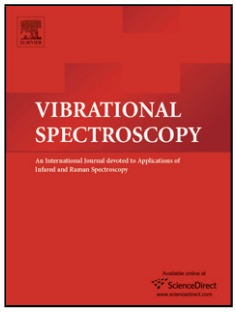
Accepted Manuscript



Title: Enhanced spectral histology in the colon using high-magnification benchtop FTIR imaging

Author: Jayakrupakar Nallala Gavin Rhys Lloyd Michael Hermes Neil Shepherd Nick Stone

PII: S0924-2031(16)30201-6
DOI: http://dx.doi.org/doi:10.1016/j.vibspec.2016.08.013
Reference: VIBSPE 2610

To appear in: *VIBSPE*

Received date: 9-6-2016
Revised date: 12-8-2016
Accepted date: 23-8-2016

Please cite this article as: Jayakrupakar Nallala, Gavin Rhys Lloyd, Michael Hermes, Neil Shepherd, Nick Stone, Enhanced spectral histology in the colon using high-magnification benchtop FTIR imaging, Vibrational Spectroscopy http://dx.doi.org/10.1016/j.vibspec.2016.08.013

Enhanced spectral histology in the colon using high-magnification benchtop FTIR imaging

Jayakrupakar Nallala[a], Gavin Rhys Lloyd[b], Michael Hermes[a], Neil Shepherd[c], Nick Stone[a*]

[a] *Biomedical Physics, School of Physics and Astronomy, University of Exeter, EX4 4QL, UK*

b Biophotonics Research Unit, Gloucestershire Royal Hospital, Gloucester, GL1 3NN, UK

[c] *Department of Pathology, Gloucestershire Hospitals NHS Foundation Trust, Gloucester, GL1 3NN, UK*

**e-mail: n.stone@exeter.ac.uk*

Highlights:

- High-magnification (63X) benchtop FTIR imaging of colon tissues of clinical origin
- Highly resolved detail of cellular features and tissue type boundaries in a label-free manner
- Indication of differences between intracellular mucin and mucin secreted outside the cells.

Abstract:

Label-free imaging of cells and tissues is a promising tool to study the molecular alterations for improved cancer diagnosis. In this regard, vibrational spectroscopic method of FTIR imaging has been employed to study the histo-pathological features and the alterations were correlated to a disease state. The aim of this study was to see if magnification to image pixel size smaller than the diffraction limit of mid-IR wavelengths with two NA objectives could provide cellular and subcellular resolution of key diagnostic features. To this effect, FTIR spectroscopic imaging has been carried out directly on paraffinized colon tissue sections using a benchtop imaging system consisting of two different IR objectives with a NA of 0.81 and 0.62. The high NA objective (0.81) at a high magnification (63X) provided a pixel size of 0.63x0.63 μm^2 in comparison to the standard magnification (12X) where the pixel size is 3.3x3.3 μm^2. The second objective (0.62 NA) at a high magnification (36X) provided a pixel size of 1.1x1.1 μm^2 in comparison to the standard magnification (7X) pixel size of 5.5x5.5 μm^2. The spectral images were corrected for paraffin and other spectral interferences using a modified EMSC algorithm and subjected to cluster analysis in order to appreciate the histopathological details. Comparison of the cluster analysis results to adjacent haematoxylin and eosin stained tissue sections which were used as the morphological controls, revealed various levels of histological and cellular organization. Together with large scale features such as the glandular and connective tissue regions, small scale features like goblet cells, interfaces between tissue types especially the peri-cryptal fibroblastic sheath were observed. Noticeable differences between intracellular mucin and secreted mucin was also observed. However, it appears that high power mid-IR radiation sources would be more beneficial especially when higher signal to noise levels would be needed at smaller time frames.

Keywords: mid-IR spectroscopic imaging, histopathology, colon cancer, goblet cells, high-resolution infrared imaging, FTIR imaging

1. Introduction:

Detection of early molecular changes indicative of disease is vitally important to ensure effective treatment selection. This is particularly important in diseases such as cancer; which continues to be one of the most significant global burdens in terms of both incidence and mortality [1]. Early signs of a disease are manifested by changes at the very basic genomic and proteomic levels which are then expressed at a much larger cellular and tissue level [2]. Diagnostic techniques directed towards identifying these changes in most cases need the samples to be extrinsically labelled to facilitate their detection. Such precise changes detected using these techniques are specific only to a small region of the genome or the proteome. Histopathology, which is the current diagnostic 'gold-standard' for cancer, on the other hand provides information on the morphological changes occurring at a much larger level, from large areas of a diseased tissue [3]. In this case too, the samples need to be labelled for visual inspection. Alternatively, vibrational spectroscopy based in the mid-infrared (mid-IR) can provide a bio-molecular fingerprint of cells [4-7], tissues [8-11] and bio-fluids [12,13] encompassing a wide range of bio-molecular information without the need for any extrinsic labelling. Hence rather than monitoring a single entity or a single biomarker, global changes associated with a disease state could be identified using this approach. Using hyperspectral IR imaging approach, the measured tissue regions are captured simultaneously by an array of pixels on a detector through which spatial detail could be obtained.

One of the factors that determine the image spatial detail is the pixel size that allows collection of sufficient sampling frequency [14]. Based on the Rayleigh criterion, although a theoretical spatial resolution that is of the order of the wavelength (~ 2-12 microns using mid-IR spectral range) could be achieved, lack of sufficient sampling frequency is a limitation. Previously, this has been the case due to limited technical capabilities of the imaging systems including detectors, light sources and optical components. Due to these constraints, direct characterization of smaller scale histological features (such as cells and subcellular features, histological boundaries between tissue types etc) was not entirely possible.

Most of the standard benchtop FTIR imaging systems provided pixel sizes as small as 6.25x6.25 μm^2 or 5.5x5.5 μm^2 [9,11,12]. Some systems provided even smaller pixel sizes as low as 2.6x2.6 μm^2

which were used to elucidate the histological features of tissue sections [15-16]. In recent years, imaging at a smaller scale within the theoretical limits of the mid-IR range using far field or standard microscopy (not considering near field IR spectroscopy) has been attempted using improved optical set ups, detector systems and brighter radiation sources [14, 17-20]. Before this, such attempts were restricted to Synchrotron radiation sources [21,22] that exhibit higher intensity than conventional Globar® sources.

It has been shown recently that using a benchtop instrument with a Globar® and using an enhanced optical set up, it is possible to identify cellular features with a high degree of detail [18,19]. This optical set up consisted of a 0.62 NA objective and with an enhanced optical set up provided 36X magnification corresponding to a pixel size of 1.1x1.1 μm^2 that was five times smaller than that achieved using standard IR optics (usually 5.5x5.5 μm^2). We have shown that, sufficient sampling frequency is generated with this set up through which images with higher detail could be achieved and those when subjected to multivariate statistical analysis can enable detailed histological characterisation of cells and tissues.

In the current study, the aim was to see if further enhancement of magnification together with increased NA of the objective using a similar set up (Agilent benchtop instrument with a Globar® radiation source) could furthermore improve image quality and thereby its histological characterization. For this, a higher NA objective (0.85) with a magnification of 12X in standard mode and 63X in a high magnification mode was used. The high magnification mode provided a sampling frequency corresponding to a pixel size of 0.63x0.63 μm^2 in comparison to the standard magnification where the pixel size is 3.3x3.3 μm^2.

