Hyperspectral imaging for identifying non-O157 Shiga toxin-producing *Escherichia coli* (STEC) serotypes

Seung-Chul Yoon[1]*, William R. Windham[1], Scott Ladely[2], Kurt C. Lawrence[1], Bosoon Park[1], Neelam Narang[2] and William C. Cray[2]

[1]U.S. Department of Agriculture, Agricultural Research Service, Athens, GA 30605, USA
[2]U.S. Department of Agriculture, Food Safety and Inspection Service, Athens, GA, 30605, USA
*Corresponding author: seungchul.yoon@ars.usda.gov

Introduction

Rapid detection and identification of pathogenic microorganisms in food is important for the development of intervention and verification strategies for the food industry and regulatory agencies. Traditional culture-based direct plating methods are still the "gold standard" for presumptive-positive pathogen detection.[1,2] These methods typically involve an enrichment process to amplify microbial cell numbers up to detectable levels, followed by plating on selective or differential agar media to grow colonies showing distinctive colour and other morphological characteristics such as shape and texture. Culture-based methods have good specificity, sensitivity and low costs, and can provide both colony count estimation and qualitative information about the microorganisms present in the food samples. However, culture-based methods are labour-intensive and time-consuming as the microorganisms must multiply to visible colonies for a presumptive-positive result. Another challenge is that unwanted background microflora grow together with target microorganisms on agar media and often look similar. Hence, highly skilled technicians are needed to pick up presumptive-positive colonies by trial and error. Nonetheless, the presumptive-positive colonies must be confirmed by serological or molecular methods to determine type.

Escherichia coli bacteria live in the intestine of warm-blooded animals, including humans. Many *E. coli* serotypes do not cause human disease; there is however, a pathogenic group of *E. coli* that produces Shiga toxin. Symptoms of illnesses caused by the consumption of Shiga toxin-producing *E. coli* (STEC) are diarrhoea, stomach cramps, vomiting and a potentially lethal kidney complication called hemolytic uremic syndrome (HUS). The most prevalent and commonly recognised STEC serotype is *E. coli* O157:H7; non-O157 STEC serogroups such as O26, O103, and O111 are also increasingly recognised.[3-4]

It is well known that the serotype O157:H7 can be easily detected when cultured on sorbitol-MacConkey (SMAC) agar for 24 h. The reason SMAC agar is effective for the detection of O157:H7 is that O157:H7 does not ferment sorbitol within 24 h unlike most other *E. coli* serotypes. Being unable to ferment sorbitol makes O157:H7 colonies colourless after overnight incubation while most other *E. coli* colonies appear pink. However, the specificity of SMAC agar is limited to O157:H7 only. Recently, Rainbow agar O157 (in short, Rainbow agar) was developed to isolate not only STEC O157:H7 but also non-O157 STEC serotypes, and to differentiate them from other non-toxigenic *E. coli*.[2] STEC O157:H7 appear black when cultured on Rainbow agar[2] and non-O157 STEC appear purple, grey or grey-blue.[5] Unfortunately, Rainbow agar is not specific enough to differentiate one non-O157 STEC serotype from another. The screening task for STEC colonies by direct plating is further complicated by the fact that food samples harbouring STEC pathogens, such as beef products, potentially have a large number of background microorganisms which can grow together with STEC. Although time-consuming and labour-intensive, traditional culture-based methods still represent a field where progress is needed in order to more accurately differentiate pathogen colonies from one another and from background flora.

Hyperspectral imaging is an optical imaging technique that combines aspects of conventional imaging and vibrational spectroscopy so that data can provide two-dimensional spatial information on colony shapes and spectral information at every pixel in each colony under test. The spectral "fingerprints" of bacteria provided by hyperspectral imaging can be used for detection and identification of pathogens. In our previous study on *Campylobacter* detection, we developed a hyperspectral imaging technique to discriminate *Campylobacter* cultures growing on agar from non-pathogenic cultures commonly found during poultry processing.[6-7] A reflectance spectral library of *Campylobacter* and non-*Campylobacter* contaminants was constructed for characterisation of absorption features and for developing classification methods.[6-7] The hyperspectral reflectance imaging protocol developed was modified in this research to detect STEC since it might have the potential to fulfil a research gap for a rapid and accurate presumptive-positive screening method for non-O157 STEC serotypes. Thus, the objective of this paper was to evaluate visible and near-infrared (VNIR) hyperspectral imaging as a new screening method to detect the top six non-O157 STEC serogroups[3]: O26, O45, O103, O111, O121, and O145 grown on Rainbow agar plates.

