

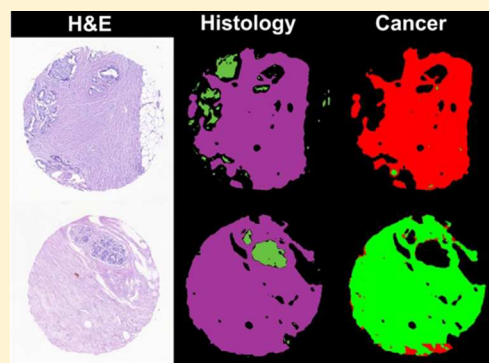
# Quantum Cascade Laser Spectral Histopathology: Breast Cancer Diagnostics Using High Throughput Chemical Imaging

Michael J. Pilling,<sup>1</sup> Alex Henderson, and Peter Gardner\*

Manchester Institute of Biotechnology, University of Manchester, 131 Princess Street, Manchester M1 7DN, U.K.

## Supporting Information

**ABSTRACT:** Fourier transform infrared (FT-IR) microscopy coupled with machine learning approaches has been demonstrated to be a powerful technique for identifying abnormalities in human tissue. The ability to objectively identify the prediseased state and diagnose cancer with high levels of accuracy has the potential to revolutionize current histopathological practice. Despite recent technological advances in FT-IR microscopy, sample throughput and speed of acquisition are key barriers to clinical translation. Wide-field quantum cascade laser (QCL) infrared imaging systems with large focal plane array detectors utilizing discrete frequency imaging have demonstrated that large tissue microarrays (TMA) can be imaged in a matter of minutes. However, this ground breaking technology is still in its infancy, and its applicability for routine disease diagnosis is, as yet, unproven. In light of this, we report on a large study utilizing a breast cancer TMA comprised of 207 different patients. We show that by using QCL imaging with continuous spectra acquired between 912 and 1800  $\text{cm}^{-1}$ , we can accurately differentiate between 4 different histological classes. We demonstrate that we can discriminate between malignant and nonmalignant stroma spectra with high sensitivity (93.56%) and specificity (85.64%) for an independent test set. Finally, we classify each core in the TMA and achieve high diagnostic accuracy on a patient basis with 100% sensitivity and 86.67% specificity. The absence of false negatives reported here opens up the possibility of utilizing high throughput chemical imaging for cancer screening, thereby reducing pathologist workload and improving patient care.



Histopathology, the study of diseased or prediseased tissue, is currently the gold standard in studying the manifestation of disease. Utilizing exogenous stains to highlight tissue architecture and morphology<sup>1</sup> enables visualization of cellular reorganization indicative of disease. However, manual inspection of stained tissue biopsies is a laborious process, and there can be significant delays between acquiring a biopsy and a clinical diagnosis being made. Furthermore, variations in pathologist experience, training, and working practices makes for a subjective diagnosis, which is prone to intra- and interobserver error.<sup>2</sup> Increased throughput and the desire for nonsubjective diagnosis are clear drivers for automated histopathology. However, despite some recent advances in digital pathology such as computer aided diagnosis (CAD),<sup>3</sup> manual inspection of stained tissue sections remains standard practice in clinics worldwide.

The pursuit of automated disease diagnosis has led to the emergence of infrared chemical imaging as a leading candidate with the promise of complementing current histopathological practice. Relying on biochemical information rather than changes in tissue architecture and morphology, chemical imaging enables tissue to be interrogated in a nondestructive and label-free manner. Numerous examples exist in the literature demonstrating its effectiveness in discriminating between normal and cancerous tissue with high levels of sensitivity and specificity.<sup>4–10</sup> Despite this, the technology remains firmly rooted in the research laboratory rather than as

an effective diagnostic tool in the clinic. The problem therein lies with infrared chemical imaging being inherently a low throughput technique. Fourier transform infrared (FTIR) microscopy utilizing a focal plane array (FPA) detector is currently the state-of-the-art in infrared chemical imaging. Exploiting the multiplex advantage of an FPA enables the rapid acquisition of tens of thousands of spectra simultaneously with each spectrum consisting of hundreds of data points. However, a typical tissue microarray often consisting of tens of millions of spectra can take days to acquire and occupy hundreds of gigabytes of storage space.

In principle, increased throughput can be achieved using lower magnification optics such as a 4× magnification objective, enabling a 2.4 × 2.4 mm field of view and a corresponding pixel size of  $\approx 19 \mu\text{m}$ . Utilizing this approach for imaging large areas of laryngeal carcinoma, Beleites<sup>11</sup> demonstrated that acquisition times could be reduced by an order of magnitude compared to using a standard 15× objective. However, this approach increases pixel contamination at cell boundaries,<sup>12</sup> and there is also the risk of missing vital diagnostic information. Bassan previously demonstrated<sup>13,14</sup> that large areas of tissue could be imaged with higher throughput by modifying the acquisition

Received: February 3, 2017

Accepted: June 19, 2017

Published: June 19, 2017

protocol to maximize the duty cycle and reduce dead time. However, ultimately, the trade-off between spatial resolution and throughput means that, at present, chemical images with acceptable spatial resolution are extremely difficult to acquire with an FTIR instrument on a clinically relevant time scale.

Increasing demand for high throughput chemical imaging has led to a renewed interest in discrete frequency infrared spectroscopy. Targeting key frequencies instead of acquiring continuous spectra has the potential to dramatically increase throughput. Studies have shown<sup>15–17</sup> that high classification accuracy can be achieved on tissue using a relatively small number of spectral biomarkers. Several potential technologies have been proposed for discrete frequency imaging, from narrowband infrared filters<sup>18</sup> and optical parametric oscillators<sup>19</sup> to super continuum light sources.<sup>20</sup> Arguably the most promising contender to date is discrete frequency imaging utilizing a tunable, high brightness external cavity infrared quantum cascade laser (QCL). Exploiting the high brightness of a QCL source<sup>21</sup> enables the optical system to be coupled to an uncooled large area microbolometer, thereby allowing large areas of tissue to be imaged with a single measurement. Recently, Bassan<sup>22</sup> demonstrated that an infrared chemical image of an entire TMA consisting of 19 million pixels could be acquired using a single wavelength in just 9 min.

