Received: 31 October 2019 | Accepted: 14 February 2020

DOI: 10.1111/jvim.15742

STANDARD ARTICLE

Journal of Veterinary Internal Medicine Open Access

ACVIM American College of Veterinary Internal Medicine

Differentiation of lymphocytic-plasmacytic enteropathy and small cell lymphoma in cats using histology-guided mass spectrometry

Sina Marsilio[1,2] | Shelley J. Newman[3] | James Scot Estep[4] | Paula R. Giaretta[5] | Jonathan A. Lidbury[2] | Emma Warry[6] | Andi Flory[7] | Paul S. Morley[8] | Katy Smoot[9] | Erin H. Seeley[9] | Matthew J. Powell[9] | Jan S. Suchodolski[2] | Jörg M. Steiner[2]

[1]Department of Veterinary Medicine and Epidemiology, UC Davis School of Veterinary Medicine, Davis, California

[2]Gastrointestinal Laboratory, Department of Small Animal Clinical Sciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, Texas

[3]Long Island University CVM, Brookville, New York

[4]Texas Veterinary Pathology, LLC, San Antonio, Texas

[5]Department of Veterinary Pathobiology, Texas A&M University, College Station, Texas

[6]Department of Small Animal Clinical Sciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, Texas

[7]Veterinary Specialty Hospital, San Diego, California

[8]Veterinary Education, Research, and Outreach Center, Texas A&M University, Canyon, Texas

[9]New River VDL, LLC, Morgantown, West Virginia, USA

Correspondence
Sina Marsilio, Department of Veterinary Medicine and Epidemiology, UC Davis School of Veterinary Medicine, Tupper Hall, 1275 Med Science Dr, Davis, CA 95616.
Email: smarsilio@ucdavis.edu

Funding information
New River VDL, LLC

Abstract

Background: Differentiation of lymphocytic-plasmacytic enteropathy (LPE) from small cell lymphoma (SCL) in cats can be challenging.

Hypothesis/Objective: Histology-guided mass spectrometry (HGMS) is a suitable method for the differentiation of LPE from SCL in cats.

Animals: Forty-one cats with LPE and 52 cats with SCL.

Methods: This is a retrospective clinicopathologic study. Duodenal tissue samples of 17 cats with LPE and 22 cats with SCL were subjected to HGMS, and the acquired data were used to develop a linear discriminate analysis (LDA) machine learning algorithm. The algorithm was subsequently validated using a separate set of 24 cats with LPE and 30 cats with SCL. Cases were classified as LPE or SCL based on a consensus by an expert panel consisting of 5-7 board-certified veterinary specialists. Histopathology, immunohistochemistry, and clonality testing were available for all cats. The panel consensus classification served as a reference for the calculation of test performance parameters.

Results: Relative sensitivity, specificity, and accuracy of HGMS were 86.7% (95% confidence interval [CI]: 74.5%-98.8%), 91.7% (95% CI: 80.6%-100%), and 88.9% (95% CI: 80.5%-97.3%), respectively. Comparatively, the clonality testing had a sensitivity, specificity, and accuracy of 85.7% (95% CI: 72.8%-98.7%), 33.3% (95% CI: 14.5%-52.2%), and 61.5% (95% CI: 48.3%-74.8%) relative to the panel decision.

Conclusions and Clinical Importance: Histology-guided mass spectrometry was a reliable technique for the differentiation of LPE from SCL in duodenal formalin-fixed paraffin-embedded samples of cats and might have advantages over tests currently considered state of the art.

Abbreviations: CE, chronic enteropathy; CI, confidence interval; FFPE, formalin-fixed paraffin-embedded; H&E, hematoxylin and eosin; HGMS, histology-guided mass spectrometry; IBD, inflammatory bowel disease; IHC, immunohistochemistry; LDA, linear discriminant analysis; LPE, lymphocytic-plasmacytic enteropathy; MALDI, matrix-assisted laser desorption ionization; MSI, mass spectrometry imaging; PARR, PCR for antigen receptor rearrangements; PCA, principal component analysis; PCR, polymerase chain reaction; SCL, small cell lymphoma.

KEYWORDS
chronic enteropathy, clonality testing, feline, HGMS, inflammatory bowel disease, inflammatory enteropathy, MALDI mass spectrometry, PARR, PCR for antigen receptor rearrangements, small cell lymphoma

1 | INTRODUCTION

Chronic enteropathy (CE) is the most common gastrointestinal disorder in older cats, with a rising prevalence over the past decade.[1] The disorder mostly comprises lymphocytic-plasmacytic enteropathy (LPE) and small cell lymphoma (SCL).[2]

Routine work-up for cats with CE often includes the collection of biopsy specimens and the histopathologic evaluation of formalin-fixed paraffin-embedded (FFPE), hematoxylin and eosin (H&E)-stained tissue sections.[3,4]

Immunohistochemistry (IHC) with stains specific for T- and B-cells, as well as clonality assays, is considered state of the art by some authors.[4,5] However, the specificity of polymerase chain reaction (PCR)-based clonality tests has been questioned[6,7] and some consider a combination of histopathology and immunohistochemistry as the gold standard.[8] The sensitivity and specificity of clonality assays in human and veterinary medicine vary widely with sensitivities between 70.0 and 97.6% and specificities between 54.3-98.7%.[7,9-15] These wide ranges are the result of a combination of technical and biological variability.[16] Because of the inherently high error rate for clonality assays, there are rigorous preanalytical, analytical, and postanalytical standards published in human medicine under the EuroClonality/BIOMED-2 clonality standardization group guidelines.[16,17] However, there are no standards for the conduction of clonality assays in veterinary medicine.[5,18] To the authors' knowledge, there are no published data on the sensitivity and specificity of PCR-based clonality assays conducted on FFPE tissue from cats with either LPE or SCL.

