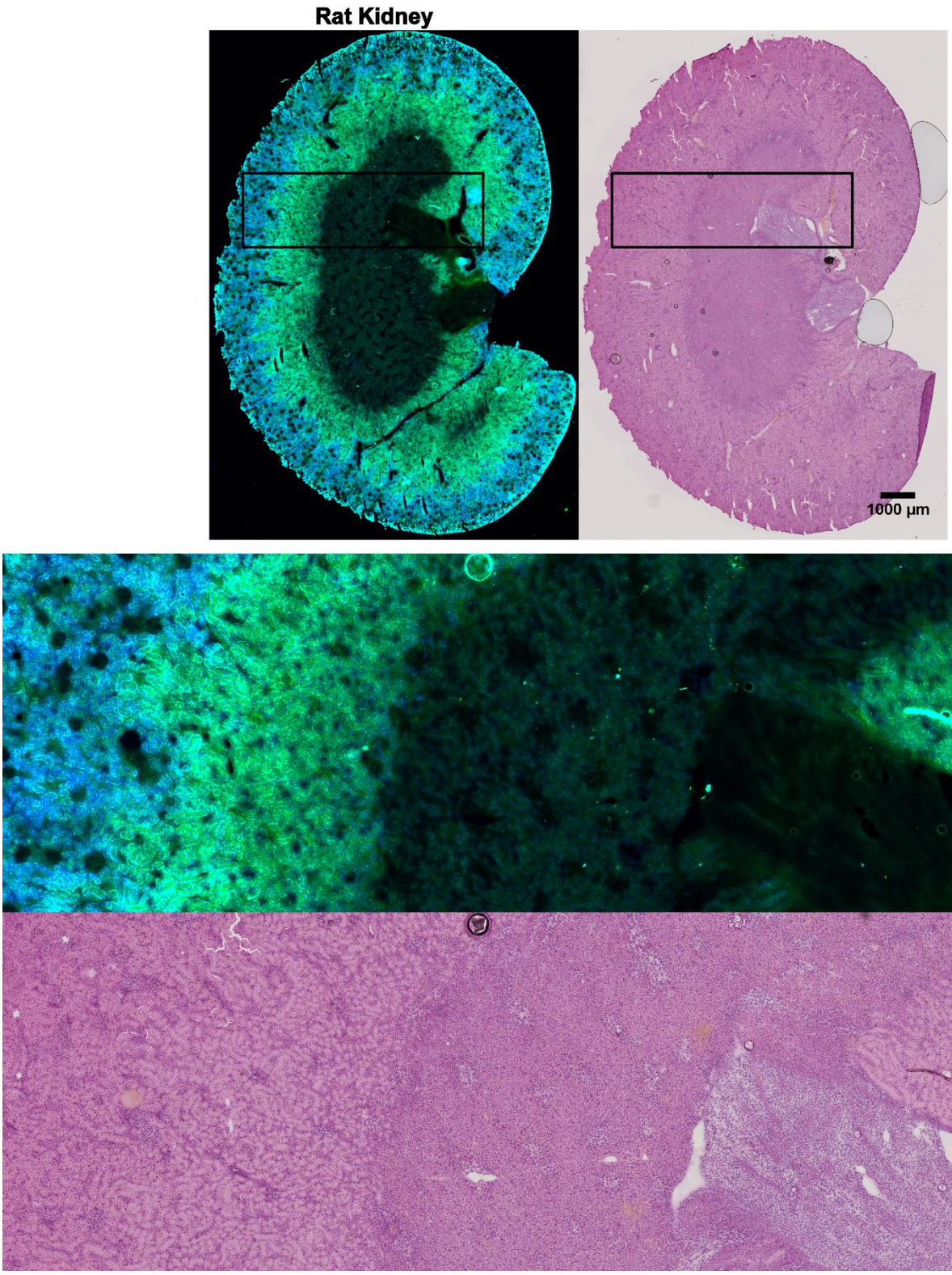


Figure S4. Rat brain registered AF and H&E. High magnification shows the transition from cortex to medulla to pelvis (left to right).

1

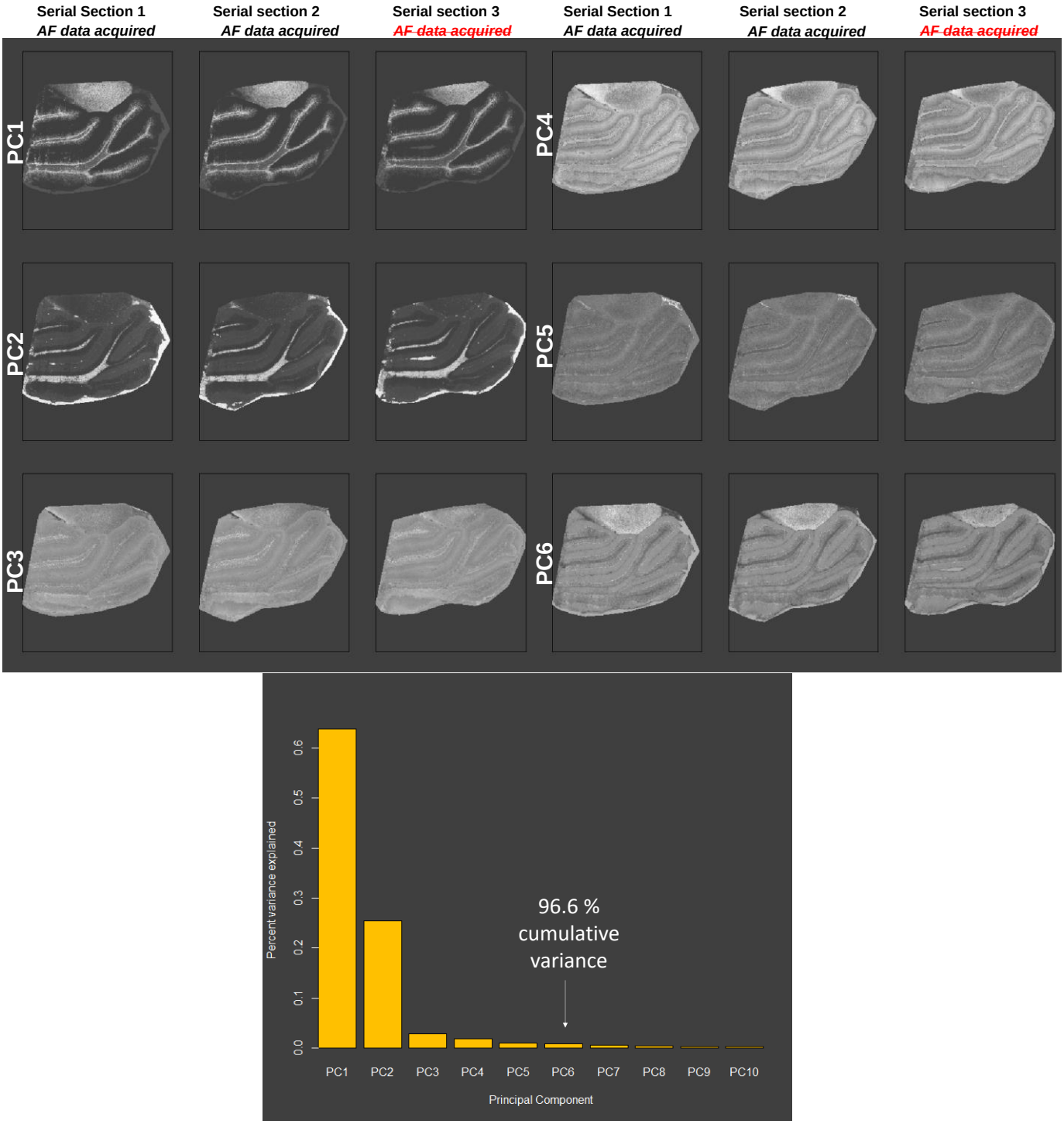


Figure S5. Evaluation of the effect of AF on mouse brain cerebellum tissue using DAN matrix and PCA.