Reference paper as:
Yoon, S., Windham, W.R., Ladely, S., Lawrence, K.C., Park, B. , Narang, N. and Cray, W.C. (2012). Hyperspectral imaging for identifying non-O157 Shiga toxin-producing Escherichia coli (STEC) serotypes, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 58-62.

Materials and Methods

For hyperspectral data collection, non-O517 STEC pathogens were grown on agar plates and imaged. The first step for the development of hyperspectral image processing methods was to build a spectral library of non-O517 STEC serotype colony spectral signatures. The second step was to develop classification and prediction algorithms to identify the six non-O157 STEC serotypes.

STEC samples and spot plating

The six STEC serotypes used were O26:H2 (strain 4), O45:H2 (strain 8), O103:H1 (strain D), O111:H1 (strain 16), O121:H19 (strain A), and O145:H- (strain K). These were obtained from Eastern Laboratory Outbreak Section of the Food Safety and Inspection Service (FSIS), Richard B. Russell Research Center, Athens, GA, USA. All strains were positive for Shiga toxin (stx1 and/or stx2) and intimin (eae) genes. Serial dilutions of each serotype were prepared from cultures grown overnight on blood agar (BA, Trypticase Soy Agar with 5% sheep blood, Remel, Lenexa, KS, USA) in physiological sterile saline. Following the initial culturing step, a cell suspension of each culture was made in sterile 0.85% NaCl. For each serotype, plates were inoculated with 5 µl of a 10^8 cfu ml^{-1} cell suspension and incubated at 37˚C for 24 h. Following the above protocol, one experiment of duplicate plates was carried out in 2010. To avoid confluent growth, cross-contamination, and to obtain sufficient colony mass to collect spectral signatures, agar plates were inoculated with the 5 µl "spots". Figure 1 shows a pseudo-colour composite image of the six STEC serotypes grown on Rainbow agar.

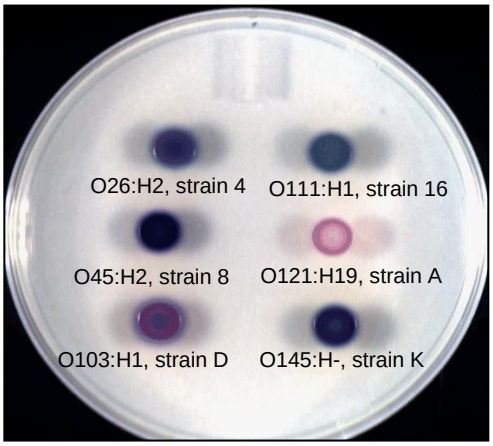

Figure 1. Spot plates of STEC serotypes.

Hyperspectral image acquisition and processing

The VNIR hyperspectral imaging system consisted of a hyperspectral camera (Institute of Technology Development, Stennis Space Center, MS, USA) attached to a copy stand and a computer that controls image acquisition. Illumination was provided by two 75 W tungsten-halogen lamps. The pixel resolution of the charge-coupled device (CCD) detector is 1376 × 1040. The nominal spectral range of the imaging system was from 400 to 1000 nm. Each agar plate was scanned line-by-line with 2 (spatial) × 4 (spectral) software binning and 40 ms exposure time. The resulting pixel resolution of one line scan image was 688 (spatial) × 260 (spectral). Each agar plate consisted of 500 lines. All image intensity values were calibrated to relative reflectance using a 75% reflectance Spectralon target. Random spectral noise was reduced by a Savitzky-Golay smoothing filter (window size = 25; order of moment = 4). Calibrated and de-noised hyperspectral images were stitched together to create a single image mosaic. The pixel resolution of a hyperspectral image dataset after pre-processing was 688 (W) × 500 (H) × 197 (λ) in the range of 400 to 900 nm. The average difference between adjacent wavebands was 2.55 nm. Two of the duplicate plate images measured were aligned side by side in a mosaic image. Ground-truth regions-of-interest (ROIs) of only pure organisms were manually selected using ENVI software (Exelis Visual Information Solutions, Boulder, CO, USA) from the mosaic image of all serotypes and agar. If colour and spectral signatures near the centre of each colony spot were visually different from the leading edge of new growth, open ellipses (like donuts) were used for ROIs of the colony. Otherwise, closed ellipses were used. The set of these ROIs was called ROISet1. Smaller sizes of ROIs, called ROISet2 and ROISet3, were prepared by random sampling ROISet1. The spectral signatures obtained from ROIs were compiled into a spectral library and used in classification algorithm development and validation.