Histology-guided mass spectrometry (HGMS) profiling is a proprietary application of matrix-assisted laser desorption/ionization (MALDI) mass spectrometry imaging (MSI) that employs histopathologic annotations for the targeted analysis of endogenous molecules from specific cell subpopulations.[19] This technology utilizes a machine learning algorithm to statistically assess the distribution of targeted biomolecules, such as proteins, lipids, or metabolites, to differentiate molecular signatures of specific disease states.[19] The ability of HGMS to distinguish between different diseases has been well demonstrated in animal models, and this technique has been successfully applied in human medicine to differentiate benign nevi from malignant melanoma.[20-22]

The aim of this retrospective study was to investigate HGMS as a potential new tool for the differential diagnosis of SCL from LPE. We hypothesized that HGMS profiling of FFPE duodenal tissue samples and machine learning could be employed to develop a classification algorithm to distinguish proteomic signatures of lymphocytic-plasmacytic inflammatory lesions and SCL in cats. In addition, we compared the performance of PCR for antigen receptor rearrangements (PARR) and HGMS for the differentiation of SCL from LPE in cats with CE.

2 | MATERIALS AND METHODS

2.1 | Sample cohort

This retrospective study was conducted at the Gastrointestinal Laboratory at Texas A&M University, College Station, Texas. The histology archive of the of the Texas A&M Gastrointestinal Laboratory was searched for tissue samples from cats with a history of CE. Cases were selected based on the history available on the Texas A&M Gastrointestinal Laboratory biopsy submission form and the initial histopathology report of either predominantly lymphocytic-plasmacytic enteritis or SCL. All histopathologic evaluations were performed by board-certified anatomic pathologists in accordance with guidelines of the World Small Animal Veterinary Association Gastrointestinal Standardization Group. World Small Animal Veterinary Association standards included morphological features (ie, villus stunting, epithelial injury, crypt distention, lacteal dilatation, and mucosal fibrosis) as well as inflammatory changes (ie, intraepithelial lymphocytes, lamina propria lymphocytes and plasma cells, eosinophils, and neutrophils).[23,24] Formalin-fixed paraffin-embedded tissue samples from all cases were sent to one of four external laboratories for immunohistochemistry (IHC) staining with CD3 and CD20 (Texas A&M Veterinary Medical Diagnostic Laboratory, College Station, Texas, USA; Michigan State University, Veterinary Diagnostic Laboratory, Lansing, Michigan, USA; University of California Davis, Leukocyte Antigen Biology Laboratory, Davis, California, USA) and for clonality testing (TDDS, SYNLAB VPG Ltd., Exeter, UK; Michigan State University, Veterinary Diagnostic Laboratory, Lansing, Michigan, USA; University of California Davis, Leukocyte Antigen Biology Laboratory, Davis, California, USA).

2.2 | Panel evaluation

A panel consisting of 5-7 board-certified veterinary specialists was formed, comprising anatomic pathologists (S. Newman, S. Estep, present for a subset of cases P. Giaretta), internists (S. Marsilio, J. Lidbury), and oncologists (A. Flory, present for a subset of cases E. Warry). The expert panel reviewed all data points available from a single case including

signalment, clinical history listed on the initial tissue submission form, completed referring veterinarian questionnaires (when available), initial histopathology reports from the previous attending pathologist, digital images of H&E- and IHC-stained slides, and clonality results. On the basis of all available information, a final diagnosis was made for each case based on unanimous votes of the panel, either through remote voting or during a panel gathering in a round table discussion. Cases that were unanimously classified with a consensus diagnosis as having either LPE or mucosal SCL were eligible to be used for algorithm development with the HGMS profiling data. Cases where no consensus could be reached, cases with isolated epitheliotropic lymphoma (ie, minimal lymphocytic infiltration in the lamina propria), or cases in which the panel found diagnoses other than those allowed within the inclusion criteria (ie, normal tissue, large cell lymphoma, eosinophilic inflammation, neutrophilic inflammation) were excluded from further analysis. Importantly, the members of the consensus panel were blinded to the HGMS results.

2.3 | Histology-guided mass spectrometry

An HGMS workflow overview is shown in Figure 1. Two serial 5-µm, duodenal tissue sections were cut from FFPE biopsy specimen blocks. One section per sample was collected onto an indium-tin oxide-coated glass slide compatible with the mass spectrometer, and the consecutive serial section was mounted onto a conventional glass slide and stained with H&E, which served as a reference section. The slides were digitized (Huron TissueScope LE120, Huron Digital Pathology, St. Jacobs, Ontario, Canada) and uploaded to an online viewing and annotation portal (proteaScope, New River VDL) for histopathological review. Areas of lymphocytic-plasmacytic infiltration within the lamina propria (~20 annotations per sample, 50 µm diameter) were annotated by a veterinary pathologist (S.N.) on the digital images. Known proteomic interferences, such as connective tissue, blood vessels, epithelium, and necrotic tissue were avoided during annotation.