Registration-driven histology-directed IMS workflow and time considerations

Figure S6 outlines the workflow in a simplistic flowchart with time highlights with the AF registration-driven workflow's unique steps highlighted with black boxes and the common steps to all histology directed workflows in red boxes.

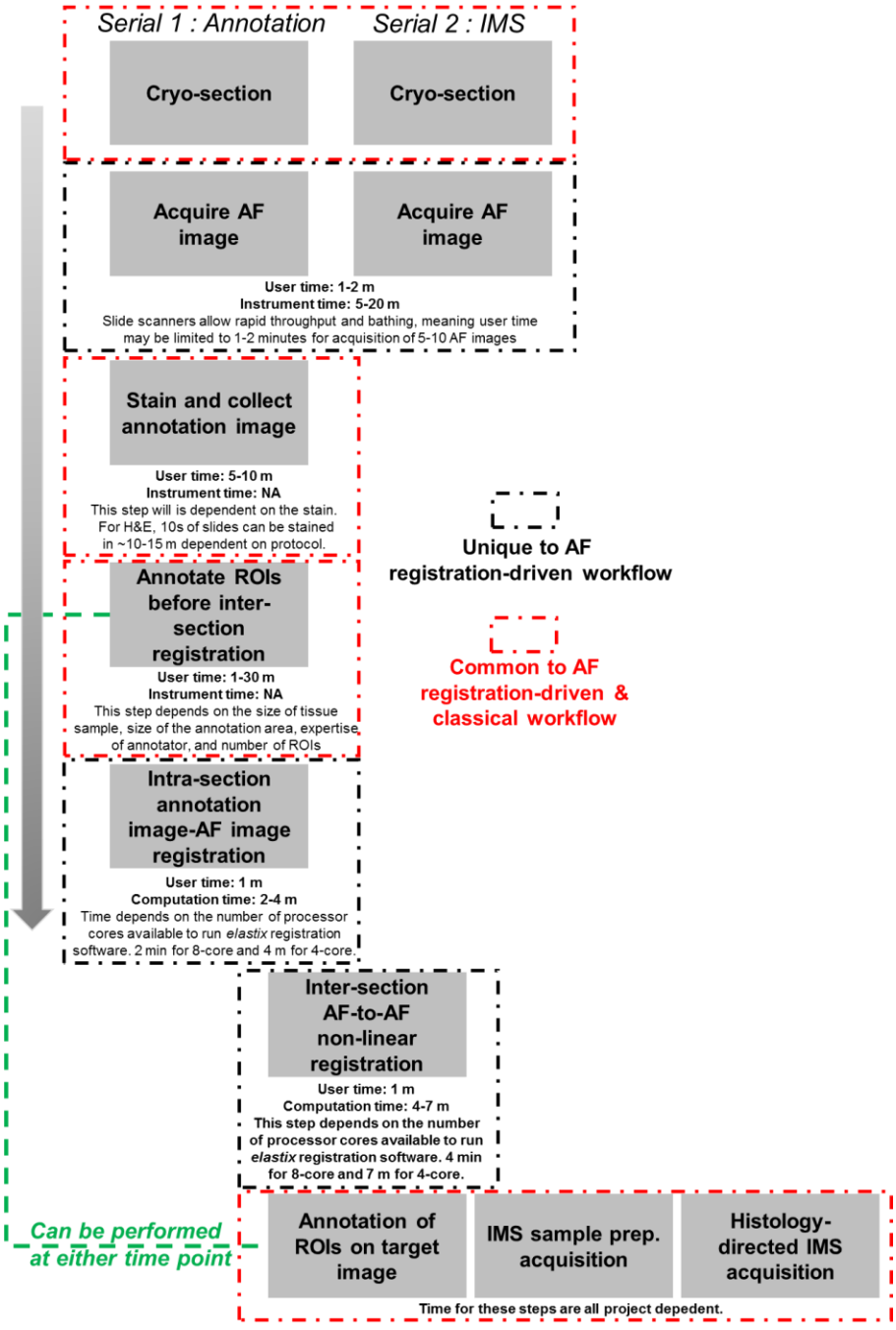


Figure S6. AF registration-driven histology-directed flowchart with time considerations. Black boxes highlight steps unique to the novel workflow while red boxes highly common steps between the new workflow and previous ones.

Figure S6 highlights the commonalities between workflow and where the AF registration-driven approach differs. The most important time factor is the AF microscopy which may add differing amounts of time dependent on the microscope and magnification used. In all experiments in this work, a 10x magnification objective was used (~1 µm / pixel). A single fluorescence channel rat kidney (~15x15 mm) AF scan required 10 minutes on the Nikon Eclipse 90i microscope. On a more modern fluorescence slide scanner, a 10x acquisition is likely to be 5-10x faster at the same magnification. Adding additional filters (not necessary for registration only experiments, see Supplement Methods), adds time linearly. Overall, the AF microscope and additional registration steps are estimated to add 5-10 minutes of user time (time where a user is interacting with a software, setting up the AF scans, etc.) and 20-30 minutes total per serial section pair (user + instrument + computational time) at 10x magnification with the system used in this work with the caveat that faster microscopes will reduce this time significantly. Lower magnification objectives could also be explored as even a 5x objective produces pixels of approximately 2 µm, allowing fine registration to be achieved in approximately a quarter of the time required by a 10x. Comparatively, manual serial registration approaches using microscope images or a mixture of high and low resolution images can take several hours or minutes, depending on desired accuracy and the user. Computational serial section registration approaches[6,11,29] using high and low resolution images may be faster at the image scanning phase (a flat-bed scan takes minutes for multiple sections) but have limited accuracy and limited ability to evaluate accuracy as described in the next section.

1 **Evaluation of autofluorescence image serial section registration using test**