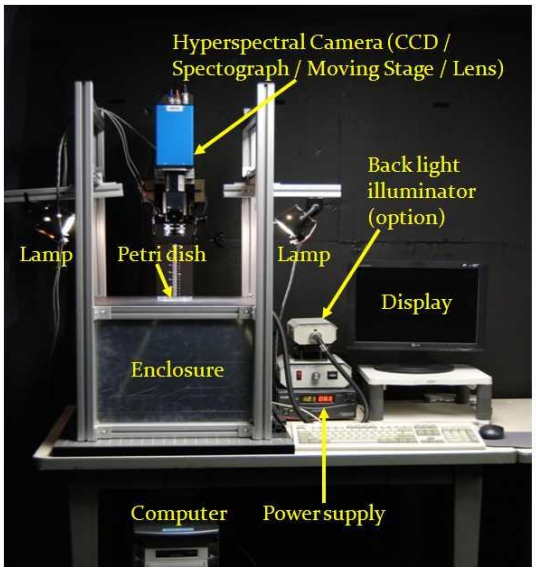

Figure 2. VNIR hyperspectral imaging system.

Reference paper as:
Yoon, S., Windham, W.R., Ladely, S., Lawrence, K.C., Park, B. , Narang, N. and Cray, W.C. (2012). Hyperspectral imaging for identifying non-O157 Shiga toxin-producing Escherichia coli (STEC) serotypes, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 58-62.

Classification algorithm development and validation

Discriminant models for each serotype were developed from two supervised classifiers: support vector machine (SVM) and Mahalanobis distance. Reflectance spectra were used for model development with the SVM classifier. First derivative spectra with a 4-point gap (10 nm) were used for the Mahalanobis distance classification. The Mahalanobis distance classification was done in the subspace spanned by the first three principal components (PCs) obtained from the principal component analysis (PCA) of ROISet3. The PC scores for each serotype were used to calculate Mahalanobis distances from each pixel in the plate image. The first derivatives of each pixel spectrum in the plate image were projected into the subspace by the loadings obtained from ROISet3. The minimum distance rule was used for the classification. Performance of both classifiers was validated on ROISet1.

Results and Discussion

Spectral characteristics of non-O157 STEC

The average ROI reflectance spectra for each serotype are shown in Figure 3 which were obtained from ROISet1 (N = 17,358). O121 had a high reflectivity across the entire visible spectrum, particularly from 500 to 700 nm. This spectral curve confirmed the bright pinkish hue of O121 pixels which were easily differentiated visually and spectrally relative to the other serotypes. O111 also showed a distinctive spectral signature linearly decreasing from 550 to 650 nm, thus, O111 appeared grey-bluish. However, O45 was very absorptive across the visible spectrum with less than 10% reflectance in the 500–700 nm while the 400–500 nm region (blue) was somewhat less absorptive. This created a very dark appearance to the human eye and made it spectrally unique relative to the other serotypes. Finally, O26, O103 and O145 showed similar spectral responses; all had grey centres surrounded by more reddish perimeters.

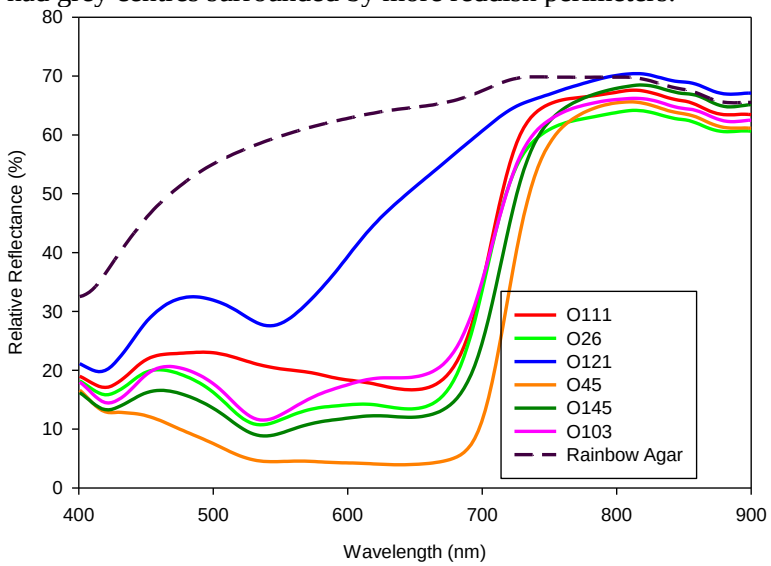

Figure 3. Average ROI reflectance spectra of non-O157 STEC serotypes.