Sections for mass spectrometry were deparaffinized with xylene and a series of graded ethanol washes. Samples were air-dried followed by antigen retrieval in Tris buffer, pH 9 (Decloaking Chamber NxGen, Biocare Medical, Pacheco, California) for 25 minutes at 95°C, and buffer exchanged to water. Images of the unstained sections were acquired using a flatbed document scanner. Annotated images of the H&E-stained sections were digitally merged with the images of the unstained sections to allow for accurate target designation in the MALDI mass spectrometer.

Sections for mass spectrometry were subjected to on-tissue digestion and matrix application using an HTX M5 Sprayer (HTX Technologies, Chapel Hill, North Carolina). Porcine trypsin (Pierce) in a concentration of 0.05 mg/mL was sprayed onto tissue over a series of 12 passes at a flow rate of 0.01 mL/min and a nozzle velocity of 750 mm/min. Samples incubated at 37°C for a total of 4 hours, followed by application of 10 mg/mL of alpha-cyano-4-hydroxycinnamic acid (Millipore Sigma) in a 70: 30: 0.1 acetonitrile/water/trifluoroacetic acid mixture over a series of 4 passes (flow rate of 0.120 mL/min and a nozzle velocity of 1200 mm/min).

Mass spectrometry data from designated areas was collected using a Bruker rapifleX™ MALDI TOF mass spectrometer (Bruker Daltonics, Billerica, MA), which operated in positive ion reflectron mode using flexImaging software for plate alignment and designation of annotated, histological targets.

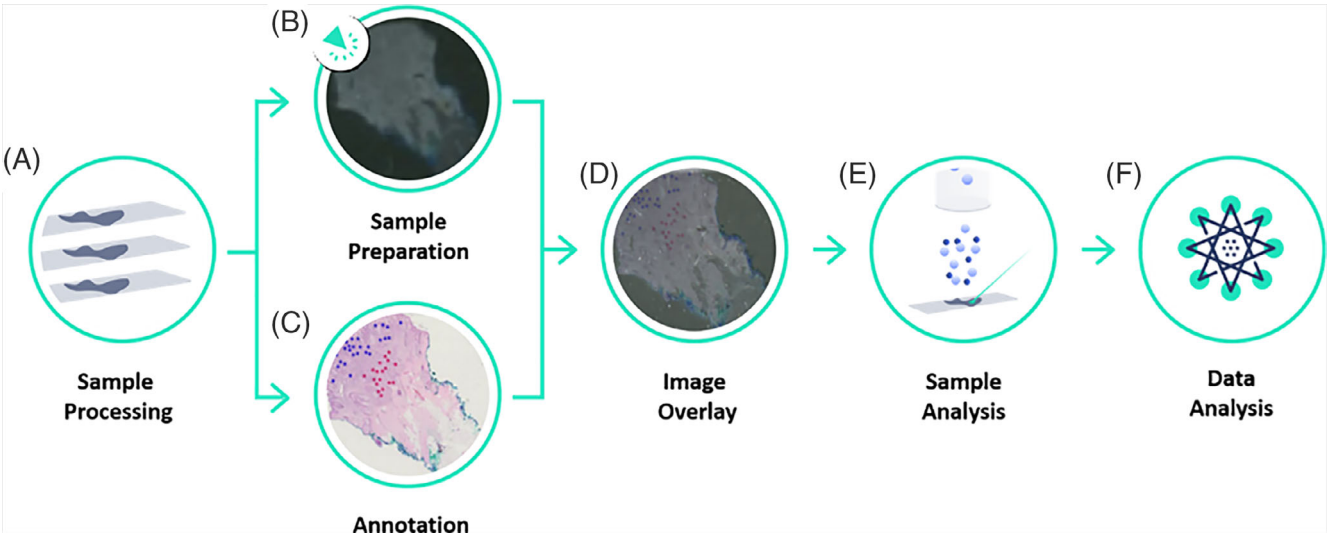

FIGURE 1 HGMS workflow: A, sample processing—FFPE embedding and cutting of 5-µm serial tissue sections; B, sample preparation—paraffin removal, antigen retrieval, trypsin digestion, and MALDI matrix application; C, annotation—anatomic pathologist marks 50-µm diameter annotations of lymphocyte cell subpopulations on H&E image; D, image overlay—digital images of prepped slide and annotated H&E image are merged to teach mass spectrometer locations for analysis; E, sample analysis—proteomic data acquired by HGMS in MALDI mass spectrometer; F, data analysis—for algorithm development, machine learning is used to generate classification algorithm; for unknown sample assessment, classification algorithm is used to classify sample as SCL or LPE. FFPE, formalin-fixed paraffin-embedded; H&E, hematoxylin and eosin; HGMS, histology-guided mass spectrometry; LPE, lymphocytic-plasmacytic enteropathy; MALDI, matrix-assisted laser desorption ionization; SCL, small cell lymphoma