While the increased throughput achievable using discrete frequency imaging is impressive, the majority of studies to date have focused on speed and image quality rather than diagnostic ability.<sup>23</sup> In a limited study, we previously utilized QCL-based discrete frequency imaging for discriminating between normal and malignant prostate epithelium.<sup>24</sup> Utilizing just 25 discrete frequencies enabled 94.60% sensitivity and 93.39% specificity to be achieved on a validation set consisting of 15 patients. However, assessing classifier performance on an independent test set resulted in poorer accuracy (sensitivity 72.14%, specificity 80.23%). Interpatient variability due to using a limited number of training patients is likely to have contributed to the reduced accuracy, but other confounding factors cannot be ruled out.

At the present time, there are several important questions which need to be answered before QCL imaging will be accepted by the biomedical community. One key issue which is yet to be fully explored is the impact of using a coherent source on spectral quality and classification accuracy. Spectra acquired using FTIR microscopes have noise characteristics which are well-understood and signal-to-noise levels which are unparalleled. In contrast, spectra acquired using a QCL imaging instrument coupled to an FPA have poorer signal-to-noise ratios and often exhibit fringes due to interference between the sample and coherent light.<sup>24,25</sup> While it is possible to mitigate coherence effects using single point acquisition,<sup>26,27</sup> this necessarily reduces throughput and negates many of the advantages of QCL imaging. Second, diagnostic accuracy using continuous frequency spectra is yet to be explored, and this is an important preliminary step in assessing this new technology which is still in its infancy. Poor classification accuracy on continuous frequency chemical images would cast doubt on the future application of discrete frequency imaging for disease diagnosis.

To the best of our knowledge, no large scale studies have been performed investigating classification accuracy for a QCL system coupled to an FPA. In light of this, we report on a large breast cancer study involving 207 different patients. Utilizing the fingerprint region between 912 and 1800  $\text{cm}^{-1}$ , we

investigate the diagnostic accuracy of QCL imaging and consider the implications for high throughput, high accuracy disease diagnosis.

## MATERIALS AND METHODS

**Sample Preparation.** Serial sections of formalin fixed paraffin embedded breast tissue microarray (TMA) cores were acquired from US Biomax, Rockville, MD, (TMA ID BR20832). The TMA used in this study consists of 207 1 mm breast tissue biopsy cores, each from a different patient. Pathological review indicated that 15 of the cores are nonmalignant with the remaining 192 being malignant. A 5  $\mu\text{m}$  section was floated onto a standard histology slide, dewaxed, and underwent hematoxylin and eosin (H&E) staining. An adjacent section was floated onto a  $\text{BaF}_2$  slide (Crystran Ltd., Poole, UK) and did not undergo any deparaffinization. Retaining the sample in wax removes the risk of chemical alteration of the sample from clearing solvents and is known to minimize resonant Mie scattering due to refractive index matching.<sup>14</sup>

**Infrared Chemical Imaging.** Infrared chemical images were collected using a Spero Quantum Cascade Laser (QCL) infrared microscope (Daylight Solutions Inc., San Diego, CA, United States). The imaging platform consists of four separately tunable QCL modules enabling unrestricted access to the fingerprint region between 912 and 1800  $\text{cm}^{-1}$ . The microscope is coupled to a room temperature 480  $\times$  480 focal plane array microbolometer, eliminating the requirement of cryogenic cooling and enabling continuous operation. Chemical images were acquired in transmission mode using the 4 $\times$  0.15 NA low magnification objective with a resultant field of view of approximately 2.02  $\times$  2.02 mm and a corresponding nominal pixel size of approximately 4.2  $\mu\text{m}$ .

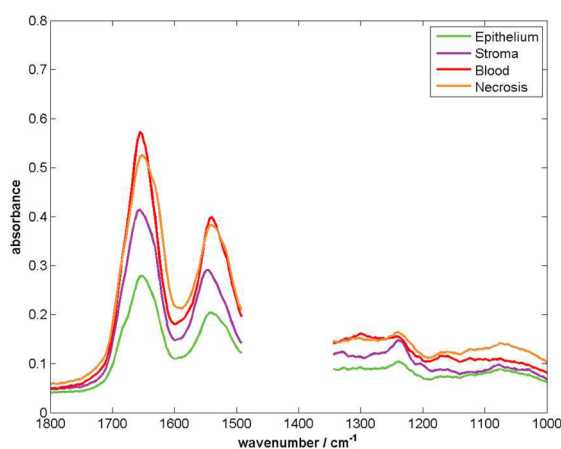
Prior to imaging, a background was acquired as a single tile from an area of the slide which had been identified as being tissue and paraffin free. Chemical images were collected in the spectral range 912–1800  $\text{cm}^{-1}$ , utilizing a step size of 4  $\text{cm}^{-1}$  to produce continuous frequency spectra. Each infrared tile consisted of 230 400 spectra, was comprised of 223 data points, and took 5 min 45 s to collect.

A chemical image of the entire TMA was collected as an 11  $\times$  13 mosaic consisting of 143 tiles and 33 million pixels acquired over 13.6 h.