Classification

The datasets for training the SVM classifier were prepared from ROISet1 (N = 17 358), ROISet2 (N = 1881), and ROISet3 (N = 294), respectively. Note that ROISet2 and ROISet3 were obtained from only one plate image on the right of the mosaic. The trained SVM model was applied to predict the class at each pixel in the image mosaic; the performance of the SVM classifier was validated using ROISet1. Classification accuracies were 99.9%, 96.0% and 91.3% for ROISet1, ROISet2 and ROISet3, respectively. Figure 4 shows the classification result image predicted by the SVM classifier trained using ROISet3. The prediction accuracies of each serotype were 100% O111, 99.8% O121, 95.1% O45, 93.7% O26, 85.9% O145, and 67.7% O103, respectively. As expected from the average spectra in Figure 3, O111, O121 and O45 were accurately classified with over 95% of pixel-level accuracy and the other three serotypes were not quite separable. Of O103 pixels, 24.8% were misclassified as O26 while only 3.2% of O26 were misclassified as O103. Also, 10.1% of O145 were misclassified as O26 while none of O26 was classified as O145. This result can be explained by the location of the average spectrum of O26 in-between O103 and O145. When qualitatively analysing the classification image, the SVM classifier performed poorly along the outer and inner rims of spots where mixed spectral phenomena dominated the reflectance features on those pixels. Thus, in the future, if the colony morphological information is incorporated into a decision-making algorithm, the

Reference paper as:
Yoon, S., Windham, W.R., Ladely, S., Lawrence, K.C., Park, B. , Narang, N. and Cray, W.C. (2012). Hyperspectral imaging for identifying non-O157 Shiga toxin-producing Escherichia coli (STEC) serotypes, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 58-62.

performance may be improved. For example, if an automated algorithm makes a decision at each segmented colony (a big spot in this case) by excluding pixels at the rims, the performance is expected to improve.

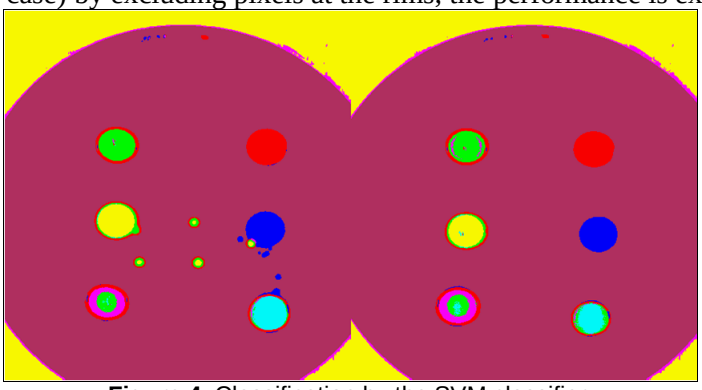

Figure 4. Classification by the SVM classifier.

The Mahalanobis distance classifier performed better than SVM. The overall classification accuracy was 94.1% versus 91.3% of SVM when trained using ROISet3. After predicting the classes of all pixels in the image mosaic, the performance of the Mahalanobis distance classifier was validated with ROISet1. Figure 5 shows the classification result image predicted by the Mahalanobis distance classifier. The prediction accuracies of each serotype were 100% O111, 100% O121, 82.8% O45, 99.0% O26, 86.3% O145 and 95.2% O103, respectively. The performance for both O111 and O121 was very close to that of the SVM. Although the detection accuracy of O45 was 12.3% worse than the SVM, most other serotypes were more accurately classified by the Mahalanobis distance classifier with improvements of 0.2% O121, 5.3% O26, 0.4% O145, and 27.5% O103, respectively. However, rim pixels still suffered from misclassifications. The PC scores obtained by ROISet1 are shown in Figure 6. This score plot explains the aforementioned poor performance on O45 because the scores of O45 were mixed into two neighbouring groups, O145 and O103, in the PC space. The other serogroups, O121, O145, O103, O26 and O111, were highly separable from one another in the PC domain.

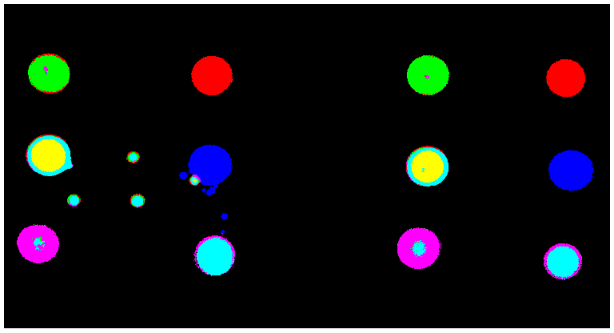

Figure 5. Classification by the Mahalanobis distance classifier.