2.4 | Algorithm development and statistical analysis

Data from biopsy specimens were loaded into SCiLS Lab Pro 2019c (Bruker Daltonics) and preprocessed via baseline correction and normalization. Data were exported and processed in R studio for the selection of discriminatory mass spectral peaks. The peak list was transferred to SCiLS Lab Pro, and samples were tagged to belong to 1 of 4 classification groups: SCL Training (22 samples), LPE Training (17), SCL Validation (30), and LPE Validation (24). Training set samples were selected from review of archival data that, at the time of diagnosis, were considered as clear-cut cases of SCL or LPE with the additional criterion that the IHC and PARR data were consistent with the diagnosis. Panel review of all training set cases yielded consensus diagnoses that were consistent with the original diagnoses. Because the training set was small, this set was later expanded by adding additional cases that were evaluated by the expert panel. The linear discriminate analysis (LDA) machine learning algorithm was developed and optimized using the training set. All of the cases were either used for the training set or the validation set, and none of the cases were used for both sets of cases. The validation set was then classified using the optimized LDA algorithm to determine assay accuracy. A two-third classification criterion was used wherein at least two-third of spectra must classify as either SCL or LPE to render a differential diagnosis. Samples that did not have at least two-third of spectra classifying as SCL or LPE were classified as indeterminant. Differences between ages and sex were calculated with Prism 8 (Version 8.2.1, Graph Pad Software, San Diego, California) using the Mann-Whitney *U* and the chi-square tests, respectively.

3 | RESULTS

3.1 | Demographics

A total of 121 cases were evaluated by a panel of veterinary experts, of which 93 cases received a consensus diagnosis as either SCL (52) or LPE (41) within duodenal tissue sections. The remaining 28 cases either did not receive a concordant diagnosis (18), exhibited no signs of SCL or LPE (4), or displayed evidence of other pathologies that made the samples incompatible for the study (6).

Cats meeting the inclusion criteria had a median age of 11 years (range: 1-17 years). There were 39 female (38 spayed) and 44 male (41 neutered) cats; the sex was unknown for 10 cats. Breeds included domestic shorthair (n = 56), domestic longhair (n = 12), domestic medium hair (n = 3), Siamese (n = 2), Burmese (n = 1), Norwegian Forest Cat (n = 1), American (n = 1), Bengal (n = 3), Maine Coon (n = 7), Ragdoll (n = 2), Russian Blue (n = 1), Persian (n = 1), and unknown (n = 3). Questionnaires were available for 17 cases. The collection method was reported as endoscopic in 91 cases and as surgical (full-thickness) in 2 cases. Tables S1 and S2 summarize the demographics and consensus diagnoses for all cats used within the training and validation data sets, respectively.

There was no significant difference between the major demographic characteristics, age, and sex between the training set and the validation set.

3.2 | Initial histopathological diagnoses

On the basis of evaluation of the 39 H&E-stained duodenal biopsy specimens comprising the training set, a diagnosis of SCL or suspected SCL was reached in 20 cases. In 19 cases, the pathologist found LPE (1 minimal to mild, 2 mild, 5 mild to moderate, 8 moderate, 2 moderate to severe, 1 mild LPE with marked lymphocytic epitheliotropism).

Histopathological evaluation of 54 biopsy specimens in the validation set, revealed a diagnosis of LPE in 26 cats (4 mild, 8 mild to moderate, 13 moderate, 1 moderate to severe). Twenty-seven cats were diagnosed with SCL or received a diagnosis of suspected SCL. One cat was diagnosed with a medium cell lymphoma. Tables S1 and S2 show the complete list of histopathological diagnoses for the training set and validation set, respectively.

3.3 | Histology-guided mass spectrometry

3.3.1 | Training set

Histology-guided mass spectrometry data representative of proteomic signatures were collected from all samples for the development of a diagnostic algorithm to differentiate SCL from LPE. Statistical software identified 18 discriminatory mass peaks based on comparisons of peak intensities between SCL and LPE data. Figure 2 illustrates the difference in intensity distribution of 2 of the discriminatory peaks, m/z 1270.7 and m/z 1402.7, both of which displayed an overall higher intensity in samples from cats with LPE. Using the 18 predictive mass peaks, a LDA machine learning algorithm was generated from a training set of 22 SCL and 17 LPE cases. An internal cross-validation assessment was performed on the algorithm using leave-20%-out repeated random subsampling, which resulted in 89.9% spectral classification accuracy. Biological variance was observed between SCL and LPE in the training set via a principal component analysis (PCA) illustrated in Figure 3.