**Data Preprocessing.** All data were preprocessed with MATLAB 2014a (The MathWorks Inc., Natick, MA, United States) using functions written in house. Infrared spectra for each biopsy core were extracted from the mosaic as a 313  $\times$  313  $\times$  223 datacube, consisting of 97 969 spectra, each with 223 data points. Spectra were quality tested to remove data obtained from areas with little or no tissue using the height of the amide I band with spectra having absorbance between 0.1 and 2 being retained. Principal component-based noise reduction was used to improve signal-to-noise with the first 40 PCs being retained. Spectra were truncated between 1000 and 1800  $\text{cm}^{-1}$ , and the region describing the absorption bands of wax (1350–1490  $\text{cm}^{-1}$ ) were removed. Each spectrum was then vector normalized to correct for different thicknesses of tissue and finally converted to its first derivative while performing Savitzky–Golay smoothing using a window size of nine data points.

## RESULTS

**QCL Chemical Imaging: Automated Histology.** Chemical images of each of the breast tissue cores were compared to the H&E stained sections, and regions of epithelium, stroma, blood, and necrosis were identified. Using the methods of Fernandez,<sup>28</sup> a spectral database was constructed from 74 cores (61 malignant and 13 nonmalignant) consisting of 171 610 epithelium, 111 960 stroma, 4431 blood, and 27 700 necrosis spectra. Mean spectra, following quality testing and noise reduction and prior to vector normalization and derivitization, for each of the histological classes are shown in Figure 1. Inspection of each mean spectrum clearly discerns significant differences in band intensity and position for each of the histological classes.



**Figure 1.** Mean spectra of epithelium, stroma, blood, and necrosis obtained using continuous frequency acquisition. The spectra were quality tested and noise reduced prior to calculating the mean. The spectral band 1350–1490  $\text{cm}^{-1}$  contaminated with paraffin response was removed for clarity.

The 74 cores were then randomly separated into training and testing cores using an 80:20 split consisting of 59 training and 15 test cores. The training and test cores used are detailed in Table S-1. Separating the cores into training and testing prior to constructing the classifier ensures that the test set is from different patients to those in the training set and therefore completely independent. A training database was then constructed from the 59 training cores which consisted of 145 386 epithelium, 88 205 stroma, 1766 blood, and 25 382 necrosis spectra. The optimal situation to prevent classifier bias is for each class to consist of equal numbers of spectra. However, this was not possible due to only a limited number of cores containing blood or necrosis. Bias was minimized in the training set by selecting equal numbers of spectra from each class with the number randomly selected being the size of the smallest class.

A Random Forest classifier<sup>29</sup> (code available from <http://code.google.com/p/randomforest-matlab/>) was then constructed using 1766 spectra per class. Each spectrum was quality checked, noise reduced, vector normalized, and derivitized prior to being used for training. Five hundred trees were used during the construction of the classifier with the number of variables selected at random to try to split each node set to 15. The node size parameter, which limits how large each decision tree can grow, was set to 10. The Random Forest classifier was then tested on the independent test set which

consisted of 26 224 epithelium, 42 602 stroma, 2665 blood, and 2318 necrosis spectra from the 15 testing cores.

Each tree in the Random Forest “votes” for the class which it predicts an unknown spectrum belongs to. The Random Forest then chooses the class having the most votes over all the trees in the forest (“majority rules”). Misclassification can occur when similar votes are cast for each class or where there is only a slim majority. Setting an acceptance threshold allows the classifier to reject pixels when there is poor agreement on class membership between the trees.

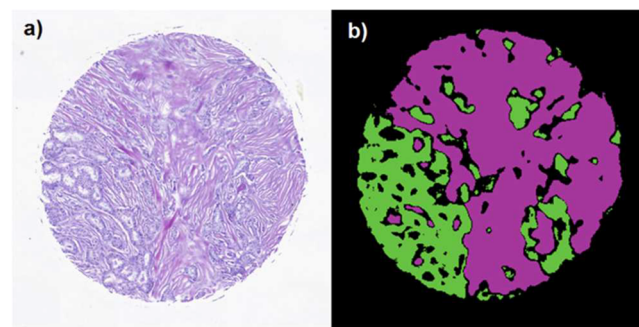
Confusion matrices provide a quantitative measure of performance and enable the correctness of classification for each class to be determined. Furthermore, the sources of misclassification can be easily identified from a confusion matrix, revealing which classes are difficult to discriminate between. Table 1 shows the confusion matrix for the four-class

**Table 1.** Confusion Matrix Showing Classification Accuracy As a Percentage Using an Acceptance Threshold of 0.6

true class	predicted class			
	epithelium	stroma	blood	necrosis
epithelium	97.02	2.97	0.01	0
stroma	5.60	94.33	0.05	0.03
blood	5.06	0	94.94	0
necrosis	0.67	0	0	99.33

system using an acceptance threshold of 0.6 (i.e., at least 60% of trees must agree on class membership). Correctness of classification rates are shown by the diagonal of the table with all four classes having accuracies >94%. Bassan<sup>15</sup> previously reported on FTIR imaging of a serial section of the same TMA on glass utilizing the amide A band for discriminating between the same four tissue types. Classification accuracies reported by Bassan (epithelium = 98.25%, stroma = 99.94%, blood = 100.00%, necrosis = 97.22%) compare favorably with the results presented here.

The Random Forest classifier was then used to classify each of the  $\approx 8$  million pixels within the TMA chemical image. Assigning each class a color allows rendering of a false color image, which provides a visual representation of the classified image. Figure 2a shows a high resolution brightfield image of a mixed core (core A5) consisting primarily of epithelium and stroma. Comparison of the classified image (Figure 2b) to the brightfield image of the H&E stained section illustrates that there is good agreement for each class. A high resolution



**Figure 2.** (a) Brightfield image of H&E stained serial section for core A5 and (b) false color image of the classified core rendered using green (epithelium) and purple (stroma).