2. Materials and Methods:

FTIR images were obtained using an Agilent FTIR imaging system (Agilent 620 FTIR microscope coupled with an Agilent 670 FTIR spectrometer) with a Globar® source, and a liquid-nitrogen cooled 128x128 FPA detector. Two different IR objectives of 15X (objective 1) and 25X (objective 2) were used for measurements. The first objective (15X) with a NA 0.62 provided a final magnification of 7X in standard mode (there is ~2X de-magnification after the 15X objective) and 36X in high magnification mode at the sample plane (there is ~2.4X extra magnification on top of the 15X),

resulting in a pixel size of 5.5x5.5 μm^2 and 1.1x1.1 μm^2 respectively on the FPA detector. The second objective (25X) with a higher NA of 0.81 provided a magnification of 12X in standard mode (there is ~2X de-magnification after the 25X objective) and 63X in high magnification mode at the sample plane (there is ~2.5X extra magnification on top of the 25X), resulting in a pixel size of 3.3x3.3 μm^2 and 0.66x0.66 μm^2 respectively. For clarity, only the final magnifications at the sample plane (i.e. 7X and 36X for objective 1; and 12 and 63X for objective 2) will be used hereafter. For 7X and 12X modes the incident radiation was used at only 25% throughput. This is required due to the high-sensitivity of the currently used FPA detector to avoid possible speckles. Alternatively, this was offset to a certain extent by increasing the attenuation time compared to what would have been used at full throughput. For 36X and 63X modes, the incident radiation was used at its full capacity of 100%. Samples included a polystyrene film, a chrome-on-glass 1951 USAF target (Edmund Optics) and five formalin fixed paraffin embedded (FFPE) benign colon tissues, obtained with the approval of the local research ethics committee for collection of additional biopsies during routine clinical investigations by colonoscopy. For the polystyrene film, different numbers of scans (16, 32, 64 and 128) was measured at 4cm^{-1} spectral resolution in order to compare the signal to noise levels of the 0.62 NA and 0.81 NA objectives in both standard and high magnifications. The background spectra of air at 256 scans were obtained for this. To calculate the signal, the difference of the maximum and the minimum peak height above the baseline in the signal range of 1500-1480 cm^{-1} was used while the difference of the maximum and the minimum peak in the signal-free range of 2100-2050 cm^{-1} was used for the noise. The median signal to noise ratio (SNR) over all the pixels in the image with error bars of ±1 median absolute deviation from the median (MAD) were calculated.

In order to test the highest resolution achievable with the instrument, an FTIR map of the inner structure of the USAF target consisting of elements 4, 5 and 6 of group 7 was measured using the 0.81 NA objective in high magnification mode. The target being manufactured from chrome on glass, required the measurements to be made in reflectance mode to enable the full mid-IR spectral range to be utilized. Background spectra were collected at 256 number of scans from the chrome on glass region, which was then ratioed against glass to get the contrast. The mosaic image itself consisted of 3x3 tiles collected with 128 scans each at 32 cm^{-1} spectral resolution. Data was loaded into R using the software package hyperSpec [23], and ggplot2 [24] was used to generate the figures. Spectra

containing detector artefacts were excluded by considering only spectra with an average total intensity of below 1.2 for further analysis.

Colon tissues were sectioned to 7 microns and placed on mid-IR transparent calcium fluoride (CaF_2) substrates. They were measured without chemical de-paraffinization. To account for the potential interferences from paraffin, which has a strong FTIR absorbance spectrum, a mathematical de-paraffinization was performed after imaging to minimize paraffin signal contributions; an approach we have previously shown to be highly effective [25-27]. The tissue sections were measured using 64 scans at 4 cm^{-1} spectral resolution and multi-mosaic images (multiple FPA tiles) were obtained depending on the size of the sample and the type of objective/magnification used. A clean area of the CaF_2 slide was used for the background, measured at 256 scans.

All the post-measurement analyses were carried out using in-house programs written in Matlab R2014b® (Mathworks, USA). The pre-processing of the raw IR spectra was carried out in a similar way to the previous work [25-27]. In brief, the IR images of the colon tissues were independently pre-processed using a modified extended multiplicative signal correction algorithm (EMSC). EMSC based mathematical de-paraffinization is now commonly being employed to minimize the variance contributions of paraffin in paraffinized tissue sections [19,28]. The algorithm employs a pure paraffin model (obtained from tissue free area of paraffin and consisting of mean paraffin spectrum and its first ten principal components) as an interference matrix incorporated into the EMSC model. Then using a fitting algorithm the paraffin variability across all the pixels of the image is minimized and only the biomolecular variability arising from the biological tissue is retained. The spectral images were also simultaneously corrected for baseline (polynomial order 4) and normalized using the same EMSC algorithm. The EMSC corrected spectral images were subjected to cluster analysis based on the K-means clustering algorithm (using Euclidean distance) with the aim to segment the various histological features of the tissue sections. The spectra corresponding to pure paraffin regions (without any tissue features) and the spectra that are noisy with low signal to noise identified by the EMSC algorithm as outliers are coloured as white pixels in all of the IR images [27]. The cluster centroids were plotted and the spectral distances between clusters were calculated (Hierarchical clustering using Ward's linkage algorithm) and represented in the form of a dendrogram in order to visualize the heterogeneity between different histological groups. K-means is an unsupervised iterative algorithm that is a well-suited method for processing large data sets such as in this case [29]. In addition,

Spatially Guided Fuzzy C-Means (sgFCM) clustering was applied to a selected subset of the samples to test if any additional information could be obtained in comparison to the K-means clustering [30]. This method has been shown to be effective in segmentation of MRI images with noise and intensity inhomogeneity and works by high probability of grouping with the neighbouring pixel if they share similar characteristics. Both pre-processing and further analysis of the colon tissue images were carried out in the spectral range of 1800-1000 cm^{-1} owing to the specificity of this region to biomolecular features. All measurements were carried out in the spectral range of 3800-1000 cm^{-1} and in transmission mode except for the USAF target where reflection mode was employed.

3. Results and Discussion:

3.1. Signal to Noise optimization: The signal to noise levels when using various experimental parameters and objectives using the benchtop FTIR imaging system were compared in order to establish sufficient SNR levels for further measurements. For this, different numbers of scans were tested on a polystyrene window using the two different IR objectives (NA of 0.62 and 0.81) at two different magnifications each (standard and high magnifications). As shown in Figure 1, there is an approximately linear trend of SNR improvement with increasing number of scans for each of the modalities. The standard magnification of both objectives (7X and 12X) on average show a higher SNR compared to the high magnification set ups (36X and 63X). This is expected as lower magnifications benefit from light throughput to the detector pixels, while higher magnifications (and smaller image pixel sizes) lead to reduced light reaching each detector pixel. Nonetheless, with the current high magnification set ups, there still appears to be comparable signal to noise levels to the standard ones although on the lower side and with an appropriate number of scans, sufficient signal to noise levels could be achieved with all configurations. Based on this and in order to establish a balance between the signal quality and time constraints especially for imaging of larger samples, 64 scans per FPA tile were used for further measurements with respect to both the objectives and their magnifications. This corresponded to 5 min of acquisition time per FPA tile.

3.2. Characterization of resolving capability: In order to measure the highest resolution of the system, rather than the pixel size, the elements 4, 5 and 6 of group 7 on the USAF target were

measured using the objective with the highest NA at the highest magnification (0.81 NA and final magnification of 63X). The lines / spacings of this group are separated by 2.76, 2.46 and 2.19 microns respectively. Figure 2 shows the images obtained by visualizing the average intensity per pixel of a low wavenumber region (Figure 2A, 1400 to 1000 cm^{-1}) and a higher wavenumber region (Figure 2B, 3900 to 3800 cm^{-1}). Figure 2C shows a zoomed area of the smallest structures (from Figure 2B) of the USAF target for the region 3900-3800 cm^{-1} and below it the corresponding intensity along the coloured cross sections. The differences in intensity along the cross section show that all of the structures of the target are clearly resolvable in the high wavenumber region. The Rayleigh criterion for optical resolution requires a differences in detected intensity of at least 25% to consider two points being separated is clearly fulfilled for all structures here [31]. However this only gives discrete resolution for specific wavelengths and estimating the resolution of an image when using multivariate analysis of an entire spectrum is still a subject of ongoing research. To resolve the USAF target features as small as ~2.19 microns (element 6 of group 7) using a pixel size of 0.66x0.66 μm^2, needed a sampling frequency of ~ 3.5. Further, pixels were binned to explore how reduction in the sampling frequency affects the resolution of the USAF target features. For this, pixels from figure 2C were binned two times (figure 2D) and four times (figure 2E) and their respective intensity profiles plotted (corresponding plots below figures 2D and E respectively). The USAF target features were still resolvable (following the Rayleigh criterion) for figure 2D (sampling frequency of ~ 1.8), while the structures become no more resolvable as can be observed by the loss of the intensity plot in figure 2E (sampling frequency of ~ 0.9). Overall, the plot shows a decreasing intensity profile for resolving the USAF target features as the sampling frequency decreases. It has to be noted that the difference in spatial resolution between the vertical and horizontal axis especially in figure 3C, is an inherent effect caused by the sampling geometry. In transflection mode, only half of the objective is illuminated by the source whereas the other half of it is used to collect the transflected light. This effectively results in a lower opening angel along one of the axis and therefore decreases spatial resolution along that axis.