3.3.2 | Validation set

The optimized LDA algorithm for the HGMS test was applied to an independent set of 54 biopsy samples (30 SCL and 24 LPE) to assess the classification accuracy of the model relative to the consensus diagnoses from the review panel. Table S2 summarizes the demographics, histopathological findings, molecular results, consensus diagnoses, and HGMS results for all cats within the validation data set. Of the 30 cases that received a consensus diagnosis of SCL by the review panel, 26 cases were classified correctly as SCL by HGMS, 2 cases classified incorrectly as LPE, and 2 cases received indeterminant scores, which

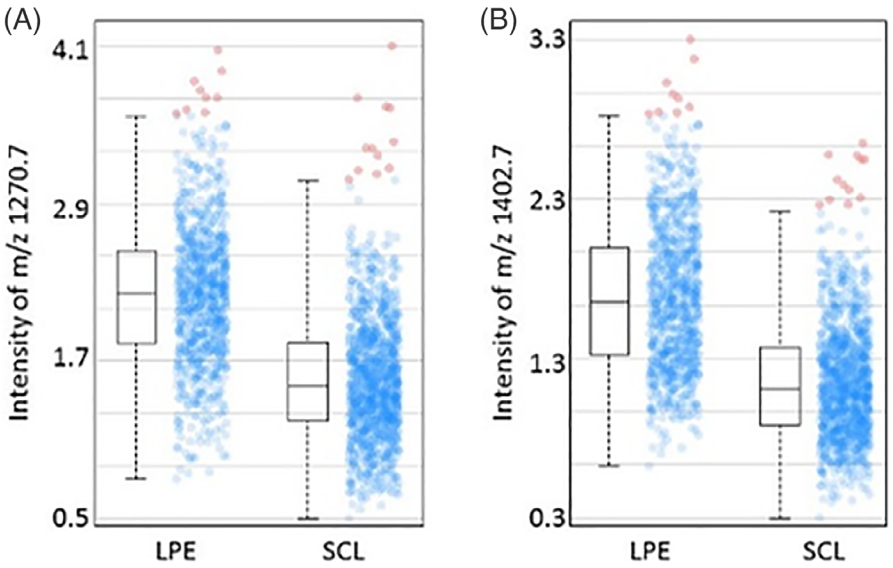

FIGURE 2 Intensity box plot comparisons of (A) *m/z* 1270.7 and (*B*) *m/z* 1402.7 from SCiLS software. Each dot represents the normalized peak intensity of a unique mass spectrum, corresponding to a single annotation on a tissue section. All data from the cases receiving LPE and SCL diagnoses from the review panel are plotted. The center line in the box indicates the median, and the lower and upper boundaries of the box indicate the second and third quartiles. The vertical lines protruding from the box designate the lower (0%) and upper quantile (99%)

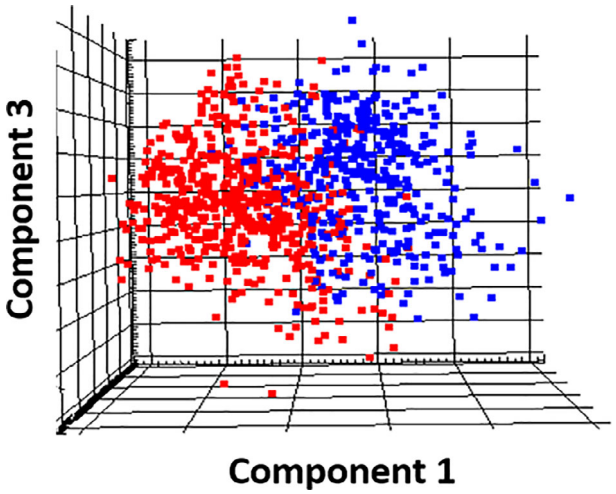

FIGURE 3 Three-dimensional PCA of training set data showing clustering of SCL (red) and LPE (blue) spectra. LPE, lymphocytic-plasmacytic enteropathy; PCA, principal component analysis; SCL, small cell lymphoma

resulted in a relative sensitivity of 86.7% (95% confidence interval [CI]: 74.5%-98.8%). Of the 24 cases that received a consensus diagnosis of LPE, 22 cases were classified correctly as LPE by HGMS, 2 cases were classified incorrectly as SCL, and 0 cases received indeterminant scores, which resulted in a relative specificity of 91.7% (95% CI: 80.6%-100%). For all 54 cases of the validation set, the accuracy of the HGMS test relative to the consensus diagnoses of the review panel was calculated to be 88.9% (95% CI: 80.5%-97.3%). Overall, there was an indeterminant rate of 3.7% for the validation set. For the calculation of relative sensitivity, specificity, and accuracy, cases that were classified as indeterminant were considered as incorrect.

3.4 | PARR results

Of the 54 cats assigned to the validation data set, 34 were clonal on FFPE sections from duodenal tissue biopsy specimens, whereas 8 were polyclonal, 3 as oligoclonal, and 7 as pseudoclonal. PARR data could not be obtained for 2 of the validation samples because of insufficient tissue material.

Of the 34 cases found to be clonal, 23 cats received a consensus diagnosis of SCL, whereas 11 were diagnosed as LPE by the review panel. For the 8 cases with reportedly polyclonal results, all 8 cats received a final consensus diagnosis of LPE. Of the 3 cases exhibiting oligoclonality, 2 cats were diagnosed as LPE by the review panel, whereas 1 cat was diagnosed with SCL. Pseudoclonality was observed in 3 cases diagnosed as LPE and in 4 cases as SCL. For calculation of relative sensitivity, specificity, and accuracy, it was assumed that oligoclonality is indicative of clonality and therefore interpreted as a positive result for SCL. Pseudoclonality was assumed to be inconclusive or indeterminant. For the calculation of relative sensitivity, specificity, and accuracy, cases that were determined to be pseudoclonal were considered as incorrect. The 2 cases of the validation set for which PARR data could not be obtained were excluded from calculations. The sensitivity and specificity of PARR relative to the consensus panel diagnoses was found to be 85.7% (95% CI: 72.8%-98.7%) and 33.3% (95% CI: 14.5%-52.2%), respectively. The relative accuracy of PARR for the 52 cases of the validation set was calculated to be 61.5% (95% CI: 48.3%-74.8%).