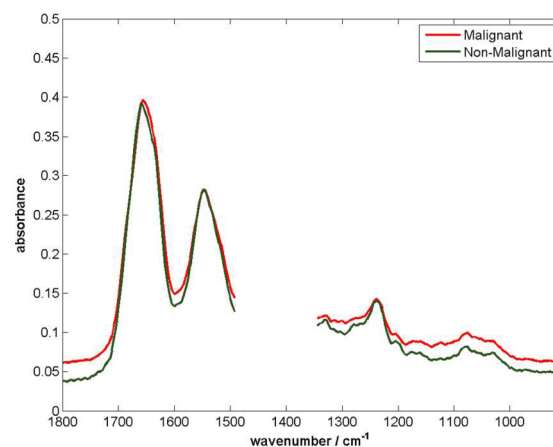


brightfield H&E image of the entire TMA and the corresponding classified image are shown in Figures S1 and S2. In the majority of cores, there is good agreement between the H&E and the classified image. Several classified cores are observed to have only a limited number of pixels. An extreme case of this is observed in the classified image of the last core of the third row, where the majority of pixels are unclassified due to lack of agreement between the trees on class membership. In situations where the core cannot be satisfactorily classified, then this would be flagged as a suspect core which would require expert pathological review.

**QCL Chemical Imaging: Cancer Diagnostics.** Ultimately, QCL chemical imaging will be adopted by clinicians only if it can add value to current methods and lead to improvements in patient treatment and care. Because pathological review is currently the gold standard, improved diagnostic accuracy is an unrealistic target and, in the absence of longitudinal studies, difficult to prove. QCL chemical imaging needs to be competitive with current diagnostic capabilities while being capable of high throughput. At the present time, no large scale studies have explored diagnostic accuracy on tissue for high throughput QCL imaging with an FPA. To address this, we investigate the discriminatory power of QCL imaging for segmenting malignant and nonmalignant tissue from a large number of patients.

Numerous examples exist in the literature of the discriminatory power of infrared chemical imaging for differentiating between normal and cancerous epithelium. Recent studies<sup>30,31</sup> have questioned this approach and suggest that stroma, specifically adjacent stroma, may have a key role to play in the initiation and progression of cancer. Recently, Pounder<sup>32</sup> reported on the difficulties associated with utilizing breast epithelium pixel spectra for differentiating between normal and cancerous tissue. Using 8 spectral metrics on the training data, they produced an receiver operator characteristics (ROC) curve with a cancer pixel level AUC of just 0.81. Furthermore, because nonmalignant breast tissue is typically not particularly glandular, this limits the number of epithelium spectra which can be used for training and testing the classifier. In light of this, we elected to use stroma rather than epithelium to explore the diagnostic capabilities of QCL chemical imaging.

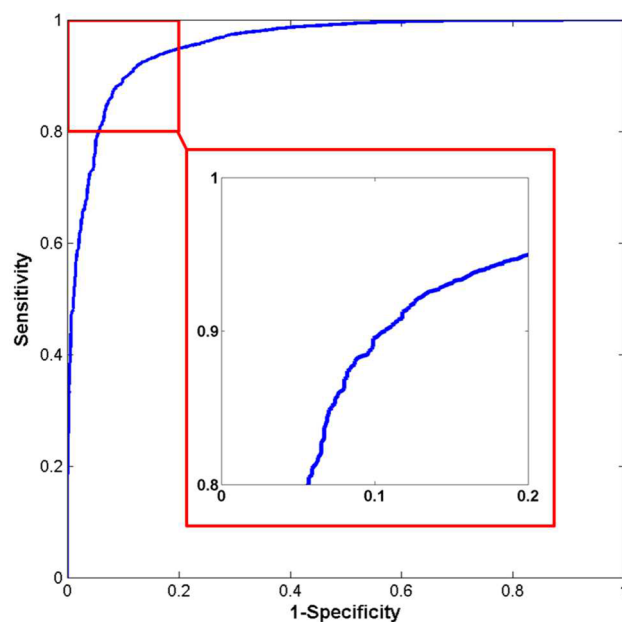
Stromal spectra were isolated from each of the 207 cores using the Random Forest classifier to remove all pixels belonging to any other class. Cores were then randomly split into five subsets using fivefold cross-validation with one subset to be used as an independent test and the remainder as a training set. To improve data handling and computation times, 1000 spectra were randomly selected from each core, and a training database was constructed. The training database for each repeat typically consisted of approximately 140 000 malignant and 13 000 nonmalignant spectra from 166 training cores. Each stroma spectrum was then labeled as either malignant or nonmalignant which was dependent on whether it came from a malignant or nonmalignant core. We elected not to distinguish between adjacent and distal stroma because this requires manual annotation of each individual stromal pixel, which can be cumbersome. Figure 3 shows the mean spectra obtained following quality testing and noise reduction for malignant and nonmalignant stroma. In contrast to the mean spectra obtained for histology (Figure 1), the stroma spectra are similar, and there does not appear to be significant differences in peak intensity and position for each class.



**Figure 3.** Mean spectra of malignant and nonmalignant stroma. The spectra were quality tested and noise reduced prior to calculating the mean. The spectral band at 1350–1490  $\text{cm}^{-1}$  contaminated with paraffin response was removed for clarity.

A Random Forest classifier was constructed from the training database with bias minimized by using all the nonmalignant spectra and an equal number of randomly selected malignant spectra. The classifier was constructed using 500 trees, and with the number of variables used to attempt to split each node, set to 15. The node size parameter which limits how large each tree can grow was set to 1. The independent test set for each of the 5 repeats typically consisted of approximately 36 000 malignant and 2800 nonmalignant spectra from 41 testing cores.