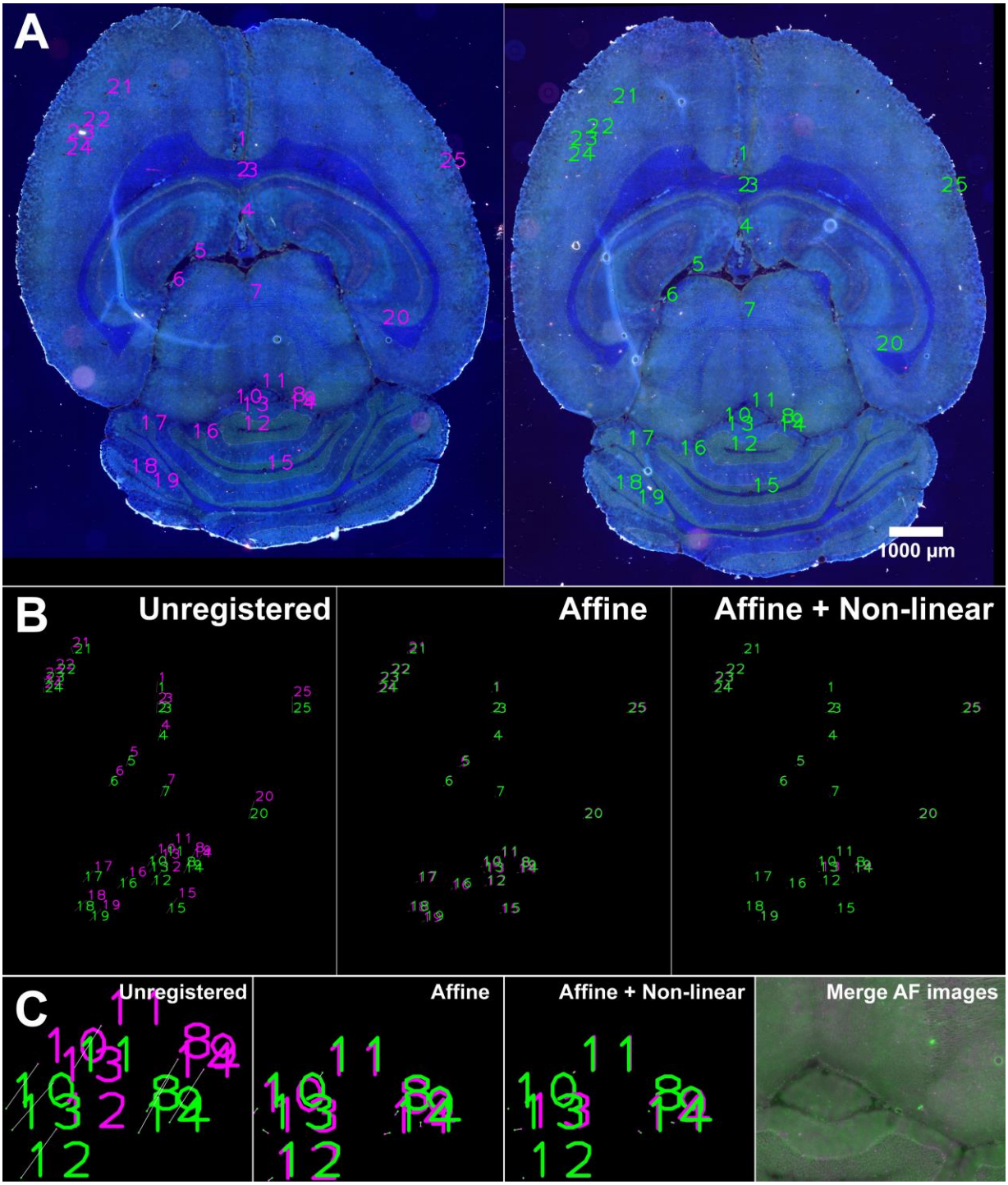


Figure S7. Landmark distance evaluation mouse brain images. **A.** Autofluorescence images of the two mouse brain sections pre-registration. Numbers indicate corresponding points. **B.** Point sets displayed with different registrations at low magnification. **C.** Point sets at high magnification in first 3 panels, left to right, then a merged AF image where the registered image is purple and the target image is green. Lines are between the points and were used to calculate distance.

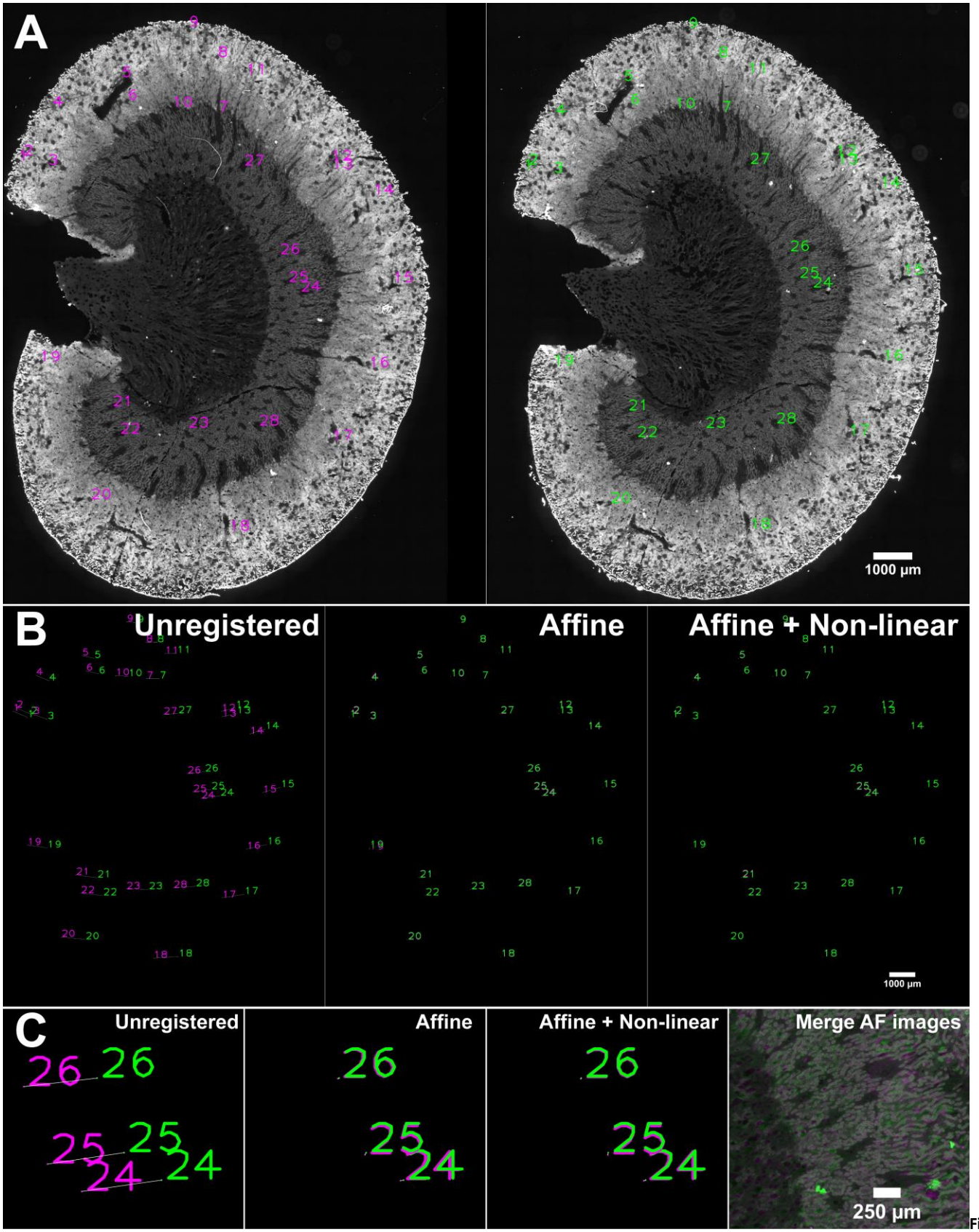


Figure S8. Landmark distance evaluation rat kidney images. **A.** Autofluorescence images of the two rat kidney sections pre-registration. Numbers indicate corresponding points. **B.** Point sets displayed with different registrations at low magnification. **C.** Point sets at high magnification in first 3 panels, left to right, then a merged AF image where the registered image is purple and the target image is green. Lines are between the points and were used to calculate distance.