Figure 4 shows plots of the positive (A) and negative (B) predictive values, respectively, for the HGMS and PARR methods as function of different disease prevalence.

4 | DISCUSSION

This is the first study investigated HGMS as a tool for the differentiation of LPE from SCL in FFPE duodenal biopsy specimens from cats with CE. Results showed that HMGS might be a suitable method for the detection of SCL in cats. Mass spectrometry imaging is a technology utilized for visualizing the spatial distribution of peptides and other molecules within a tissue sample. Mass spectrometry imaging has a wide array of applications, such as pharmacological studies, drug distribution,

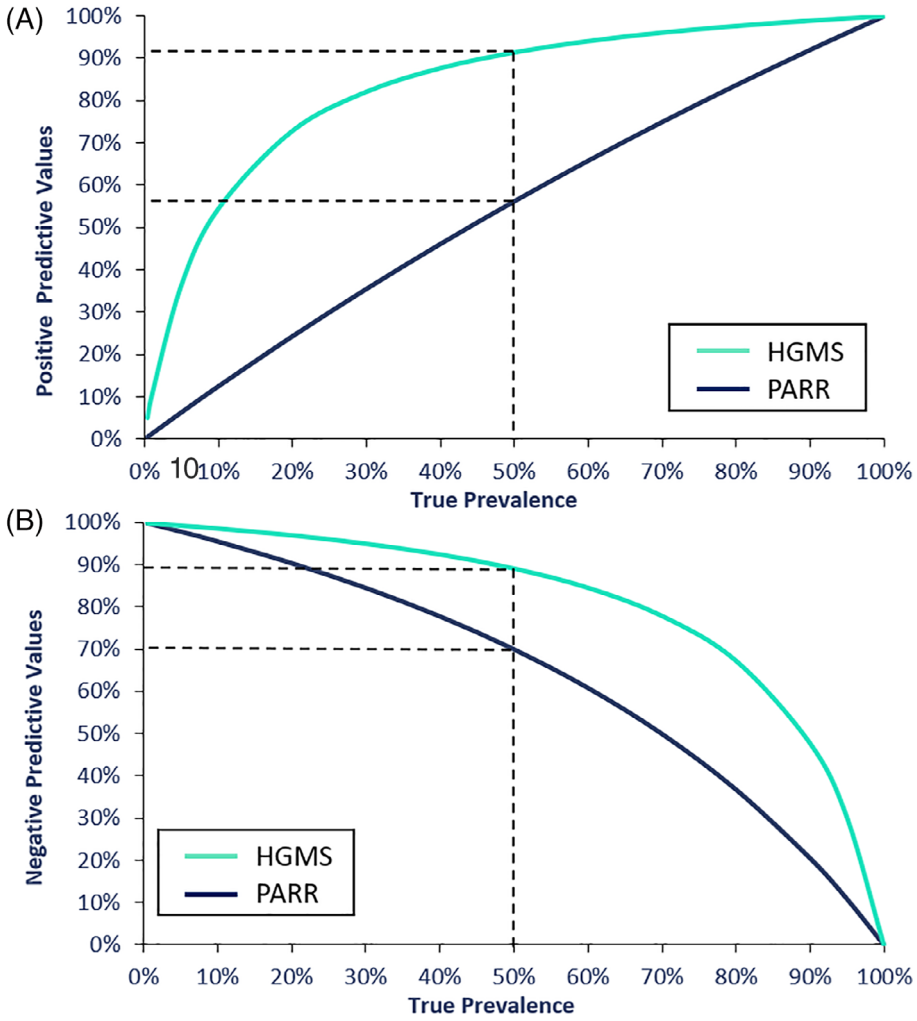

FIGURE 4 Positive (A) and negative (B) predictive values for the HGMS and PARR methods as function of different disease prevalence. HGMS, histology-guided mass spectrometry; PARR, polymerase chain reaction for antigen receptor rearrangements

histological studies, microbial identification, and biomarker discovery.[21,22,25] As a special application of MSI, HGMS integrates pathologist annotations within the mass spectrometry workflow to selectively look at histologically relevant cell subpopulations. This technology can be applied to both fresh frozen and FFPE tissues, and it is also suitable for tissue microarrays. Histology-guided mass spectrometry can be used in either a targeted manner, where the molecular target is known, such as in the differentiation of different protein isoforms or for pharmacological studies, or in an untargeted manner where the emphasis is on the detection of a mass spectral peak pattern consisting of differentially expressed molecules between samples.[19,26] This approach is often referred to as mass spectral fingerprinting. With the application of HGMS, mass spectral profiles are collected from each annotation site in a tissue sample and statistically assessed by a machine learning algorithm to identify differences in peak intensities that are indicative of specific molecules. An HGMS classification algorithm therefore relies on the signature mass spectral fingerprints instead of individual biomarkers for disease classification. This technology is used in human medicine, for the differentiation of benign nevi from malignant melanoma which, similar to cats with CE, can pose a diagnostic challenge for pathologists with a high interobserver variability.[20-22,27]