The Random Forest classifier output provides an estimate of the probability that a spectrum belongs to a particular class. Adjusting the class probability threshold enables the trade off between true positives and false positives to be visualized in the form of ROC curves. A typical ROC curve for classification of the independent test set is shown in Figure 4. Considering the similarities between malignant and nonmalignant stromal spectra, the ROC curves appear surprisingly good. The



**Figure 4.** ROC plot obtained using Random Forest to classify the independent test set for stroma. (AUC = 0.9582).

resulting AUC value of 0.9582 indicates that there is good differentiation between malignant and nonmalignant pixels.

The AUC values for the independent test sets for each of the five repeats are shown in Table 2. All values are consistently

**Table 2. AUC Values for Malignant Stroma for the Independent Test Sets Based on the Five Repeats Using Different Sets of Test Patients**

run	1	2	3	4	5
AUC	0.9582	0.9184	0.9581	0.8972	0.9466

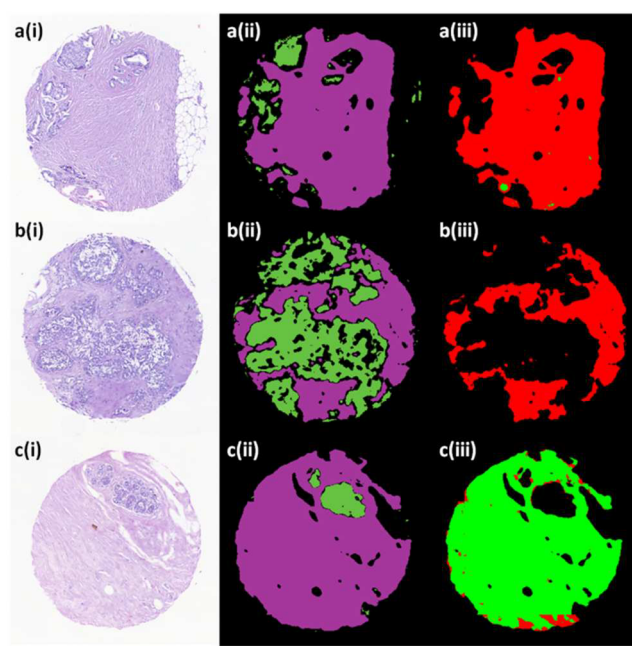
high, indicating that the choice of training and test patients has only a minimal effect on classifier performance. Finally, a confusion matrix was constructed for all independent test set pixels for the five repeats. An acceptance threshold of 0.6 was used, which allowed pixels to be rejected when fewer than 60% of trees agree on the predicted class while retaining approximately 90% of all stroma spectra. The resulting confusion matrix in Table 3 reveals that 93.56% of malignant stroma and 85.64% of nonmalignant stroma are correctly classified.

**Table 3. Confusion Matrix Obtained for All Independent Test Set Pixels Using an Acceptance Threshold of 0.6**

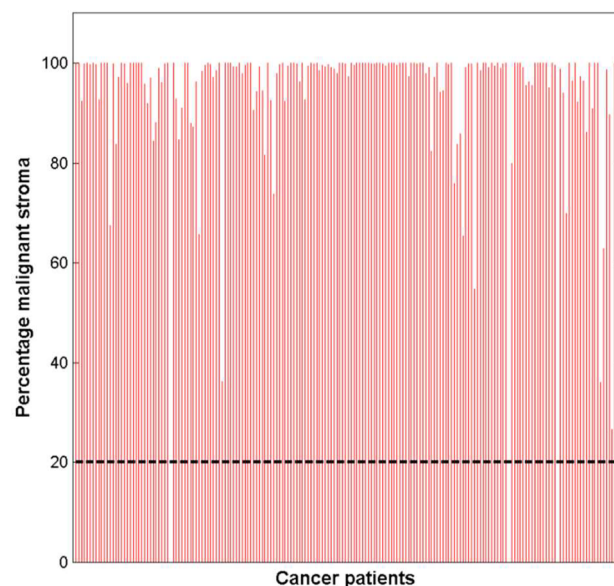
true class	predicted class	
	malignant (%)	nonmalignant (%)
malignant	93.56	6.44
nonmalignant	14.36	85.64

**Automated Histopathology: Patient Cancer Diagnosis.** Accurate discrimination between malignant and nonmalignant spectra is an important proof of concept for QCL chemical imaging. However, practical application of QCL imaging in the clinic will require accurate differentiation between malignant and nonmalignant biopsy cores. We assess the potential for QCL automated patient diagnostics by subjecting each stromal pixel within each core to the Random Forest classifier. Figures 5a(i)–c(i) show the brightfield images of the H&E stained sections for two malignant and one nonmalignant cores. A comparison of these images to Figures 5a(ii)–c(ii) representing the histology demonstrates that there is good agreement between the histologically classified image and the H&E. Finally, all nonstromal pixels are removed, and the cancer classifier is used to assign each stromal pixel as either malignant or nonmalignant. Rendering malignant stromal pixels red and nonmalignant green enables a false color image to be formed. Figures 5a(iii) and b(iii) and Figure 5c(iii) show the classification result for the two malignant and the nonmalignant core, respectively. The malignant cores are dominated by red pixels, indicating that the classifier can accurately identify malignant stromal pixels. The nonmalignant core (Figure 5c(iii)) is predominantly green, indicating that the classifier correctly classifies most stromal pixels as being nonmalignant. A small proportion of pixels at the edges of the core were classified as malignant (red), but the vast majority of pixels were classified as nonmalignant.