3.3. Characterization of tissue features: While the USAF standard provided a good reference to test the imaging system, the ultimate resolving capabilities were tested directly on real samples from clinical origin. This provides a direct evidence for the features of interest that could be resolved for

diagnosis of a disease. To this effect, initially the same region of interest of a paraffinized colon tissue sample was measured using both IR objectives and at two different magnifications each corresponding to different pixel sizes. The details corresponding to these objectives and the achievable pixel sizes is shown in Table 1. A total of 4 images using the different parameters were obtained from this sample, pre-processed and subjected to independent K-means cluster analysis. As shown in Figure 3, a comparison of the cluster analysis results demonstrates the glandular or epithelial regions (crypts) and the connective tissue regions (regions between the glands/crypts), but with a different degree of detail in each of the 4 images. These variations were prominent especially in and around the epithelium corresponding to nuclear regions, cytoplasmic regions, mucin features and goblet cells as discussed in detail later in this section.

Although, a higher NA is associated with higher spatial resolution, optimal sampling frequency is still essential to attain the necessary image detail. This is exemplified in figure 3C, where although a setup of higher NA (NA 0.81) was used in comparison to figure 3B (NA 0.62), the obtained image detail is lesser compared to the latter. It has to be noted that although a higher NA objective was used in figure 3C, the pixel size was 3.3x3.3 μm^2 as compared to the lower NA objective for which the pixel size was 1.1x1.1 μm^2. In terms of sampling frequency for the given size of the image which was ~ 416x228 microns; figure 3C was obtained using 126x69 pixels while figure 3b was obtained using 378x207 pixels respectively. Therefore, obtaining a sufficient sampling frequency using optimal magnifications could enhance the resolving capabilities of a system.

This is directly translated into histopathological characterization of the analyzed tissue where the degree of elucidated tissue features differed; examples in this case are the cellular features such as goblet cells. As can be seen in Figure 3B, there is a clear separation of goblet cells features (cluster 8) while in Figure 3C these regions/clusters appear averaged across pixels. Nonetheless, the best image for cellular detail was achieved using the high magnification mode of the 0.81 NA objective, where the final magnification was 63X with a pixel size of 0.66x0.66 μm^2 as shown in Figure 3D. In this case, a very clear pattern of the glandular structures/crypts (clusters 8, 4, 6, 2, 7, 9 and 1) and the connective tissue (clusters 5 and 3) could be resolved.

In comparison to previous studies based on mid-IR imaging including our own, several interesting features were observed in the current study. Firstly, as reported previously [19], the goblet cell

features were immediately distinguishable along the edge of the crypts (clusters 6 and 2) based on the dominant mucin spectral features and the corresponding dendrogram (Figure 4) and their localization with in and along the outer region of the crypt [30]. Additionally in this case, features that can be correlated to mucin secretion from the goblet cells to the lumen of the crypt were observed. Cluster 8 in Figure 3D, which is the extra cellular mucin present in the central lumen of the crypt also appears as part of some of the goblet cells (clusters 6 and 2) but always facing towards the internal part of the gland. This continues from the cellular part towards the crypt lumen. This observation gives rise to the hypothesis if the differences in spectral signatures are discriminating the mucin internalized within the goblet cells from the mucin secreted outside of the cells. The clusters 6 and 8 although both correspond to mucin they are distinctively localised with a specific pattern. This distinct clustering within and outside the cell potentially indicates that extracellular mucin is distinct from the intracellular mucin. Whether this is associated with localisation between secreted or trans-membrane type mucins or is just related to the pure form of mucin secreted out of the cells needs to be further evaluated. Additional studies in conjunction with glycoprotein assays could provide more insights into this observation, the validation of which could enable label-free and rapid characterization of mucins which are implicated in a wide range of diseases including cancer.

Secondly, the boundaries between glandular tissue and connective tissue were more apparent compared to previous studies [32-34]. In particular, cluster 1 established a clear separation at the epithelial-connective tissue boundary which is followed by a second layer as cluster 9. Previously this boundary was always accredited to the outer nuclear region of the crypt due to lack of high degree of detail [32]. It is however known that this region is in the vicinity of the basement membrane and the peri-cryptal fibroblastic sheath (PCFS). The presence of two distinct regions surrounding the crypt indicate that the layer of PCFS corresponding to the basement membrane of the colonic epithelium specifically in conjunction with the outer nuclear part is being resolved. The ability to resolve these features using a label-free spectral histopathological approach could enable rapid identification of invasive cancers as the degradation of the PCFS in conjunction with the basement membrane is one of the vital signs of tumor invasiveness [35].

In some of the other samples analyzed, pixelation of the connective tissue region was observed and sufficient segmentation of the histological features could not be obtained. On closer examination, a low variance between histological groups was noted which potentially could be due to the combination

of the current optical set up with the low brightness Globar® radiation source, with the point spread function (PSF) at the longer wavelength covering several image pixels. This observation points out to the importance of employing high intensity radiation sources such as synchrotrons, or laser sources such as quantum cascade lasers (QCL) or supercontinuum sources, that could potentially improve the signal to noise levels to overcome such issues.

To explore if the connective tissue group pixel membership could be improved using an alternative less constrained approach, cluster analysis based on sgFCM was tested on these images. This method has been used elsewhere to address similar issues in magnetic resonance imaging (30). As an example, the same FTIR image of a tissue interphase between an adenomatous and a normal gland processed using K-means clustering and sgFCM is shown in Figure 5. As can be seen in the Figure 5A (K-means image), most of the connective tissue region (between the glands) appear pixelated (clusters 4 and 6). By the use of sgFCM (Figure 5B), this could be overcome where a better segmentation of both the epithelial (clusters 1,2,4,5 and 7) and the connective tissue regions (cluster 3) was obtained. Although no major segmentation difference could be observed between the adenomatous and the normal gland, clusters 5 and 7 appears to belong exclusively to the normal gland. However, the data is not sufficient in this case to imply any pathological evidence at this point and it is the segmentation of different tissue types (epithelium and connective tissue) that is of interest.

Finally, four different tissue interfaces (epithelium, lamina propria, muscularis mucosa and submucosa) from the same tissue section were measured and analyzed using K-means cluster analysis. As shown in Figure 6, all the tissue interfaces were segmented in a way directly comparable to the HE stained tissue (Epithelial glands – cluster 1 and 7, lamina propria – cluster 3, muscularis mucosa – cluster 6 and submucosa – clusters 2; clusters 4 and 5 could not be attributed to any histological features). Further, the identification of boundaries especially between epithelium and the remaining connective tissue of lamina propria, submucosa and muscularis mucosa is much defined and easier to identify in the IR cluster image.

The exploration of label-free diagnostic methods is a highly sought after area of research and mid-IR spectral imaging is a promising tool in this regard. This is now well documented where several studies have shown the potential of mid-IR imaging as a diagnostic tool applied to various types of cancer

[9,10,36-38]. To precisely interpret the biomolecular changes either in a manual or an automatic manner, it is essential to have highly resolved images. Until recently, although mid-IR wavelengths as low as ~2 microns were used for imaging, due to the limitations in instrument capabilities, such spatial resolutions in an image could not be achieved. Although, to discriminate a diseased tissue from a non-diseased tissue, such high resolution imaging may not be required [9,11], it is of great importance from a clinical perspective to obtain as pure spectral information as possible to monitor molecular changes at the smallest level possible. Such capabilities could potentially reduce the chances of mis-diagnosis (over or under diagnosis) as the boundaries between cells or tissue types would be much more defined and easier to interpret.

One of the important reasons for not being able to achieve diffraction limited resolution of the order of the wavelength, was the lack of optimal sampling frequency or smaller pixel sizes. With the very recent developments in various instrument parameters, it is now becoming possible to obtain a higher sampling frequency by effectively reducing the pixel sizes by improved magnifications at the sample plane [14,17-19]. According to the Rayleigh criterion, the diffraction limited spatial resolution = $0.61\lambda/NA$, where λ is the wavelength of the light and NA is the numerical aperture of the objective. Therefore using the IR objective of NA 0.81, the achievable theoretical spatial resolution at 3900-3800 cm^{-1} (equivalent to ~2.6 micrometres) is ~2 microns. Using this at 63X magnification corresponding to a pixel size of 0.66x0.66 μm^2, it was possible to resolve the USAF target features as small as ~2.19 microns (element 6 of group 7, figure 2) which therefore needed a sampling frequency of ~ 3.5.

However, the analysis of the tissue sections was performed in the fingerprint region of 1800-1000 cm^{-1} (equivalent to ~5.5 to 10 micrometres). With the high NA objective of 0.81, the achievable theoretical spatial resolution in this region is 4 to 7.5 microns. The features of interest that were targeted to be resolved were still within these scales which included the cellular features, the epithelial-connective tissue boundaries and especially the boundary that corresponded to the PCFS. Such improvements were possible due to the improved magnifications using smaller pixel sizes.