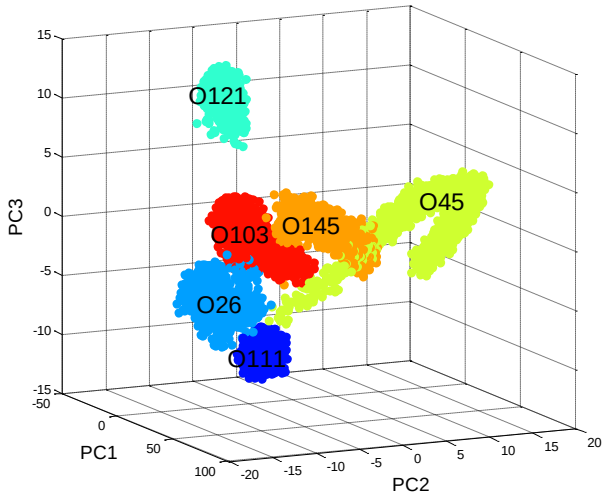

Figure 6. Principal component scores used for Mahalanobis distance classification.

Conclusion

A VNIR hyperspectral imaging technique for detection of non-O157 STEC serotypes on agar plates was investigated by building spectral libraries and developing classification models. Pure cultures of six non-O157 STEC serotypes were grown on agar plates. The classification results obtained by SVM and Mahalanobis distance classifiers were mapped into the agar plate images and the results were validated. Serotypes O111 and O121 consistently showed over 99% accuracy regardless of the classification algorithms. However, the classification accuracies of serotypes O26, O45, O103 and O145 had varying results from 67% up to 99%, depending on the classification algorithm adopted. The mixed results were caused by the similar VNIR spectral characteristics among O26, O103 and O145. However, first derivatives of spectra helped to increase the detection accuracies of O26, O103 and O145. In conclusion, the hyperspectral imaging methods have the potential to be developed as a non-contact and non-destructive

Reference paper as:
Yoon, S., Windham, W.R., Ladely, S., Lawrence, K.C., Park, B. , Narang, N. and Cray, W.C. (2012). Hyperspectral imaging for identifying non-O157 Shiga toxin-producing Escherichia coli (STEC) serotypes, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 58-62.

screening tool for detection of non-O157 STEC serotypes. But, to be more practical, the methods will have to be refined to detect bacterial samples grown on spread plates.

Acknowledgements

The authors would like to thank Jerry Heitschmidt for his technical assistance in hyperspectral image data collection.

References

1. C.S. Beauchamp and J.N. Sofos, "Diarrheagenic *Escherichia coli*", in *Pathogens and toxins in foods: challenges and interventions*, Ed by V. K. Juneja and J. K. Sofos, ASM, Washington, DC, p. 82 (2009).
2. J. Meng and M.P. Doyle, "Microbiology of Shiga toxin-producing *Escherichia coli* in foods", in *Escherichia coli 0157:H7 and other Shiga toxin-producing E. coli strains*, Ed by J.B. Kaper and A.D. O'Brien, ASM, Washington, DC, p. 97 (1998).
3. J.M. Bosilevac and M. Koohmaraie, *Appl. Environ. Microbiol.* **77**, 2103-2112 (2011).
4. P.M. Griffin, "Epidemiology of Shiga toxin-producing *Escherichia coli* infections in humans in the United States", in *Escherichia coli 0157:H7 and other Shiga toxin-producing E. coli strains*, Ed by J. B. Kaper and A. D. O'Brien ASM, Washington, DC, p. 15 (1998).
5. K.A. Bettelheim, *Microbiol. Immunol*. **42**, 265-269 (1998).
6. S.C. Yoon, K.C. Lawrence, G.R. Siragusa, J.E. Line, B. Park and P.W. Feldner, *Trans. ASABE* **52**, 651-662 (2009).
7. S.C. Yoon, K.C. Lawrence, J.E. Line, G.R. Siragusa, P.W. Feldner, B. Park and W.R. Windham, *Sens. Instrum. Food Qual. Saf.* **4**, 35-49 (2010).

Reference paper as:
Yoon, S., Windham, W.R., Ladely, S., Lawrence, K.C., Park, B. , Narang, N. and Cray, W.C. (2012). Hyperspectral imaging for identifying non-O157 Shiga toxin-producing Escherichia coli (STEC) serotypes, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 58-62.