A bar chart displaying the proportion of stroma pixels classified as malignant for each of the malignant cores is displayed in Figure 6. The bar chart shows that for nearly all patients the proportion of malignant stroma is close to 1. Out of 192 malignant cores, four did not have any identifiable

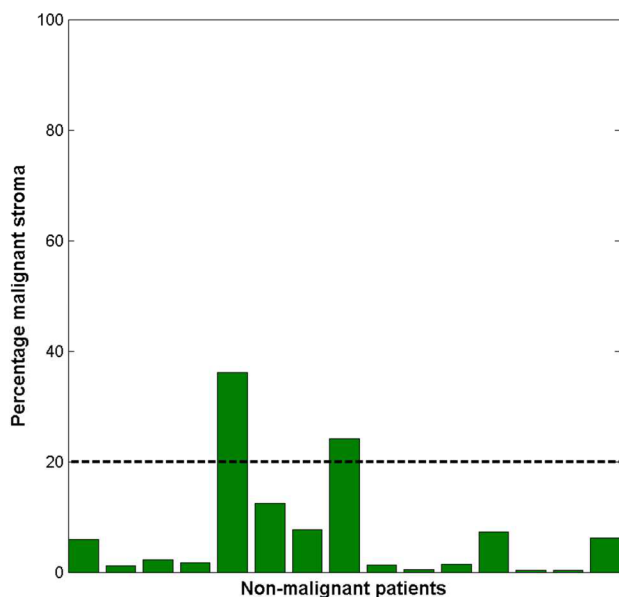


**Figure 5.** a–c(i) Brightfield image of H&E stained section of two malignant and one nonmalignant core, a–c(ii) false color image of histology rendered using green (epithelium) and purple (stroma) to illustrate predicted histological class, and a–c(iii) false color image of malignancy rendered using green (nonmalignant stroma) and red (malignant stroma).



**Figure 6.** Bar chart showing malignant stroma pixels as a percentage of total stroma pixels for each of the 192 cancer patients. The dashed line represents a threshold of 0.2.

stroma (as determined by Random Forest) and could not be classified. The mean proportion of malignant stroma for the remaining malignant cores is 95.3%, suggesting stroma is an effective indicator of malignancy. Figure 7 shows the resulting bar chart for the nonmalignant cores and there are striking differences compared to the bar chart obtained for the malignant cores. Two of the nonmalignant cores had high levels of malignant stroma (36.17 and 24.12%), and these appear to have been misclassified. Given the limited number of



**Figure 7.** Bar chart showing malignant stroma pixels as a percentage of total stroma pixels for each of the nonmalignant core patients. The dashed line represents a threshold of 0.2, which enables 86.7% correctness of classification of the nonmalignant cores.

nonmalignant cores available on the TMA, it is likely that the classifier did not include sufficient interpatient variability to classify the pixels within these two cores with high accuracy. However, the mean malignant stroma proportion is just 7.25% compared to 95.3% for the malignant cores, which suggests that we can diagnose cancer by choosing an appropriate threshold which we consider indicates malignancy. Utilizing a threshold of 20% (0.2), Fernandez<sup>28</sup> demonstrated that, in the case of prostate cancer, highly accurate segmentation between patient-matched benign and malignant epithelium could be achieved.

Applying a threshold of 0.2 to the bar chart in Figure 6 enables 100% accurate classification of all cores which had identifiable stroma. Recall that 4 cores out of 192 did not contain any stroma and could not be classified. It is important to note that there were no false negatives (malignant diagnosed as nonmalignant), which is a key deliverable for cancer screening. In the nonmalignant case, all cores had identifiable stroma, but there were 2 false positives (nonmalignant diagnosed as malignant), resulting in 86.67% specificity. For cancer screening, this would be an acceptable level because all abnormal or suspect samples would be reviewed by a pathologist to make a final diagnosis.

## DISCUSSION

Infrared chemical imaging has the potential to provide complementary information to aid diagnosis, leading to improved patient treatment and care. Despite significant progress over the past decade, sample throughput is still a key barrier to clinical translation.

QCL chemical imaging has emerged as a leading candidate for high throughput infrared chemical imaging with a diagnostic window compatible with clinical time scales. Current investigations on QCL imaging for diagnostics are broadly in two directions: focal plane array detection and single point detection. QCL-based systems utilizing a large ( $480 \times 480$ ) FPA allow the simultaneous acquisition of 230 400 spectra, thereby imaging large areas of tissue quickly and at high spatial

resolution. Recently, doubts have been expressed as to the veracity of the acquired data due to the poorly understood impact of coherence effects. Some authors believe single point acquisition mitigates, to some extent, coherence effects through operating confocally.<sup>26</sup> However, the clear disadvantage of single point collection is that the sample has to be rastered at every single point to image a given area of tissue. Nevertheless, if a limited number of discrete frequencies are required, then single point acquisition can achieve acquisition times similar to those of FPA imaging. Tiwari<sup>27</sup> previously utilized a single point QCL infrared microscope to image a  $2.5 \text{ mm}^2$  area of tissue. Utilizing a single frequency enabled the full area to be acquired with a  $5 \mu\text{m}$  pixel size in approximately 15 s. However, acquiring the same area using the fully accessible spectral range ( $800\text{--}1800 \text{ cm}^{-1}$ ) increased the acquisition time to approximately 50 min. For comparison, current FPA-based systems would measure  $4 \text{ mm}^2$  at a single frequency, with  $4.3 \mu\text{m}$  pixel size in approximately 5 s, and a continuous frequency spectrum ( $912\text{--}1800 \text{ cm}^{-1}$ ) in 5.75 min. Given the throughput advantage of an FPA-based QCL microscope, a key question is the impact of coherence on spectral quality and classification accuracy. If diagnostic accuracy can be competitive with the current gold standard, i.e., the pathologist, then FPA-based QCL imaging has a significant throughput advantage over single point imaging.