Stages of biomarker discovery commonly involve the discovery and identification of biomarkers within a training set, and the validation of the assay using an independent set of samples.[28,29] In the initial stages of assay development, samples should be well characterized. In the light of a lacking gold standard for the diagnosis and differentiation of LPE from SCL in cats, this study used the decisions of a panel of experts consisting of board-certified anatomic pathologists, internists, and oncologists. The members of the consensus panel were blinded as to the results of HGMS. Some of the consensus diagnoses served as a standard for selection of samples to comprise the training set that was used to develop the classification algorithm, along with previously classified samples. Also, the consensus diagnoses served as a benchmark against which relative test parameters were calculated in the validation set. This multidisciplinary approach is commonly used in human medicine, especially in the field of oncology, where tumor boards are considered good clinical practice and are recommended by the National Comprehensive Cancer Network guidelines.[30] Tumor boards consist of different experts who participate to provide balanced consensus opinions on the diagnostic and treatment plans for patients with complex cancer diagnoses.[31] A multidisciplinary approach is also explicitly recommended by the current EuroClonality/BIOMED-2 group,

who is dedicated to the standardization of clonality assays in human medicine.[16,17] The EuroClonality/BIOMED-2 group was founded in response to a high interassay and intraassay variability and in an attempt to provide analytical[17] as well as preanalytical and postanalytical standards for the conduction and interpretation of clonality assays.[16] However, despite these extensive efforts, clonality assays in human medicine have been shown to have a multitude of pitfalls[32] with a recent study reporting a specificity as low as 54.3% for the T-cell receptor assay.[7] There is currently no standardization of clonality assays in veterinary medicine. Clonality assays in healthy cats with similar demographic characteristics to cats with CE revealed that the specificity of clonality assays for the detection of SCL in veterinary medicine might be of equal concern.[6] Clonality assays can reveal positive results in the absences of malignancy in many instances, such as benign clonal expansions, pseudoclonality, and oligoclonality, and with false-positive results.[5,17] False-negative results can occur because of insufficient primer coverage or with a polyclonal (ie, inflammatory) background.[5,17] Therefore, the guidelines by the EuroClonality/ BIOMED-2 group state that results of molecular clonality studies should always be interpreted in the context of the clinical, morphological, and immunophenotypic diagnosis, and in close collaboration with different experts in the field.[17] These recommendations were reflected in the study design of the current study. Relative to the panel consensus diagnoses for the validation set cases, HGMS demonstrated a sensitivity of 86.7% (95% CI: 74.5%-98.8%), a specificity of 91.7% (95% CI: 80.6%-100%), and an overall accuracy of 88.9% (95% CI: 80.5%-97.3%).

When comparing results of HGMS and results of PARR analysis for the validation set cases, HGMS showed superior test parameters relative to the panel consensus. Relative to the panel consensus diagnoses for the validation set cases, PARR exhibited a sensitivity of 85.7% (95% CI: 72.8%-98.7%), a specificity of 33.3% (95% CI: 14.5%-52.2%), and an overall accuracy of 61.5% (95% CI: 48.3%-74.8%). Even though the true disease prevalence in a tested population of cats with CE is usually unknown, we calculated the positive and negative predictive values for HGMS and PARR as a function of true disease prevalence (Figure 4). Results showed that in this cohort of cats, HGMS generally showed more favorable results compared to PARR.

Our study has several limitations. Given the retrospective nature of the study, procedures around the collection of intestinal biopsies were not standardized. However, this reflects the true conditions under which a test for the differentiation of SCL from LPE has to perform. This concept is further supported by recent studies showing, that standardization is a major source of poor reproducibility in preclinical trials.[33,34]

Despite best efforts, follow-up information was not available for all cases. However, the median survival time of cats with SCL can be substantial,[35] hence follow-up information might only allow for very limited conclusions.

Even though tumor boards are considered to be the good clinical practice by the National Comprehensive Cancer Network as well as the EuroClonality/BIOMED-2 group, expert opinion is not equal to evidence, leading to a degree of uncertainty for results in this study. However, in the authors' opinion, this approach is the best possible compromise in the absence of a true gold standard for the differentiation of LPE and SCL.

Some samples were removed from the analysis because of a lack of consensus between the panelists. Any test for the differentiation of LPE from SCL is likely to be most useful using samples that are considered equivocal based on histopathology evaluations. Therefore, removing samples from the analysis that are equivocal may have led to biased results. However, the authors have undertaken every effort to mitigate this bias. Laboratory tests should always be developed on the full range of the disease spectrum and limiting the test development to a specific population can also severely bias test results.[36] This study included duodenal samples representative of the full disease spectrum of LPE and SCL. Every sample was assessed with all currently available diagnostic tests for the differentiation of LPE from SCL, namely, histopathology, IHC, and PARR. Compared to the routine analysis of histopathological samples, the expert panel took a substantial amount of time to review each case then took all available data into consideration to formulate a final diagnosis. However, in the face of a lacking gold standard, the panel could not come to a consensus for 18 out of 121 samples. Current PCR-based clonality assays in cats were developed and tested in cats with SCL compared to healthy cats,[37,31] instead of cats with LPE. Therefore, the authors believe that results of this study are more representative of the true disease spectrum and reflect genuine test superiority over currently available tests for the differentiation of LPE from SCL in cats.