1

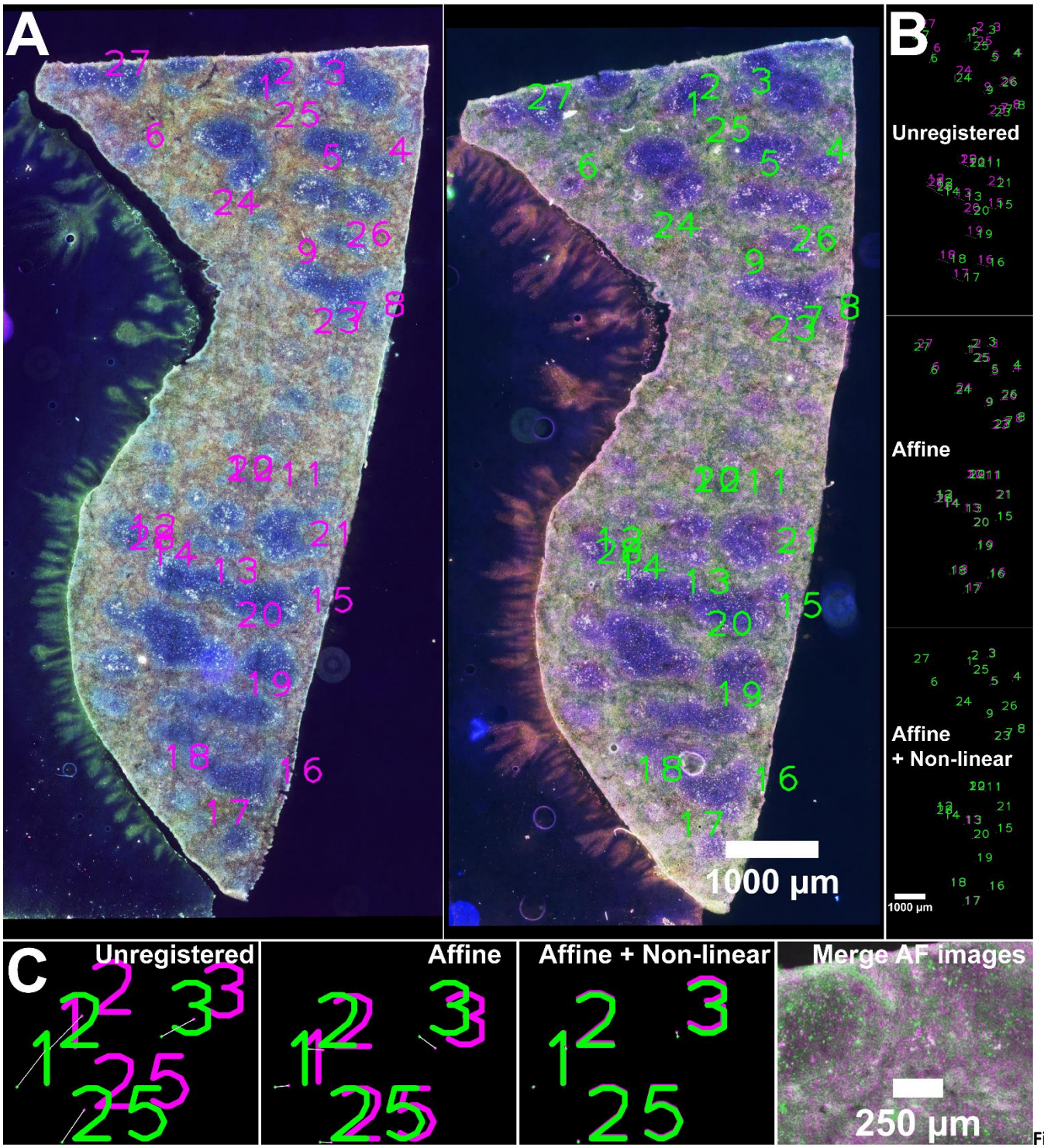


Figure S9. Landmark distance evaluation mouse spleen images. **A.** Autofluorescence images of the two mouse spleen sections pre-registration. Numbers indicate corresponding points. **B.** Point sets displayed with different registrations at low magnification. **C.** Point sets at high magnification in first 3 panels, left to right, then a merged AF image where the registered image is purple and the target image is green. Lines are between the points and were used to calculate distance.

On the difficulty of comparing the novel method to previously published approaches

It is important to note the difficulty confronted in comparing our novel approach with previous histology-directed approaches[6,11,29] as such fine (< 10 µm) comparisons are impossible when working in a histology-directed platform mixing low and high-resolution images. The low-resolution images are not resolved to the level of detail to select cellular and micron level morphological features as corresponding landmarks. Previous works using this approach have had success targeting large histological structures but was not presented as targeting cellular-level foci. Further, the mean distance error of the novel method's non-linear registration was ~10 µm across multiple tests and most flat-bed scanners images achieve maximum resolving of > 10 µm, meaning comparison at these levels is not technically possible. Figure S10 demonstrates this difficult using a rat kidney H&E image scanned with a high-resolution slide scanner and a low resolution flatbed scan where the resolving power of the scanner is estimated to be between 15-30 µm, already above the mean distance error of our AF method. Besides poor resolving power, this registration is also multi-modal between a stained whole slide digital image and flatbed scan image of an unstained section, which reduces certainty.

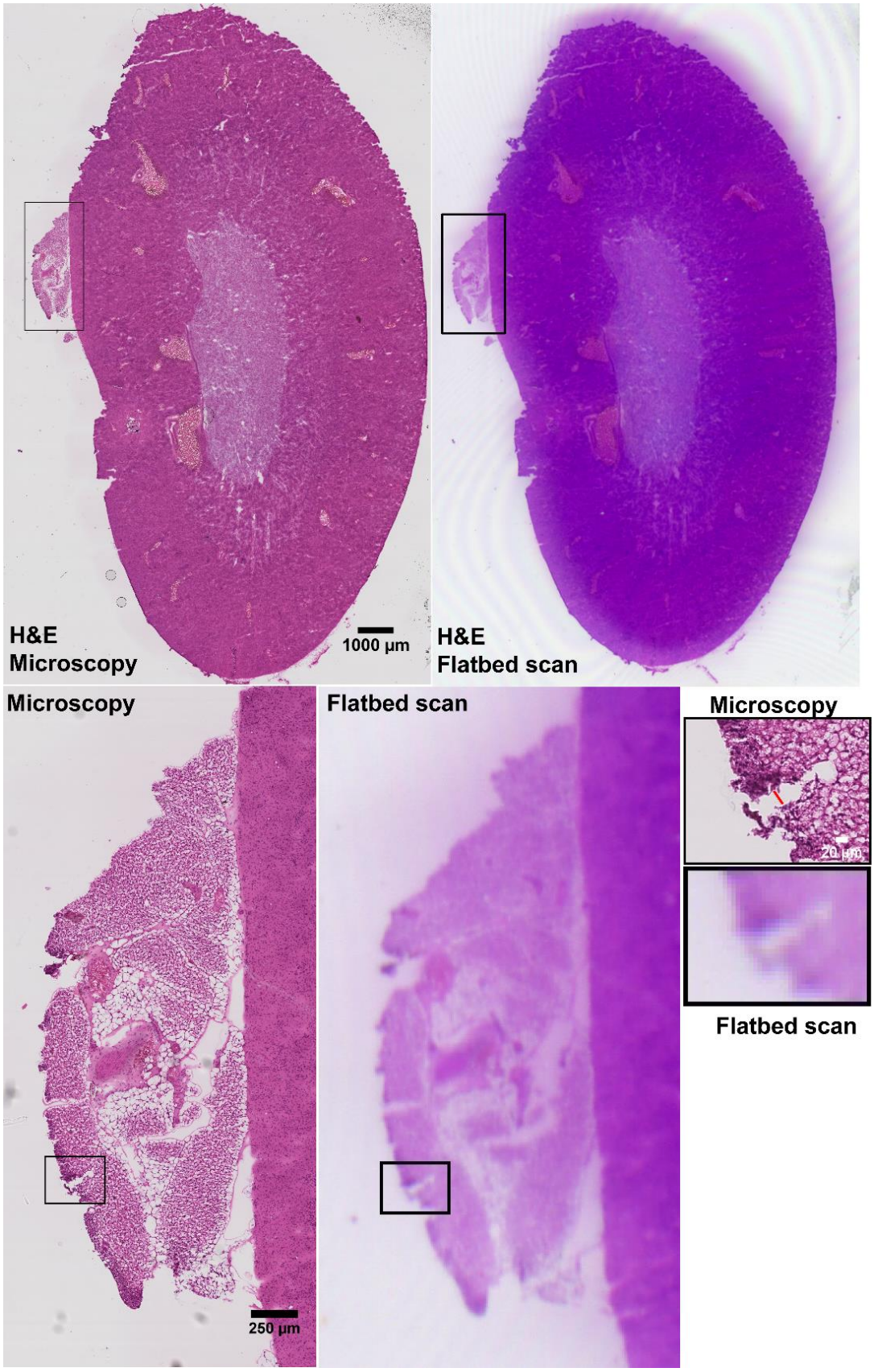


Figure S10. Comparison of microscopy to flatbed scan of same tissue section of a rat kidney. The red bar in the highest resolution microscopy panel is 30 µm. We qualitatively assess the true resolving power of the flatbed scan to be somewhere around 15-30 µm, below the mean distance we calculated for corresponding landmarks across multiple tissue types highlighting the difficulty of assessing registration at the level our novel approach achieves.