In our previous study, imaging of colon tissues was carried out using a 0.62 NA objective at 36X magnification and a pixel size of 1.1×1.1 μm^2 [19]. In comparison to this, the current study demonstrates an important improvement in the technical capability of the current FTIR bench top imaging systems, where the suitability of using a higher NA objective at 63X magnification and a pixel

size of 0.66×0.66 μm^2 is demonstrated. Furthermore, in going from 1.1×1.1 μm^2 to 0.66×0.66 μm^2, the improvement in resolving tissue interfaces between epithelium and connective tissue, the PCFS and mucin associated patterns are highlighted in the current study.

Although the spectral images of colon tissue shown here provide potentially important detail, there are certain limitations when using such high-resolution imaging methods. Due to the higher magnifications and smaller pixels sizes, the time taken to measure tissue regions increases to a large extent and thereby the amount of data collected. Therefore with the current technical capabilities such application would be feasible only to screen specific regions of interest that are of diagnostic importance and that are not possible to resolve using standard approaches. Nevertheless, such specific details would provide vital information to clinicians/pathologists. However, with improvements in detector areas, pixel densities, sizes such as megapixel detectors, and further improvements to optical set ups, it might become possible to scan large areas at high magnifications.

Alternatively, discrete frequencies could be used provided sub-spectral sampling is sufficient, reducing the amount of data and associated acquisition time to a large extent [39]. Such approaches could be made possible using high power radiation sources such as QCLs [40-42] or supercontinuum sources that are under development [43, 44]. Such sources would also mean improved signal to noise levels thereby providing a chance to mitigate the time constraints using the current systems. One of the challenges in this study was that the variance in the images with the pixel size of 0.66x0.66 μm^2 was quite low to be able to exploit the spectral information. This could have been improved to an extent by increasing the number of scans per FPA tile with a compromise with the time constraints. However, in the future the use of high power radiation sources could potentially overcome this problem allowing correlating the morphological differences to the spectral differences.

4. Conclusions:

The current study demonstrates a label-free biomolecular approach based on FTIR spectroscopic imaging to characterize cellular and histological features of FFPE colon tissues from clinical origin at a high degree of detail. We have shown that FTIR imaging at reduced pixel sizes using higher magnifications improves the degree of histological detail that could be obtained. This is demonstrated by detailed segmentation of goblet cell features and tissue interphases between epithelium and

connective tissue especially the PCFS. Further there appears to be an indication that the intracellular mucin is different from the mucin secreted outside the cells. For conventional IR radiation sources, obtaining sufficient signal to noise levels might still be a limitation with respect to time constraints where traditional multivariate algorithms need to be adapted for effective segmentation of the spectral data. Alternatively, high power radiation sources could potentially overcome these issues. FTIR imaging at higher magnifications appears to be an advance revealing further small scale details for obtaining molecular information for the diagnosis of cancers.

5. Acknowledgments:

This work is part of the MINERVA project (http://www.minerva.eu) which is supported by the European Union's Seventh Framework Programme (FP7) for research, technological development and demonstration under grant agreement no 317803.

6. References:

1. J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, F. Bray, GLOBOCAN 2012 v1, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet], Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr.

2. B. Vogelstein, K. W. Kinzler, Cancer genes and the pathways they control, Nat. Medicine, 10, 2004, 789–799.

3. M. V. Brown, J. E. McDunn, E. M. Smith, M. V. Milburn, D. A. Troyer, K. A. Lawton, Cancer detection and biopsy classification using concurrent histopathological and metabolomic analysis of core biopsies, Genome Med., 4:33, 2012.

4. D. Townsend, M. Miljković, B. Bird, K. Lenau, O. Old, M. Almond, C. Kendall, G. Lloyd, N. Shepherd, H. Barr, N. Stone, M. Diem, Infrared micro-spectroscopy for cyto-pathological classification of esophageal cells, Analyst, 140, 2215-2223, 2015.

5. S. Yao, M. Moenner, A. Engdahl, C. Petibois, Use of synchrotron-radiation-based FTIR imaging for characterizing changes in cell contents, Anal. Bioanal. Chem., 404, 1311-1316, 2012.

6. J. M. Schubert, B. Bird, K. Papamarkakis, M. Miljković, K. Bedrossian, N. Laver, M. Diem, Spectral cytopathology of cervical samples: detecting cellular abnormalities in cytologically normal cells, Lab. Invest., 90, 1068-1077, 2010.
7. J. K. Pijanka, D. Kumar, T. Dale, I. Yousef, G. Parkes, V. Untereiner, Y. Yang, P. Dumas, D. Collins, M. Manfait, G. D. Sockalingum, N. R. Forsyth, J. Sulé-Suso, Vibrational spectroscopy differentiates between multipotent and pluripotent stem cells, Analyst, 135, 3126-3132, 2010.

8. J. T. Kwak, A. Kajdacsy-Balla, V. Macias, M. Walsh, S. Sinha, R. Bhargava, Improving prediction of prostate cancer recurrence using chemical imaging, Sci. Rep., 5, 8758, 2015.

9. D. C. Fernandez, R. Bhargava, S. M. Hewitt, I. W. Levin, Infrared spectroscopic imaging for histopathologic recognition, Nat. Biotechnol., 23, 469-474, 2005.

10. C. Peng, S. Kaščáková, F. Chiappini, N. Olaya, C. Sandt, I. Yousef, D. Samuel, P. Duma, C. Guettier, F. LeNaour, Discrimination of cirrhotic nodules, dysplastic lesions and hepatocellular carcinoma by their vibrational signature, J. Transl. Med., 12, 14:9, 2016.

11. J. Nallala, M. D. Diebold, C. Gobinet, O. Bouche, G. D. Sockalingum, O. Piot, M. Manfait, Infrared spectral histopathology for cancer diagnosis: a novel approach for automated pattern recognition of colon adenocarcinoma, Analyst, 139, 4005-4015, 2014.

12. C. Hughes, G. Clemens, B. Bird, T. Dawson, K. M. Ashton, M. D. Jenkinson, A. Brodbelt, M. Weida, E. Fotheringham, M. Barre, J. Rowlette, M. J. Baker, Introducing discrete frequency infrared technology for high-throughput biofluid screening, Sci. Rep., 6, 20173, 2016.

13. J. R Hands, G. Clemens, R. Stables, K. Ashton, A. Brodbelt, C. Davis, T. P. Dawson, M. D. Jenkinson, R. W. Lea, C. Walker, M. Baker, Brain tumour differentiation: rapid stratified serum diagnostics via attenuated total reflection Fourier-transform infrared spectroscopy, J. neuro-oncol., 127 (3), 463-472, 2016.

14. R. K. Reddy, M. J. Walsh, M. V. Schulmericha, P. S. Carney, R. Bhargava, High-definition infrared spectroscopic imaging, Appl. Spectrosc., 67(1), 93-105, 2013.

15. C. Krafft, D. Codrich, G. Pelizzo, V. Sergo, Raman and FTIR microscopic imaging of colon tissue: a comparative study, J. Biophoton., 1, 154-169, 2008.

16. R. Noreen, R. Pineau, C. C. Chien, M. Cestelli-Guidi, Y. Hwu, A. Marcelli, M. Moenner, C. Petibois,
Functional histology of glioma vasculature by FTIR imaging, Anal. Bioanal. Chem., 401, 795-801, 2011.

17. C. R. Findlay, R. Wiens, M. Rak, J. Sedlmair, C. J Hirschmugl, J. Morrison, C. J. Mundy, M. Kansiz, K. M. Gough, Rapid biodiagnostic *ex vivo* imaging at 1 μm pixel resolution with thermal source FTIR FPA, Analyst, 140, 2493-2503, 2015.

18. C. Hughes, A. Henderson, M. Kansiz, K.M. Dorling, M. J. Hernandez, M.D. Brown, N.W. Clarke, P. Gardner, Enhanced FTIR bench-top imaging of single biological cells, Analyst, 140, 2080-2085, 2015.

19. J. Nallala, G. R. Lloyd, N. Shepherd, N. Stone, High-resolution FTIR imaging of colon tissues for elucidation of individual cellular and histopathological features, Analyst, 141 (2), 630-639, 2016.

20. L. S. Leslie, T. P. Wrobel, D. Mayerich, S. Bindra, R. Emmadi, R. Bhargava, High definition infrared spectroscopic imaging for lymph node histopathology, PLoS ONE, 10 (6), 2015.