In this study, we investigated the potential of QCL (FPA) chemical imaging for disease diagnosis. We demonstrated that rapid QCL chemical imaging using continuous frequency spectra can accurately discriminate between four histological classes in breast tissue. Moreover, we showed that we can easily discriminate between malignant and nonmalignant stroma spectra with 93.56% sensitivity and 85.64% specificity on an independent test set. These excellent rates of correct classification suggest that coherence effects have little or no impact on classification accuracy. On a patient basis, the 188 out of 192 malignant cores, which could be classified, were classified with a sensitivity of 100%. Importantly, there were no cases where a malignant core was classified as nonmalignant. Nonmalignant patients were classified correctly with an accuracy of 86.67%. These results compare favorably to those reported by Sattlecker<sup>33</sup> for ensemble support vector machine (SVM) breast cancer type prediction using microcalcification spectra, where 87.5% sensitivity and 75% specificity was obtained.

The high sensitivity reported here opens up the possibility of high throughput automated screening, enabling biopsies to be triaged for pathological review. Screening requires that all cancer cases are detected because any missed cases will result in patient under-treatment. The 86.67% specificity reported here is acceptable because the 13.33% of nonmalignant cores misclassified would be reviewed by a pathologist to rule out cancer. The key advantage of screening with this system is that 86.67% of the nonmalignant cores that were correctly classified would not need to be reviewed by a pathologist. Clearly, there are significant gains to be made with high throughput screening, considering that approximately a quarter of all breast tissue biopsies are benign but still undergo pathological review.<sup>34</sup>

The preliminary results presented here demonstrate that high classification accuracy of malignancy is achievable when using a single TMA for training and testing the classifier. While promising, we appreciate that training and testing on separate TMAs would be a more robust demonstration of the potential of QCL imaging for disease diagnosis. Utilizing multiple TMAs



could potentially introduce confounding factors such as the influence of the substrate, different thickness of tissue, and variations in focus. Therefore, we believe that, although promising, further work needs to be performed using a larger number of patients over several separate TMAs to fully demonstrate the potential of the technology for disease diagnosis.

The key question that arises from this study relates to the speed of continuous frequency imaging and whether it provides sufficient throughput to be clinically viable. Rapid throughput has always been the key advantage of discrete frequency chemical imaging. While this is true, we believe that high throughput chemical imaging is achievable using continuous frequency spectra. Utilizing a breast cancer tissue microarray, we analyzed 207 breast tissue biopsy cores in just 13.6 h, which is equivalent to a core being acquired on average every 3 min 56 s. Even higher throughput could have been achieved by optimizing the acquired spectral range. In this study, we elected to acquire the entire fingerprint region between 912 and 1800  $\text{cm}^{-1}$ . Given the presence of wax bands between 1350 and 1490  $\text{cm}^{-1}$  and the limited biochemical information between 912 and 1000  $\text{cm}^{-1}$ , a more intelligent approach would be to omit these regions. We calculate that acquiring the range 1000–1350 and 1490–1800  $\text{cm}^{-1}$  would require 168 discrete frequencies instead of 223. Assuming a linear relationship between acquisition time and discrete frequencies, we estimate the entire TMA could have been acquired in approximately 10.25 h instead of 13.6 h. Optimizing the acquisition parameters would translate to each core being acquired in approximately 2 min 58 s, which is a time scale likely to be acceptable to clinicians.

## CONCLUSIONS

In this study, we showed that FPA-based QCL imaging using continuous frequency spectra enables highly accurate discrimination between malignant and nonmalignant stroma. We further showed that we can use high throughput automated histopathology to accurately diagnose biopsy cores on a patient basis. Coherence effects associated with FPA-based QCL imaging appear to have only a minimal effect on classification accuracy because the sensitivity and specificity reported here are in broad agreement with similar studies using FTIR imaging. The results reported here pave the way for FPA-based discrete frequency imaging for disease diagnosis. Future studies of FPA-based discrete frequency imaging using a large number of patients are required to assess its full potential for high throughput automated histopathology.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.7b00426.

High resolution brightfield image of H&E stained TMA section, corresponding histology classified image, and BR20832 TMA specification sheet including patient age, pathologists' assessment on malignancy, identification of cores used for training and testing of histology classifier, and correctness of diagnosis using QCL IR imaging (PDF)

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: peter.gardner@manchester.ac.uk; Phone: +44 (0) 161 306 4463.

### ORCID

Michael J. Pilling: 0000-0003-1634-3326

### Notes

The authors declare no competing financial interest. Additional research data supporting this publication are available from the Zenodo repository at DOI: 10.5281/zenodo.808456.

## ACKNOWLEDGMENTS

P.G., A.H., and M.J.P. would like to acknowledge the EPSRC (Grants EP/K02311X/1 and EP/L012952/1). The authors would also like to acknowledge Daylight Solutions for the loan of the Spero for the duration of this study.