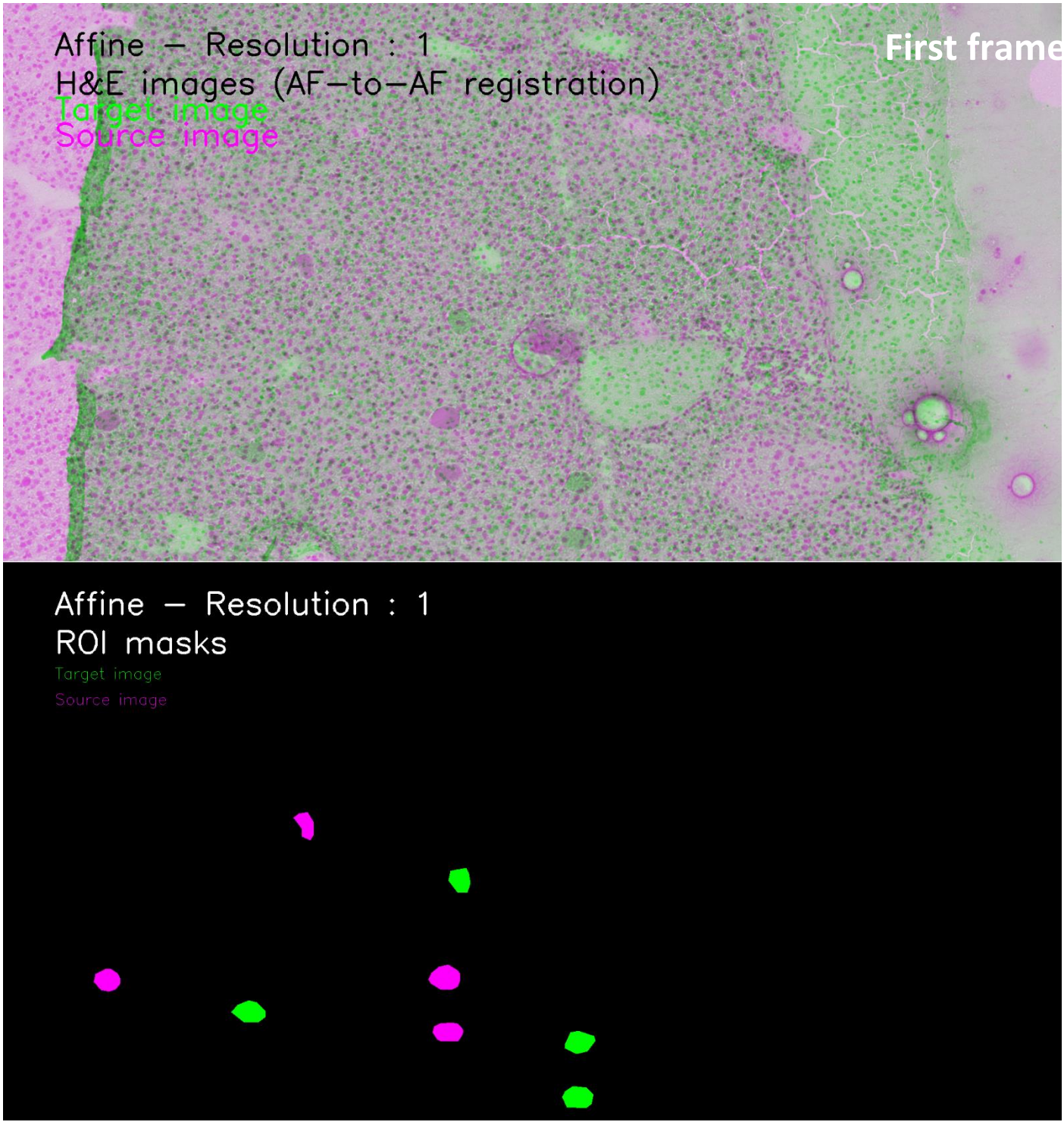


Figure S11. First and final video frames of Video S2 showing registration of the malaria infected mouse liver. The green mask are infected hepatocytes from the target image and the magenta are from the source (warped) image. Resolutions refer to the multi-resolution registration scheme where registration starts on a down-sampled image and the final registration ends on the full resolution image.

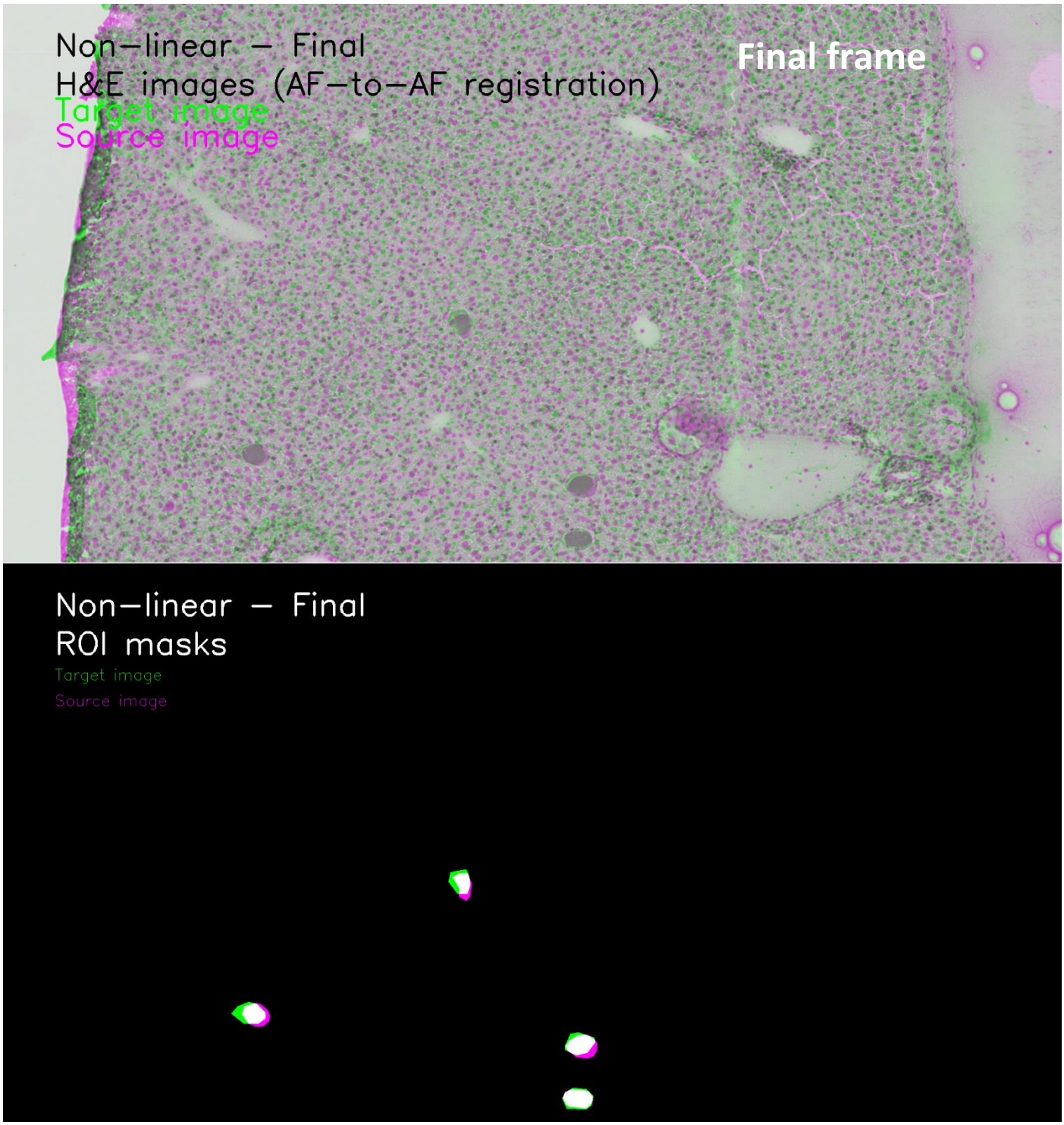


Figure S11 (cont.).