21. M. Z. Kastyak-Ibrahim, M. J. Nasse, M. Rak, C. Hirschmugl, M. R. Del Bigio, B. C. Albensi, K. M. Gough, Biochemical label-free tissue imaging with subcellular-resolution synchrotron FTIR with focal plane array detector, Neuroimage., 60(1), 376-83, 2012.

22. M. J. Nasse, M. J. Walsh, E. C. Mattson, R. Reininger, A. Kajdacsy-Balla, V. Macias, R. Bhargava, C. J. Hirschmugl, High-resolution Fourier-transform infrared chemical imaging with multiple synchrotron beams, Nat. Methods, 8, 413–416, 2011.

23. C. Beleites, V. Sergo, hyperSpec: a package to handle hyperspectral data sets in R, R package, v ersion 0.98-20150304. http://hyperspec.r-forge.r-project.org.

24. H. Wickham, ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009.

25. R. Wolthuis, A. Travo, C. Nicolet, A. Neuville, M. P. Gaub, D. Guenot, E. Ly, M. Manfait, P. Jeannesson, O. Piot, IR spectral imaging for histopathological characterization of xenografted human colon carcinomas, Anal. Chem., 80, 8461-8469, 2008.

26. E. Ly, O. Piot, R. Wolthuis, A. Durlach, P. Bernard, M. Manfait, Combination of FTIR spectral imaging and chemometrics for tumour detection from paraffin-embedded biopsies, Analyst, 133, 197-205, 2008.

27. J. Nallala, C. Gobinet, M. D. Diebold, V. Untereiner, O. Bouché, M. Manfait, G. D. Sockalingum, O. Piot, Infrared spectral imaging as a novel approach for histopathological recognition in colon cancer diagnosis, J. Biomed. Opt., 17, 1-12, 2012.

28. J. Nallala, G. R. Lloyd, N. Stone, Evaluation of different tissue de-paraffinization procedures for infrared spectral imaging, Analyst, 140, 2369-2375, 2015.

29. P. Lasch, W. Haensch, D. Naumann, M. Diem, Imaging of colorectal adenocarcinoma using FT-IR microspectroscopy and cluster analysis, Biochim. Biophys. Acta., 1688(2), 176-186, 2004.

30. S. K. Adhikari, J. K. Sing, D. K. Basu, M. Nasipuri, Conditional spatial fuzzy C-means clustering algorithm for segmentation of MRI images, Appl. Soft. Computing, 34, 758-769, 2015.

31. Rayleigh XXXI. Investigations in optics, with special reference to the spectroscope Philosophical Magazine Series 5, Philosophical Magazine Series 5, Taylor & Francis, 8, 261-274, 1879.

32. J. Nallala, O. Piot, M. D. Diebold, C. Gobinet, O. Bouché, M. Manfait, G. D. Sockalingum, Infrared and Raman imaging for characterizing complex biological materials: a comparative morpho-spectroscopic study of colon tissue, Appl. Spectrosc., 68, 1, 2014.

33. J. Nallala, O. Piot, M. D. Diebold, C. Gobinet, O. Bouché, M. Manfait, G. D. Sockalingum, Infrared imaging as a cancer diagnostic tool: introducing a new concept of spectral barcodes for identifying molecular changes in colon tumors, Cytometry Part A, 83, 294-300, 2013.

34. A. Travo, O. Piot, R. Wolthuis, C. Gobinet, M. Manfait, J. Bara, M. E. Forgue-Lafitte, P. Jeannesson, IR spectral imaging of secreted mucus: a promising new tool for the histopathological recognition of human colonic adenocarcinomas, Histopath., 56, 921-931, 2010.

35. H. Mutoh, S. Sakurai, K. Satoh, H. Osawa, T. Tomiyama, H. Kita, T. Yoshida, K. Tamada, H. Yamamoto, N. Isoda, K. Ido, K. Sugano, Pericryptal fibroblast sheath in intestinal metaplasia and gastric carcinoma, Gut, 54, 33-39, 2005.

36. R. Baker, K. D. Rogers, N. Shepherd, N. Stone, New relationships between breast microcalcifications and cancer, Br. J. Cancer, 103, 1034-1039, 2010.

37. R. K. Sahu, S. Argov, S. Walfisch, E. Bogomolny, R. Moreha, S. Mordechai, Prediction potential of IR-micro spectroscopy for colon cancer relapse, Analyst, 135, 538-544, 2010.

38. J. Einenkel, U. D. Braumann, W. Steller, H. Binder, L. C. Horn, Suitability of infrared microspectroscopic imaging for histopathology of the uterine cervix, Histopath., 60, 1084-1098, 2012.

39. G. R. Lloyd, N. Stone, Method for identification of spectral targets in discrete frequency infrared spectroscopy for clinical diagnostics, Appl Spectrosc., 69(9),1066-73, 2015.

40. M. R. Kole, R. K. Reddy, M. V. Schulmerich, M. K. Gelber, R. Bhargava, Discrete frequency infrared microspectroscopy and imaging with a tunable quantum cascade laser, Anal. Chem., 84, 10366-10372, 2012.

41. K. Yeh, S. Kenkel, J. N. Liu, R. Bhargava, Fast infrared chemical imaging with a quantum cascade laser, Anal. Chem., 87 (1), 485-493, 2015.

42. N. Kröger, A. Egl, M. Engel, N. Gretz, K. Haase, I. Herpich, B. Kränzlin, S. Neudecker, A. Pucci, A. Schönhals, J. Vogt, W. Petrich, Quantum cascade laser–based hyperspectral imaging of biological tissue, J. Biomed. Opt., 19(11), 111607, 2014.

43. C. R. Petersen, U. Møller, I. Kubat, B. Zhou, S. Dupont, J. Ramsay, T. Benson, S. Sujecki, N. Abdel-Moneim, Z. Tang, D. Furniss, A. Seddon, O. Bang, Mid-infrared supercontinuum covering the 1.4-13.3 μm molecular fingerprint region using ultra-high NA chalcogenide step-index fibre, Nat. Photonics, 8, 830-834, 2014.

44. I. D. Lindsay, S. Valle, J. Ward, G. Stevens, M. Farries, L. Huot, C. Brooks, P. M. Moselund, R. M. Vinella, M. Abdalla, D. de Gaspari, R. M. von Wurtemberg, S. Smuk, H. Martijn, J. Nallala, N. Stone, C. Barta, R. Hasal, U. Moller, O. Bang, S. Sujecki, A. Seddon, Towards supercontinuum-driven hyperspectral microscopy in the mid-infrared, Proc. SPIE 9703, Optical Biopsy XIV: Toward Real-Time Spectroscopic Imaging and Diagnosis, 970304, 2016.

Figure captions:

Figure 1: Signal to noise ratio comparison between the O.62 NA (7X and 36X magnification) and 0.81 NA (12X and 63X magnification) infrared objectives at different number of scans. Plots indicate median signal to noise ratio over all the pixels of the IR image of a polystyrene film with error bars of ±1 median absolute deviation from the median.

Figure 2: FTIR image of chrome on glass USAF 1951 target of the elements 4, 5 and 6 of group 7 measured using 0.81 NA infrared objective at 63X magnification corresponding to a pixel size of 0.66x0.66 μm^2. A. Average intensity per pixel in the region 1400-1000 cm^{-1}; B. Average intensity per pixel in the region 3900-3800 cm^{-1}; C. Zoomed in area of the boxed region from B with its corresponding average intensity along the coloured lines; D. 2X2 binned image of C with its corresponding average intensity along the coloured lines; E. 4X4 binned image of C with its corresponding average intensity along the coloured lines.

Figure 3: Comparison of histological resolving capabilities of FTIR benchtop imaging with cluster analysis of a normal colon tissue with respect to the HE stained morphological control using different NA objectives. A. 0. 62 NA objective at 7X magnification and a pixel size of 5.5×5.5 μm^2; B. 0. 62 NA objective at 36X magnification and a pixel size of 1.1×1.1 μm^2; C. 0. 81 NA objective at 12X magnification and a pixel size of 3.3×3.3 μm^2; D. 0. 81 NA objective at 63X magnification and a pixel size of 0.66×0.66 μm^2. All the images were analyzed independently using different cluster numbers and the color are randomly assigned for a given image. The size of the measured tissue was ~ 416x228 microns.

Figure 4: Cluster centroid spectra and dendrogram corresponding to figure 3D in the spectral region of 1800-1000 cm^{-1}. The cluster centroids show typical glycoprotein features corresponding to mucin indicated at the respective wavenumbers. The corresponding dendrogram representing the heterogeneity of the clusters is also shown.