## REFERENCES

- (1) Luna, L. G. *Manual of histologic staining methods of the Armed Forces Institute of Pathology*; Blakiston Division, McGraw-Hill: New York, 1968.
- (2) Lattouf, J. B.; Saad, F. *BJU Int.* **2002**, *90*, 694–698.
- (3) Bahlmann, C.; Patel, A.; Johnson, J.; Ni, J.; Chekkoury, A.; Khurd, P.; Kamen, A.; Grady, L.; Krupinski, E.; Graham, A.; et al. *Proc. SPIE* **2012**, *8315*, 831504.
- (4) Bird, B.; Miljkovic, M.; Remiszewski, S.; Akalin, A.; Kon, M.; Diem, M. *Lab. Invest.* **2012**, *92*, 1358–1373.
- (5) Baker, M. J.; Gazi, E.; Brown, M. D.; Shanks, J. H.; Clarke, N. W.; Gardner, P. *J. Biophotonics* **2009**, *2*, 104–113.
- (6) Bergner, N.; Romeike, B. F. M.; Reichart, R.; Kalff, R.; Krafft, C.; Popp, J. *Analyst* **2013**, *138*, 3983–3990.
- (7) Kallenbach-Thieltges, A.; Großerüschkamp, F.; Mosig, A.; Diem, M.; Tannapfel, A.; Gerwert, K. *J. Biophotonics* **2013**, *6*, 88–100.
- (8) Kochan, K.; Heraud, P.; Kiupel, M.; Yuzbasiyan-Gurkan, V.; McNaughton, D.; Baranska, M.; Wood, B. R. *Analyst* **2015**, *140*, 2402–2411.
- (9) Hughes, C.; Iqbal-Wahid, J.; Brown, M.; Shanks, J. H.; Eustace, A.; Denley, H.; Hoskin, P. J.; West, C.; Clarke, N. W.; Gardner, P. *J. Biophotonics* **2013**, *6*, 73–87.
- (10) Kwak, J. T.; Hewitt, S. M.; Sinha, S.; Bhargava, R. *BMC Cancer* **2011**, *11*, 62.
- (11) Beleites, O. G.-L. C.; Ernst, G.; Popp, J.; Krafft, C. *Biomed. Spectrosc. Imaging* **2015**, *4*, 57–66.
- (12) Bhargava, R. *Anal. Bioanal. Chem.* **2007**, *389*, 1155–1169.
- (13) Bassan, P.; Sachdeva, A.; Shanks, J. H.; Brown, M. D.; Clarke, N. W.; Gardner, P. *Analyst* **2013**, *138*, 7066–7069.
- (14) Bassan, P.; Sachdeva, A.; Shanks, J. H.; Brown, M. D.; Clarke, N. W.; Gardner, P. *Proc. SPIE* **2014**, 90410D.
- (15) Bassan, P.; Mellor, J.; Shapiro, J.; Williams, K. J.; Lisanti, M. P.; Gardner, P. *Anal. Chem.* **2014**, *86*, 1648–1653.
- (16) Lloyd, G. R.; Stone, N. *Appl. Spectrosc.* **2015**, *69*, 1066–1073.
- (17) Mittal, S.; Wrobel, T. P.; Leslie, L. S.; Kadjacsy-Balla, A.; Bhargava, R. *Proc. SPIE* **2016**, 979118.
- (18) Kodali, A. K.; Schulmerich, M.; Ip, J.; Yen, G.; Cunningham, B. T.; Bhargava, R. *Anal. Chem.* **2010**, *82*, 5697–5706.
- (19) Amrania, H.; McCrow, A. P.; Matthews, M. R.; Kazarian, S. G.; Kuimova, M. K.; Phillips, C. C. *Chem. Sci.* **2011**, *2*, 107–111.
- (20) Ishibe, K.; Nakada, S.; Mera, Y.; Maeda, K. *Microsc. Microanal.* **2012**, *18*, S91–S95.
- (21) Faist, J.; Capasso, F.; Sivco, D. L.; Sirtori, C.; Hutchinson, A. L.; Cho, A. Y. *Science* **1994**, *264*, 553.
- (22) Bassan, P.; Weida, M. J.; Rowlette, J.; Gardner, P. *Analyst* **2014**, *139*, 3856–3859.
- (23) Sreedhar, H.; Varma, V. K.; Gambacorta, F. V.; Guzman, G.; Walsh, M. J. *Biomed. Opt. Express* **2016**, *7*, 2419–2424.

- (24) Pilling, M. J.; Henderson, A.; Bird, B.; Brown, M. D.; Clarke, N. W.; Gardner, P. *Faraday Discuss.* **2016**, *187*, 135–154.
- (25) Kole, M. R.; Reddy, R. K.; Schulmerich, M. V.; Gelber, M. K.; Bhargava, R. *Anal. Chem.* **2012**, *84*, 10366–10372.
- (26) Yeh, K.; Bhargava, R. *Biomedical Vibrational Spectroscopy 2016: Advances in Research and Industry*; Proc. SPIE: San Francisco, California, 2016; Vol. 9704.
- (27) Tiwari, S.; Raman, J.; Reddy, V.; Ghetler, A.; Tella, R. P.; Han, Y.; Moon, C. R.; Hoke, C. D.; Bhargava, R. *Anal. Chem.* **2016**, *88*, 10183–10190.
- (28) Fernandez, D. C.; Bhargava, R.; Hewitt, S. M.; Levin, I. W. *Nat. Biotechnol.* **2005**, *23*, 469–474.
- (29) Breiman, L. *Machine Learning* **2001**, *45*, 5–32.
- (30) Kumar, S.; Desmedt, C.; Larsimont, D.; Sotiriou, C.; Goormaghtigh, E. *Analyst* **2013**, *138*, 4058–4065.
- (31) Kong, R.; Reddy, R. K.; Bhargava, R. *Analyst* **2010**, *135*, 1569–1578.
- (32) Pounder, F. N.; Reddy, R. K.; Bhargava, R. *Faraday Discuss.* **2016**, *187*, 43–68.
- (33) Sattlecker, M.; Baker, R.; Stone, N.; Bessant, C. *Chemom. Intell. Lab. Syst.* **2011**, *107*, 363–370.
- (34) Weaver, D. L.; Rosenberg, R. D.; Barlow, W. E.; Ichikawa, L.; Carney, P. A.; Kerlikowske, K.; Buist, D. S. M.; Geller, B. M.; Key, C. R.; Maygarden, S. J.; Ballard-Barbash, R. *Cancer* **2006**, *106*, 732–742.