Grading accuracy of the annotation overlap of the infected mouse hepatocytes

In order to grade the overlap between the annotation section and IMS section for all ROIs numerically presented in Figure 4, Dice similarity coefficients (DSCs) [37] were calculated post-registration. Equation 1 shows the mathematical definition of the DSC

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (1)$$

where in this context TP is a true positive (overlapping pixel), and FP is a false positive, a target pixel that does not overlap with a registered source pixel, and FN is a false negative or a registered source pixel that does not overlap with a target pixel. A DSC of 1 indicates perfect overlap between the ROI masks. Thus, the mean DSC of 0.725 of the infected hepatocytes in section displayed in Figure 3 indicates there was a mean 72.5% overlap across all 14 annotations. Figure S12 shows images of the overlap, together with corresponding DSCs. Of the twelve annotated ROIs on the staining, the minimum DSC value was 0.57 indicating that even in the worst-case ROI, infected hepatocytes, which are individual cellular structures, still had >50% overlap after AF-based registration, allowing for reliable targeting post-registration. The compared structures (infected hepatocytes) change in size and shape from one section and secondarily, these values are being calculated on pixel areas with an average area of 1.7 mm^2 and average widths and heights of 43 and 40 µm, respectively. This latter point differs from many applications of DSC where whole organs are annotated in medical images and compose a large portion of the image.

To further supplemental the infected hepatocyte targeting analysis for the IMS section, serial sections from 3 biological replicates were made in triplicate (3 technical replicates) and a microscopy only accuracy evaluation was performed. Figure S13 outlines the competitive registration evaluation workflow. In this assay, three microscopy modality images were made per serial section pair for competitive evaluation. First, AF microscopy was performed, followed by unstained brightfield microscopy (BF), and finally, the sections were stained by H&E and a microscopy image was obtained. Unstained BF images were obtained to make a comparison of histology-directed targeting accuracy of AF with another simple, non-destructive microscopy modality and H&E was taken as a gold standard and for annotation of the infected hepatocyte ROIs.

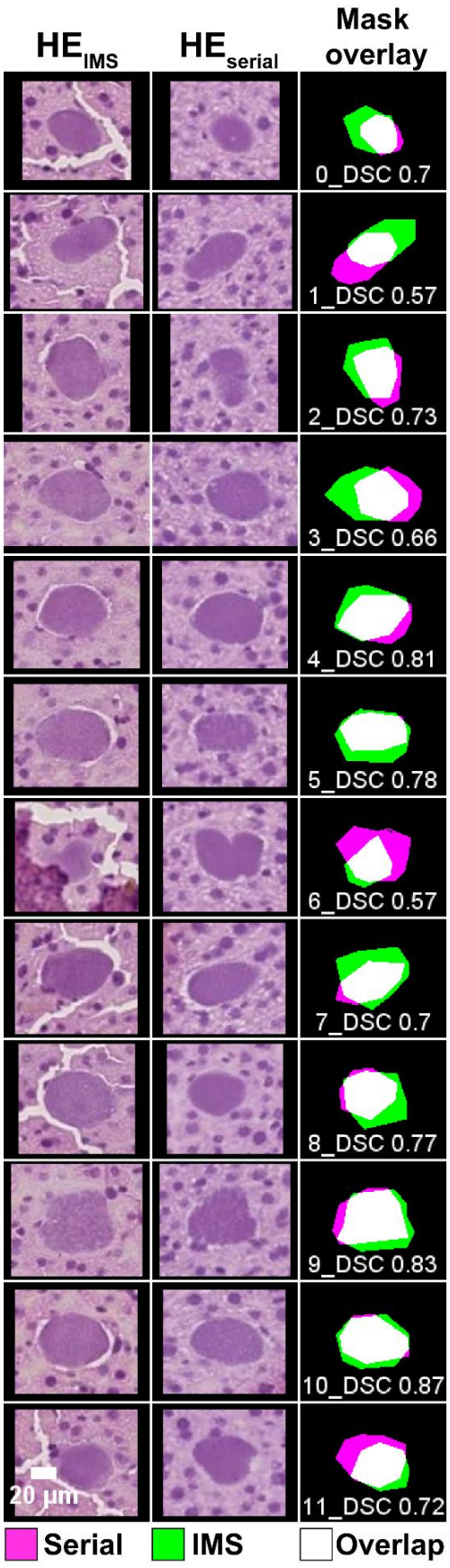


Figure S12. Dice-Sorenson coefficient for overlap of infected hepatocytes for 12 malarial lesions after AF-to-AF registration. HE_{IMS} = H&E stain of the section where IMS was performed after IMS acquisition. HE_{serial} = H&E stain of the section that was annotated for IMS data acquisition. In this case, HE_{serial} is shown registered to HE_{IMS} by AF-to-AF. DICE mean = 0.725.

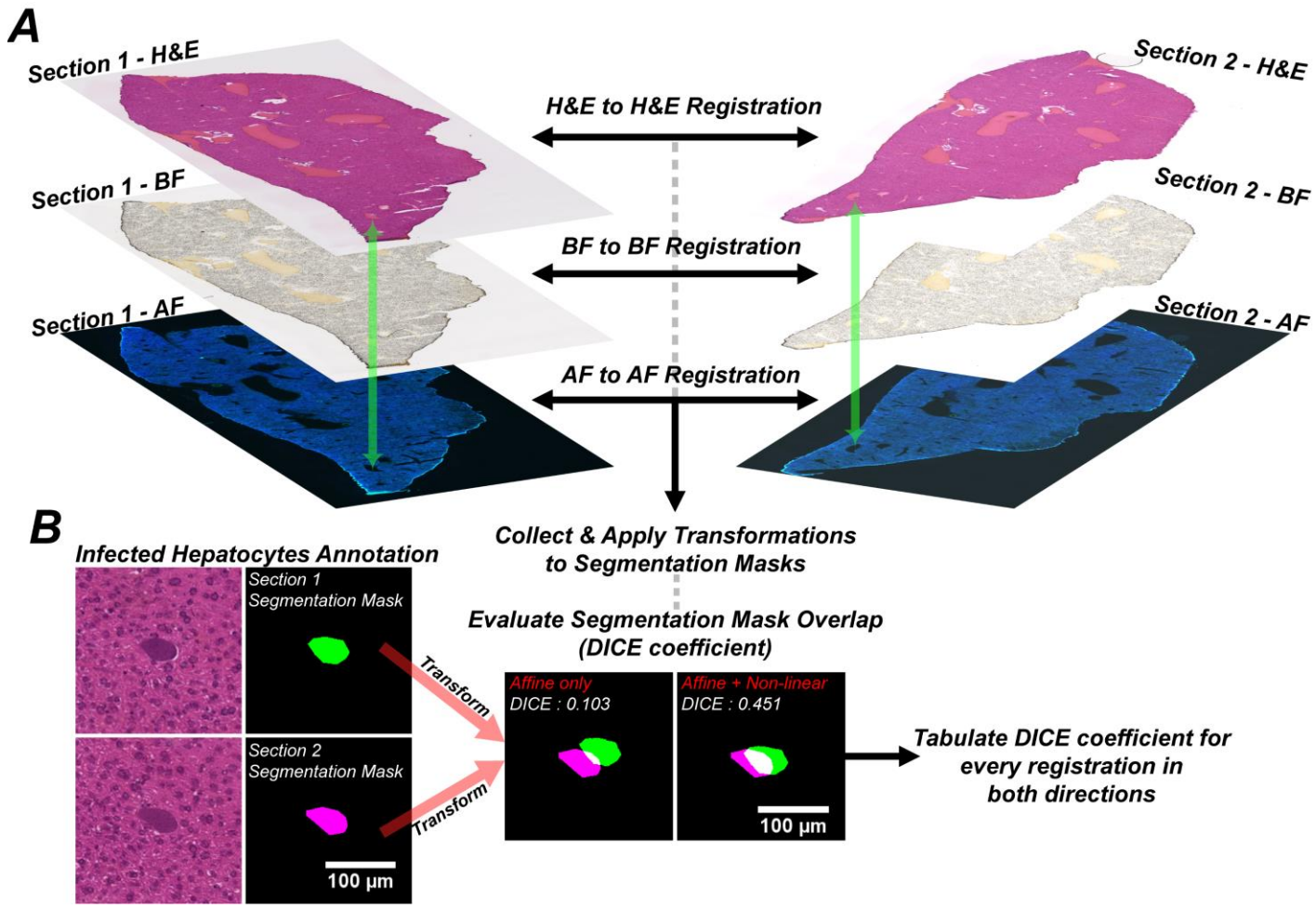


Figure S13. Schematic of the workflow for competitive evaluation of registration accuracy of infected hepatocytes between serial tissue sections using different modalities. A. This workflow shows the idea of intra-section and inter-section registration. In each vertical image stack of modalities, we see the intra-section registration of modalities (highlighted by the green arrow), then between inter-section between the labelled Section 1 and Section 2 images (with the horizontal black arrow). B. Evaluation is done by making polygonal ROIs on the intra-section registered H&E and then applying the inter-section registration to the masks, in both directions, and computed the resulting overlapping with a DICE coefficient. This is tabulated for every possible pair (n = 9) in both directions (Serial Section 1 → Serial Section 2, Serial Section 2 → Serial Section 1).