Figure 5: Comparison of FTIR-cluster analysis results of a benign colon tissue using K-means clustering and spatially guided Fuzzy C-means clustering with respect to the HE stained morphological control. Images were obtained using a 0.81 NA objective at 63X magnification and a pixel size of 0.66×0.66 μm^2. The size of the measured tissue was ~ 169x169 microns.

Figure 6: Histological characterization of a normal colon tissue using FTIR benchtop imaging with cluster analysis in comparison to the HE stained morphological control. The FTIR image was acquired using a 0.81 NA objective at 63X magnification and a pixel size of 0.66×0.66 μm^2. The size of the measured tissue was ~169x338 microns.

Fig 1

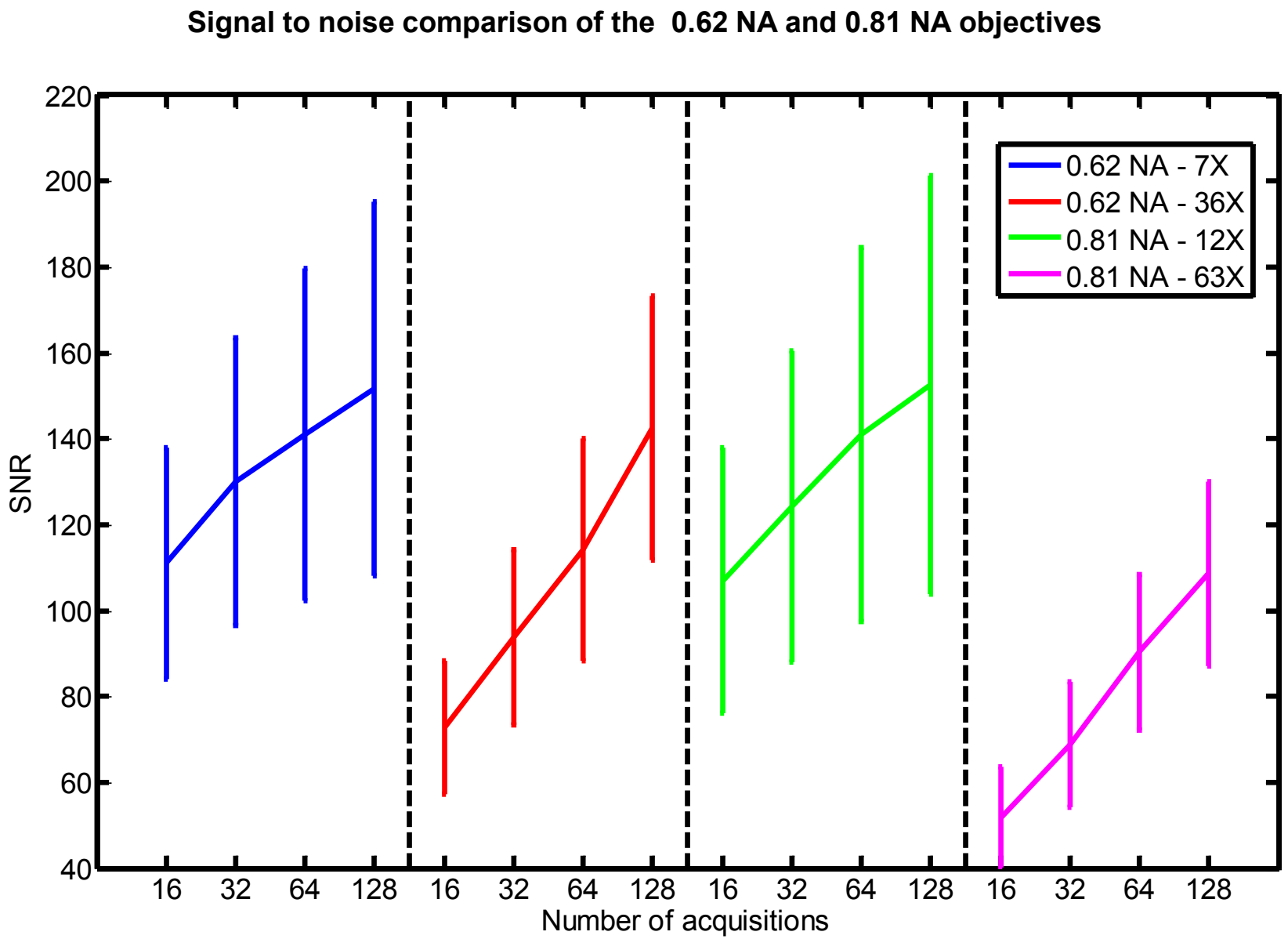

Signal to noise comparison of the 0.62 NA and 0.81 NA objectives
SNR
220
200
180
160
140
120
100
80
60
40
16
32
64
128
16
32
64
128
16
32
64
128
16
32
64
128
Number of acquisitions
0.62 NA - 7X
0.62 NA - 36X
0.81 NA - 12X
0.81 NA - 63X

Fig 2

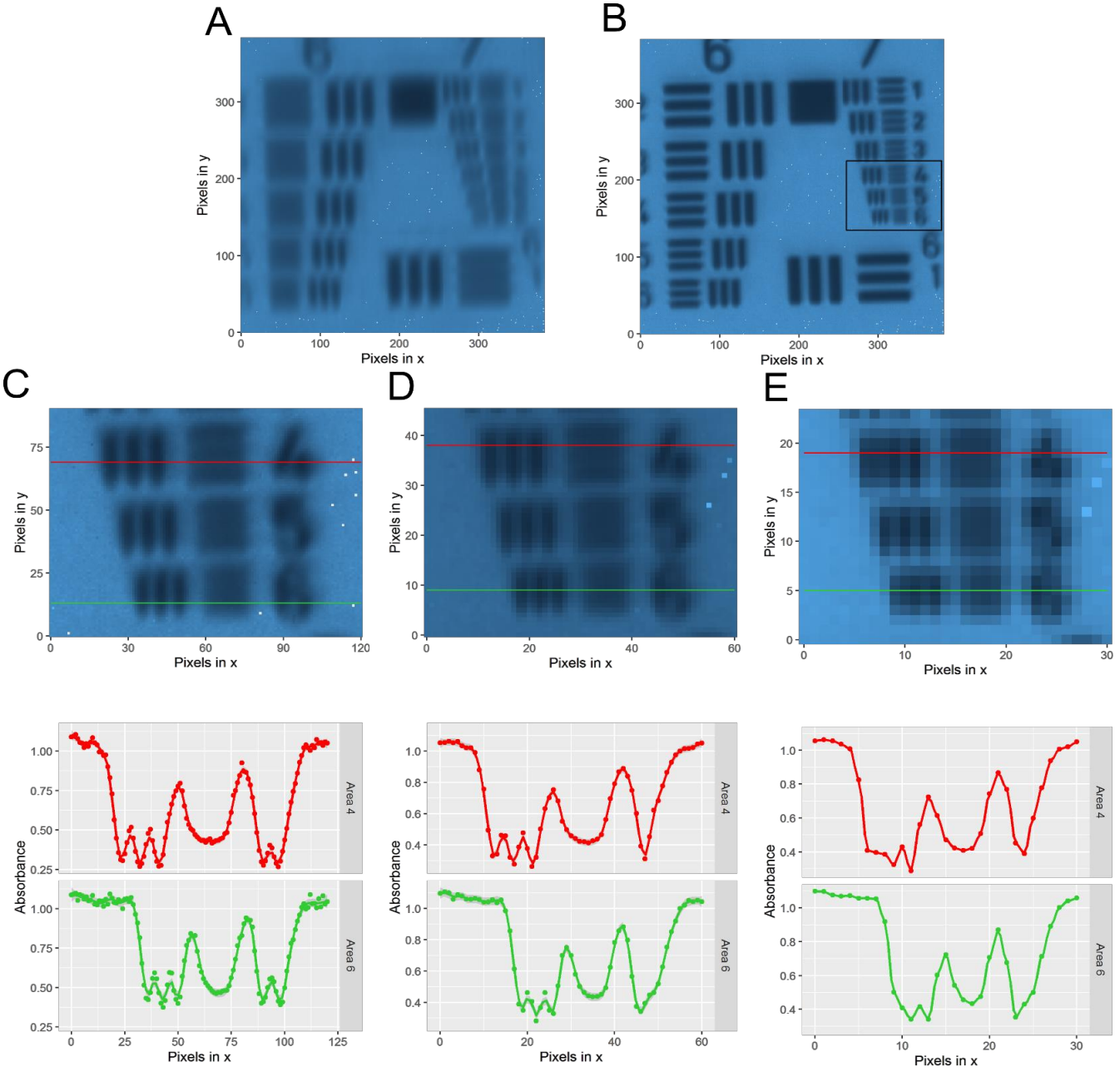
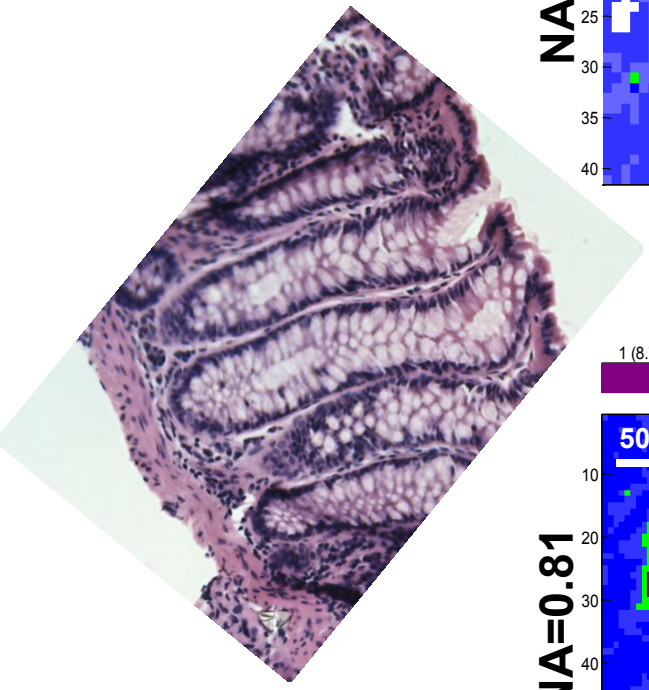