Figure S13-A shows representative data from each modality and the overall registration scheme. Initially, H&E and BF microscopy images from one section (intra-section registration) are aligned to the AF of the same section using a linear affine model resulting in a stack of AF, BF, and H&E in the same coordinate space (indicated by the transparent green arrow in the Figure S13-A). This step allows for all subsequent inter-section registration to start from the same point. Appendix I shows the intra-section registered images of 5 randomly selected patches from 3 of the 18 tissue sections (1 from each biological replicate). After intra-section registration, inter-section registration (between serial sections) is performed using an affine then non-linear transformation model for each serial section pair. In each registration, the same optimized parameters used elsewhere are employed. Since there are two serial sections, two registrations are performed from the first section to the second section, and vice versa.

Figure S13-B highlights the evaluation using the DICE metric of infected hepatocyte segmentation masks. First, annotation of the infected hepatocytes is done using polygonal ROIs on the H&E images registered to AF (Figure S13-B, annotation). In order to only compare registration accuracy and not ROI presence and absence, only infected hepatocytes visible in each section were annotated. After annotation of infected hepatocytes (mean number of ROIs per section = 22.7 ± 4.9), binary masks of the annotations were generated for each tissue section, shown in the green and magenta in Figure S13-B. As the masks were made in the intra-section coordinate space, they could be transformed using the previously developed inter-section registrations. After transformation of the masks using a nearest neighbor interpolation (to prevent interpolation error from expanding the segmentation dimensions), the DICE of the masks was calculated for the entirety of the section. Table S1 shows the raw tabulated mask data. For final comparison, the section 1 to 2 and section 2 to 1 intra-section registration DICE coefficients were averaged, although they only minimally diverged, 0.006 on average across all registrations. Based on this metric, AF outperforms all other image microscopy when chaining linear (Affine) then non-linear registration, although only statistically significantly between AF and BF when using a pair-wise t-test and a = 0.05 (p-value = 0.033). Non-linear registration globally improves overlap by 28.7% (mean DICE from all modalities = 0.287 ± 0.04), with AF performing 33.3% better when using non-linear models. Importantly, AF was also the only modality that had some overlap in every final registration, indicated by the lack of zeroes in the **DICE Non-linear** columns. BF and H&E has 2 and 3 section pairs respectively where there was no mask overlap at any step. However, for some samples, like s3_1 in Table S1, all modalities were performant.

Another critical aspect of this evaluation is determine the expected degree of overlap. As the sections are serial, the annotated ROI changes in size and shape from section to section. This was quantified across all ROI pairs by calculating the difference in area for each, shown in the histogram in Figure S15. Here there is a noticeable distribution of changes favoring increase area. The bias towards ROI expansion is due to the fact that annotations were made between ROIs visible in both sections, and substantial shrinking between sections meant a confident annotation could not be made. The mean ROI area from all annotations was 1393 ± 472 μm^2. The mean absolute change (both smaller and larger) in ROI area for all pairs was 384 ± 274 μm^2. The absolute value accounts for both expanded and shrunk

ROIs between serial sections. Relative to the mean ROI area, the mean change is ~30%, indicating the impossibility to have perfect DICE coefficients. Indeed, it is unwise to aim for perfect overlap in these structures as the registration models would need to become complex and registration artifacts are much more likely to occur with the required fine warping grids. Therefore, in targeting for IMS blind to the degree of overlap, the goal is to have some overlap as IMS ROIs can be expanded to account for target ROI size discrepancies between serial sections at a small penalty of increased instrument time. More

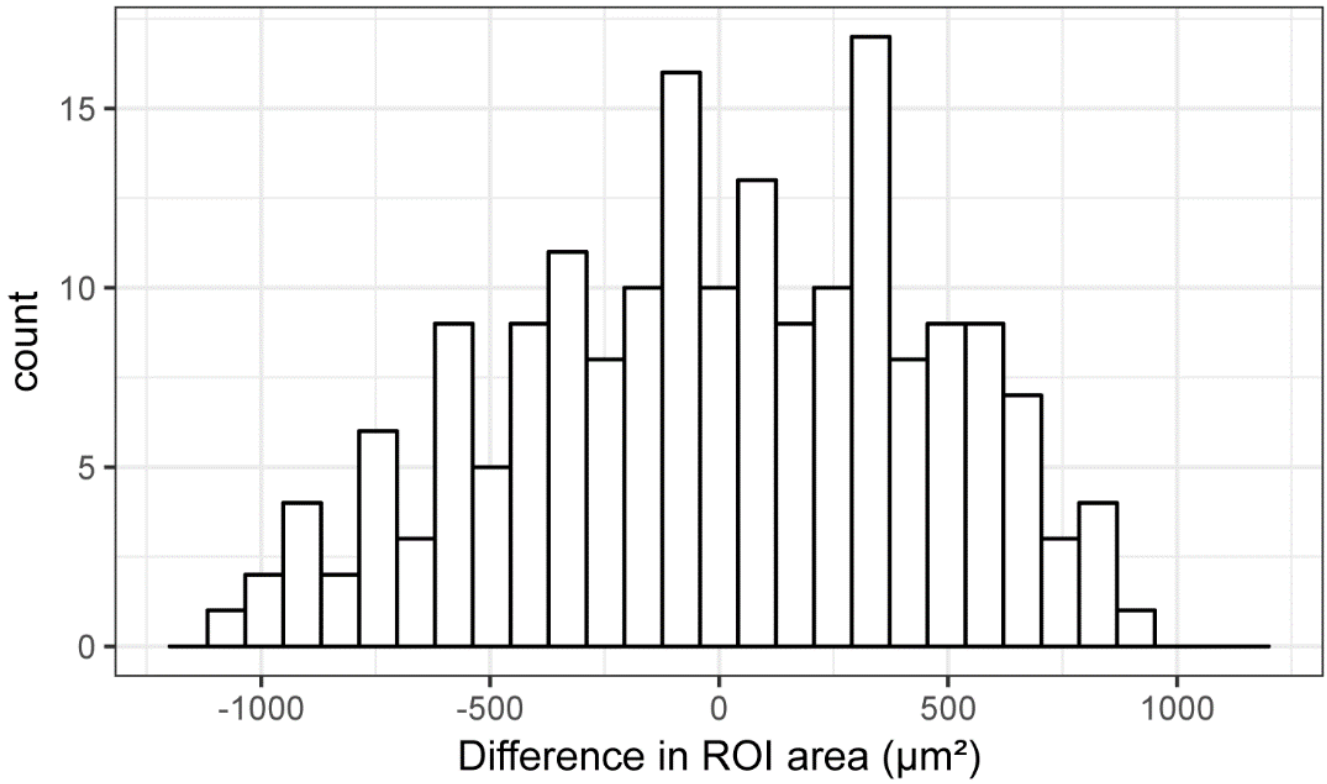


Figure S14. The difference in area between matched infected hepatocytes (n = 188) of the 9 serial sections (3 biological replicates) used for evaluation.

critical is consistent overlap for every ROI pair, *i.e.*, every annotated ROI should have some overlap with its corresponding ROI in the serial section after registration. This consistency is somewhat quantified in the overall DICE coefficient given in Table S1, but given (1), larger ROIs will bias have more weight in the overall DICE calculation. Therefore, to quantify consistency, the number of ROIs with any overlap was divided by the number of ROIs total. The results from this analysis are summarized in Table S2. Again, AF to AF proved the most effective registration with 93.9% of ROIs (n=188) having positive overlap after non-linear registration. BF to BF had 76.6 % of structures overlapping and H&E to H&E had 71.2% of structures overlapping.