Fig 3



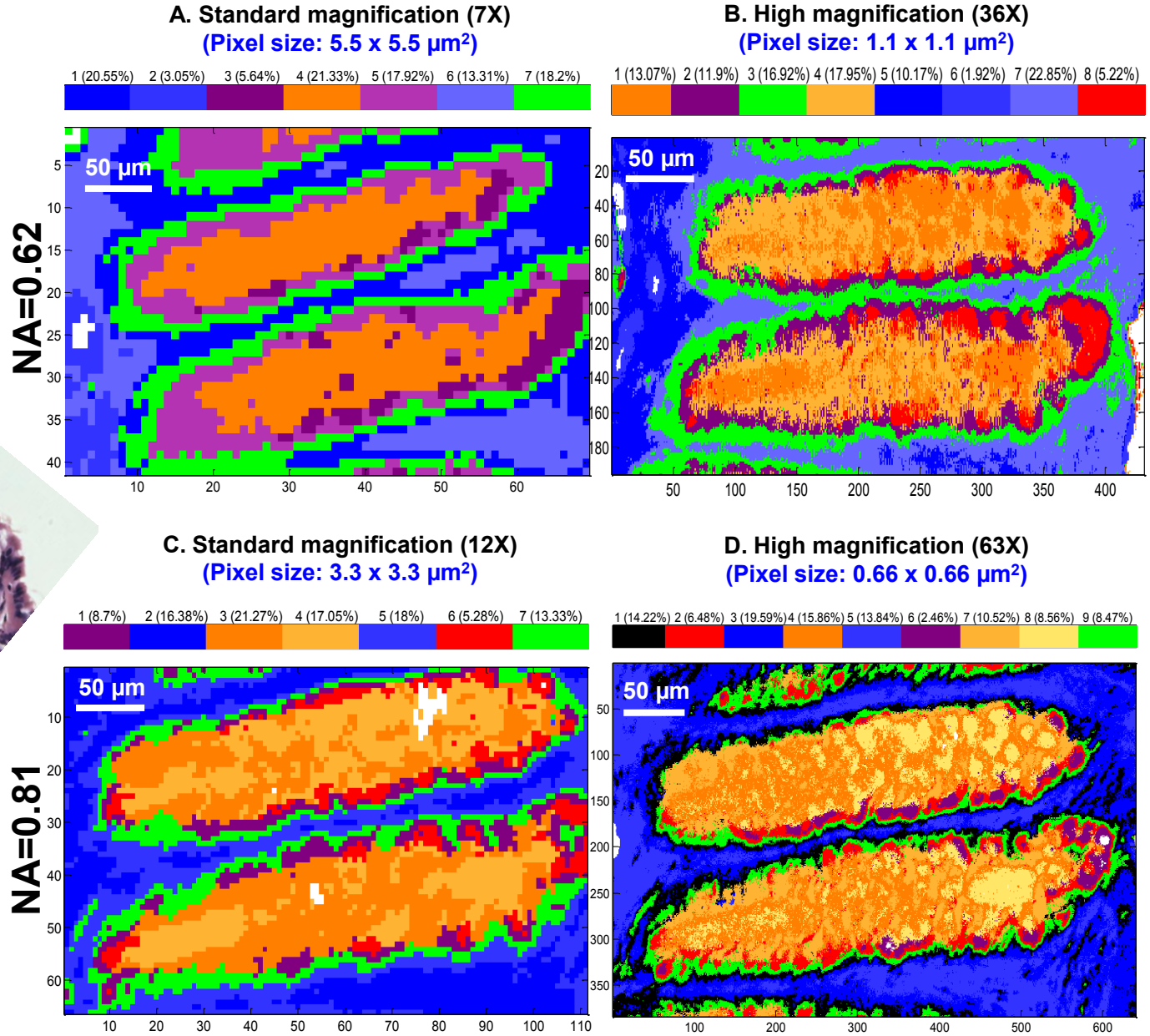
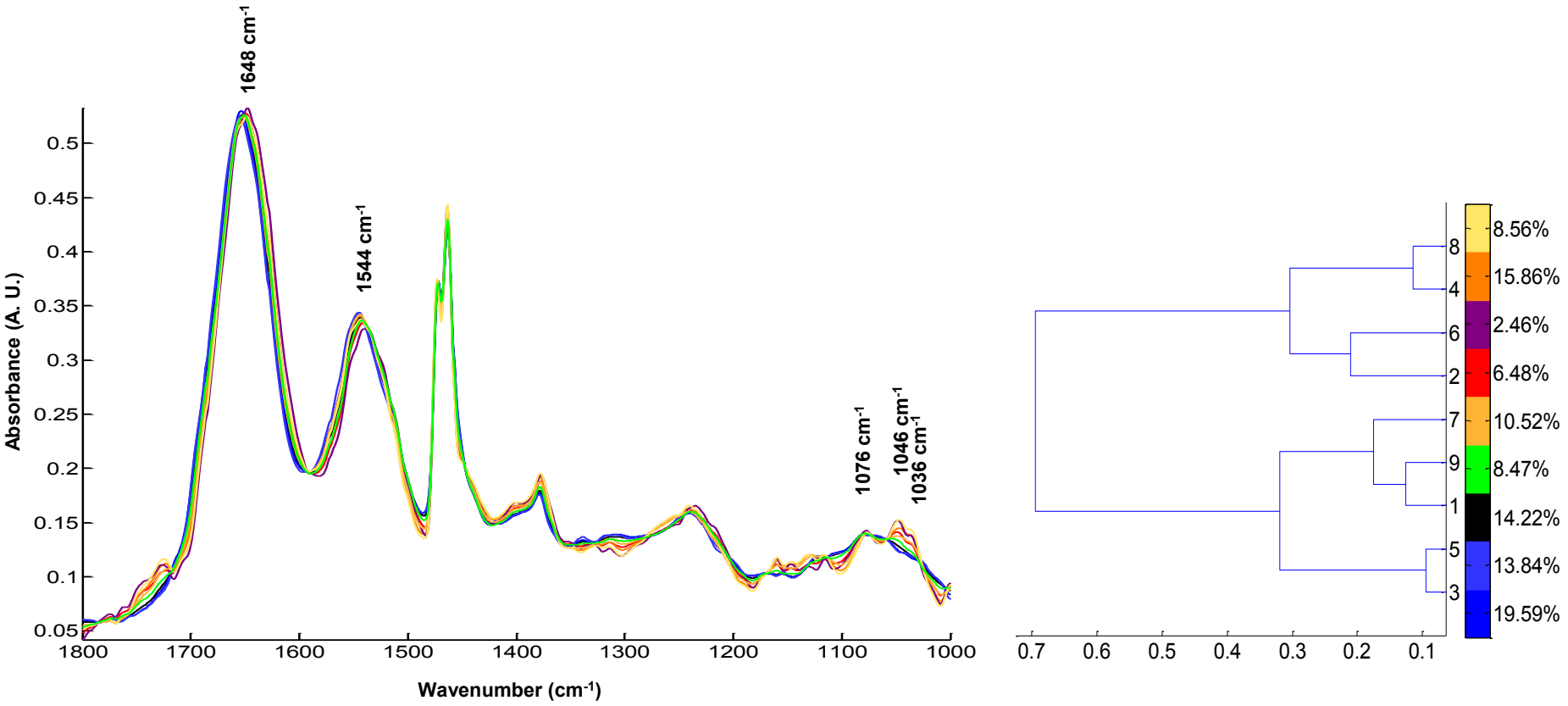