1

Modality	Sample Pair	S1->S2 DICE Affine	S2->S1 DICE Affine	S1->S2 DICE Non-linear	S2->S1 DICE Non-linear	Merge DICE Affine	Merge DICE Non-linear
AF to AF	*s1_1*	0.198	0.193	0.525	0.491	0.196	0.508
AF to AF	*s1_2*	0.215	0.215	0.433	0.418	0.215	0.426
AF to AF	*s1_3*	0.124	0.047	0.423	0.406	0.086	0.415
AF to AF	*s2_1*	0.497	0.482	0.621	0.606	0.489	0.614
AF to AF	*s2_2*	0.404	0.412	0.639	0.632	0.408	0.635
AF to AF	*s2_3*	0.409	0.404	0.723	0.729	0.406	0.726
AF to AF	*s3_1*	0.166	0	0.677	0.686	0.083	0.682
AF to AF	*s3_2*	0	0	0.390	0.391	0	0.391
AF to AF	*s3_3*	0.195	0.183	0.692	0.646	0.189	0.669
	Mean	0.245	0.215	0.569	0.556	0.230	0.563
	SD	0.150	0.172	0.121	0.123	0.159	0.122

Modality	Sample Pair	S1->S2 DICE Affine	S2->S1 DICE Affine	S1->S2 DICE Non-linear	S2->S1 DICE Non-linear	Merge DICE Affine	Merge DICE Non-linear
BF to BF	*s1_1*	0	0.037	0.009	0.441	0.018	0.225
BF to BF	*s1_2*	0	0	0	0	0	0
BF to BF	*s1_3*	0.119	0	0.386	0.162	0.060	0.274
BF to BF	*s2_1*	0.483	0.485	0.516	0.461	0.484	0.489
BF to BF	*s2_2*	0.345	0.415	0.528	0.524	0.380	0.526
BF to BF	*s2_3*	0.387	0.419	0.663	0.673	0.403	0.668
BF to BF	*s3_1*	0.200	0.226	0.642	0.630	0.213	0.636
BF to BF	*s3_2*	0.102	0	0.529	0.537	0.051	0.533
BF to BF	*s3_3*	0.134	0.128	0.500	0.506	0.131	0.503
	Mean	0.197	0.190	0.419	0.437	0.193	0.428
	SD	0.162	0.190	0.234	0.206	0.174	0.205

Modality	Sample Pair	S1->S2 DICE Affine	S2->S1 DICE Affine	S1->S2 DICE Non-linear	S2->S1 DICE Non-linear	Merge DICE Affine	Merge DICE Non-linear
H&E to H&E	*s1_1*	0	0.034	0.578	0.502	0.017	0.540
H&E to H&E	*s1_2*	0	0	0	0	0	0
H&E to H&E	*s1_3*	0	0	0	0	0	0
H&E to H&E	*s2_1*	0.496	0.497	0.623	0.627	0.497	0.625
H&E to H&E	*s2_2*	0.380	0.402	0.596	0.596	0.391	0.596
H&E to H&E	*s2_3*	0.360	0.348	0.750	0.699	0.354	0.725
H&E to H&E	*s3_1*	0.221	0.222	0.778	0.726	0.222	0.752
H&E to H&E	*s3_2*	0.094	0.128	0.588	0.591	0.111	0.590
H&E to H&E	*s3_3*	0.096	0.058	0.496	0.460	0.077	0.478
	Mean	0.183	0.188	0.490	0.467	0.185	0.478
	SD	0.178	0.177	0.274	0.262	0.177	0.268

Table S1. Table of DICE coefficients for all ROIs for registrations across all modalities. Merge DICE coefficients are the average of the S1→S2 and S2→S1 registrations. **Non-linear** is an affine followed by a non-linear transformation.

AF to AF Sample	S1->S2 Affine	S1->S2 Non-linear	S2->S1 Affine	S2->S1 Non-linear	Merge Affine	Merge Non-linear
s1_1	0.636	0.939	0.606	1	0.621	0.970
s1_2	0.500	0.750	0.500	0.833	0.500	0.792
s1_3	0.333	0.741	0.222	0.741	0.278	0.741
s2_1	0.882	1	0.882	1	0.882	1
s2_2	1	1	0.955	1	0.977	1
s2_3	0.913	1	0.875	1	0.894	1
s3_1	0.526	1	0	1	0.263	1
s3_2	0	1	0	0.941	0	0.971
s3_3	0.500	1	0.409	0.955	0.455	0.977
Mean	0.588	0.937	0.494	0.941	0.541	0.939
SD	0.282	0.099	0.330	0.083	0.306	0.091

BF to BF Sample	S1->S2 Affine	S1->S2 Non-linear	S2->S1 Affine	S2->S1 Non-linear	Merge Affine	Merge Non-linear
s1_1	0	0.061	0.273	1	0.136	0.530
s1_2	0	0	0	0	0	0
s1_3	0.333	0.815	0	0.481	0.167	0.648
s2_1	0.882	1	0.882	0.941	0.882	0.971
s2_2	0.864	1	1	1	0.932	1
s2_3	0.870	1	0.792	0.958	0.831	0.979
s3_1	0.474	1	0.421	1	0.447	1
s3_2	0.235	0.824	0	0.882	0.118	0.853
s3_3	0.409	0.909	0.409	0.909	0.409	0.909
Mean	0.452	0.734	0.420	0.797	0.436	0.766
SD	0.316	0.363	0.352	0.304	0.325	0.297

H&E to H&E Sample	S1->S2 Affine	S1->S2 Non-linear	S2->S1 Affine	S2->S1 Non-linear	Merge Affine	Merge Non-linear
s1_1	0	0.970	0.182	0.788	0.091	0.879
s1_2	0	0	0	0	0	0
s1_3	0	0	0.037	0	0.019	0
s2_1	0.882	0.941	0.882	0.824	0.882	0.882
s2_2	0.818	1	0.864	0.909	0.841	0.955
s2_3	0.870	1	0.875	0.958	0.872	0.979
s3_1	0.421	1	0.421	1	0.421	1
s3_2	0.294	0.765	0.176	0.882	0.235	0.824
s3_3	0.364	0.864	0.273	0.909	0.318	0.886
Mean	0.405	0.727	0.412	0.697	0.409	0.712
SD	0.335	0.375	0.328	0.358	0.329	0.364

Table S2. Table of ROI consistency for registration between modalities. Values in **Affine** and **Non-linear** columns are number of overlapping ROI pairs after registration / total number of ROI pairs per serial section pair. Values of 1 indicate every ROI pair has a DICE coefficient > 0 (*i.e.,* they overlap to some degree). Merge values are the average of the S1→S2 and S2→S1 registrations. **Non-linear** is an affine followed by a non-linear transformation.

References

(34) Multimodality image registration by maximization of mutual information,Maes, F.; Collignon, A.; Vandermeulen, D.; Marchal, G.; Suetens, P. *IEEE Trans. Med. Imaging* **1997**, *16*, 187-198.

(35) Adaptive Stochastic Gradient Descent Optimisation for Image Registration,Klein, S.; Pluim, J. P. W.; Staring, M.; Viergever, M. A. *Int. J. Comp. Vision* **2008**, *81*, 227.

(36) Cell and tissue autofluorescence research and diagnostic applications,Monici, M. *Biotechnol Annu Rev* **2005**, *11*, 227-256.

(37) Measures of the Amount of Ecologic Association Between Species,Dice Lee, R. *Ecology* **1945**, *26*, 297-302.