A. Standard magnification (7X)
(Pixel size: 5.5 x 5.5 μm²)
B. High magnification (36X)
(Pixel size: 1.1 x 1.1 μm²)
C. Standard magnification (12X)
(Pixel size: 3.3 x 3.3 μm²)
D. High magnification (63X)
(Pixel size: 0.66 x 0.66 μm²)
NA=0.62
NA=0.81
1 (20.55%) 2 (3.05%) 3 (5.64%) 4 (21.33%) 5 (17.92%) 6 (13.31%) 7 (18.2%)
1 (13.07%) 2 (11.9%) 3 (16.92%) 4 (17.95%) 5 (10.17%) 6 (1.92%) 7 (22.85%) 8 (5.22%)
1 (8.7%) 2 (16.38%) 3 (21.27%) 4 (17.05%) 5 (18%) 6 (5.28%) 7 (13.33%)
1 (14.22%) 2 (6.48%) 3 (19.59%) 4 (15.86%) 5 (13.84%) 6 (2.46%) 7 (10.52%) 8 (8.56%) 9 (8.47%)
50 μm
50 μm
50 μm
50 μm

Fig 4


1648 cm-1
1544 cm-1
1076 cm-1
1046 cm-1
1036 cm-1
Absorbance (A. U.)
Wavenumber (cm-1)
8.56%
15.86%
2.46%
6.48%
10.52%
8.47%
14.22%
13.84%
19.59%

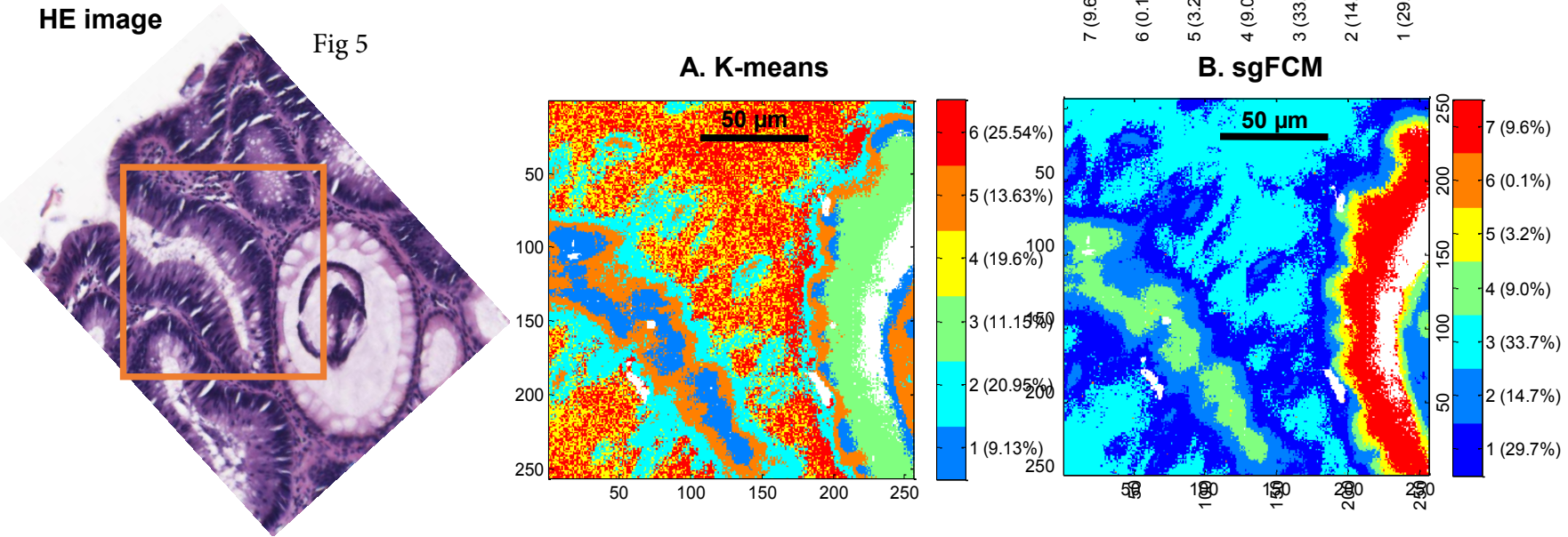
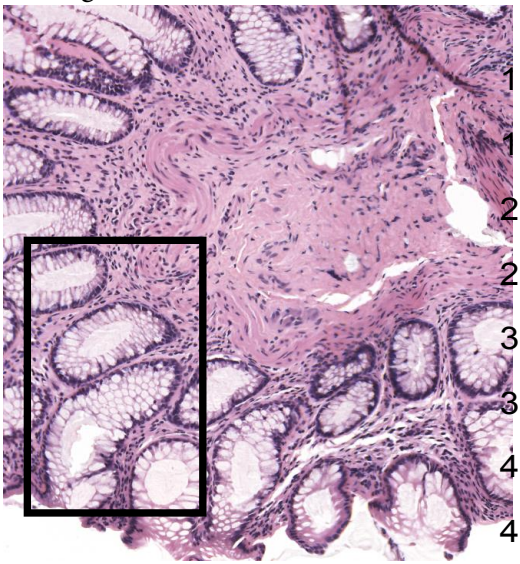

HE image
A. K-means
B. sgFCM
50 μm
50 μm
6 (25.54%)
5 (13.63%)
4 (19.6%)
3 (11.15%)
2 (20.95%)
1 (9.13%)
7 (9.6%)
6 (0.1%)
5 (3.2%)
4 (9.0%)
3 (33.7%)
2 (14.7%)
1 (29.7%)

Fig 5

Fig 6



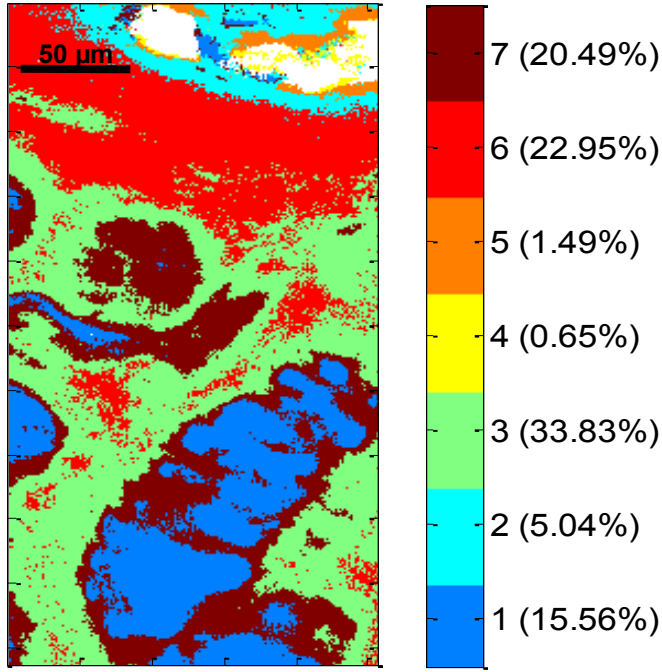

50 µm
50
100
150
200
250
300
350
400
450
500
50 100 150 200 250
7 (20.49%)
6 (22.95%)
5 (1.49%)
4 (0.65%)
3 (33.83%)
2 (5.04%)
1 (15.56%)

Table 1: Achievable pixel sizes from the two IR objectives at different magnifications

	Standard magnification	**High magnification**
Objective 1, NA 0.62	7X (Pixel size: 5.5 x 5.5 μm^2)	36X (Pixel size: 1.1 x 1.1 μm^2)
Objective 2, NA 0.81	12X (Pixel size: 3.3 x 3.3 μm^2)	63X (Pixel size: 0.66 x 0.66 μm^2)