Annotations and HGMS analyses were limited to the lamina propria of samples from the upper intestinal tract and thus data drawn from this study are applicable only to those locations. It is currently unknown, whether the same algorithm can be applied to other samples of the gastrointestinal tract, such as gastric or ileal samples. Similar to clonality assays, results of HGMS apply to the tested tissue only and do not exclude the presence of SCL in other parts of the gastrointestinal tract. Samples in which the epithelium is predominantly or exclusively affected or samples from other locations than the upper intestinal tract were not eligible for inclusion assessment as are samples from other locations of the small intestinal tract. However, we found the exclusive involvement of the epithelium to be very rare, and only 1 case was excluded from our data set because of this finding. Results of HGMS showed a small degree of overlap between groups. Within the validation set, 2 of 54 samples (3.7%) were classified as indeterminant.

Finally, samples used in this study were surplus FFPE samples stored in the tissue archive of the of the Texas A&M Gastrointestinal Laboratory. With increasing storage time, DNA quality decreases[38] and might have affected results of the clonality assay.

ACKNOWLEDGMENT

A portion of these data was presented at the 2018 American College of Veterinary Pathology meeting in Washington, District of Columbia and the 2019 American College of Veterinary Internal Medicine Forum, Phoenix, Arizona.

CONFLICT OF INTEREST DECLARATION

Marsilio, Lidbury, Suchodolski, and Steiner are or were employed by the Gastrointestinal Laboratory at Texas A&M University, at the time of the study, which offers laboratory testing, including histopathology services, on a fee-for-service basis. Smoot, Seeley, and Powell are or were employed by New River VDL, LLC at the time of study. New River VDL, LLC offers a commercially available histology guided mass spectrometry assay for cats on a fee-for-service basis. Steiner serves as a paid consultant for New River VDL. Estep is employed by Texas Veterinary Pathology, LLC, which offers histopathology for animals on a fee-for-service basis. Newman, Warry, Flory, Giaretta, and Morley have no conflicts to disclose.

OFF-LABEL ANTIMICROBIAL DECLARATION

Authors declare no off-label use of antimicrobials.

INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC) OR OTHER APPROVAL DECLARATION

This study used surplus archived tissue specimens for which no IACUC approval was deemed necessary.

HUMAN ETHICS APPROVAL DECLARATION

Authors declare human ethics approval was not needed for this study.

ORCID

Sina Marsilio https://orcid.org/0000-0002-0693-0669
Paula R. Giaretta https://orcid.org/0000-0003-2261-281X
Jonathan A. Lidbury https://orcid.org/0000-0001-5107-4577
Matthew J. Powell https://orcid.org/0000-0002-3334-4366
Jörg M. Steiner https://orcid.org/0000-0003-3336-2086

REFERENCES

1. Richter KP. Feline gastrointestinal lymphoma. *Vet Clin North Am Small Anim Pract*. 2003;33:1083-1098. vii.
2. Sabattini S, Bottero E, Turba ME, Vicchi F, Bo S, Bettini G. Differentiating feline inflammatory bowel disease from alimentary lymphoma in duodenal endoscopic biopsies. *J Small Anim Pract*. 2016;57:396-401.
3. Jergens AE. Feline idiopathic inflammatory bowel disease: what we know and what remains to be unraveled. *J Feline Med Surg*. 2012;14:445-458.
4. Moore PF, Rodriguez-Bertos A, Kass PH. Feline gastrointestinal lymphoma: mucosal architecture, immunophenotype, and molecular clonality. *Vet Pathol*. 2012;49:658-668.
5. Keller SM, Vernau W, Moore PF. Clonality testing in veterinary medicine: a review with diagnostic guidelines. *Vet Pathol*. 2016;53:711-725.
6. Marsilio S, Ackermann MR, Lidbury JA, Suchodolski JS, Steiner JM. Results of histopathology, immunohistochemistry, and molecular clonality testing of small intestinal biopsy specimens from clinically healthy client-owned cats. *J Vet Intern Med*. 2019;33:551-558.
7. Kokovic I, Novakovic BJ, Cerkovnik P, et al. Clonality analysis of lymphoid proliferations using the BIOMED-2 clonality assays: a single institution experience. *Radiol Oncol*. 2014;48:155-162.
8. Paulin MV, Couronne L, Beguin J, et al. Feline low-grade alimentary lymphoma: an emerging entity and a potential animal model for human disease. *BMC Vet Res*. 2018;14:306.
9. Bruggemann M, White H, Gaulard P, et al. Powerful strategy for polymerase chain reaction-based clonality assessment in T-cell malignancies report of the BIOMED-2 concerted action BHM4 CT98-3936. *Leukemia*. 2007;21:215-221.
10. Ponti R, Fierro MT, Quaglino P, et al. TCRgamma-chain gene rearrangement by PCR-based GeneScan: diagnostic accuracy improvement and clonal heterogeneity analysis in multiple cutaneous T-cell lymphoma samples. *J Invest Dermatol*. 2008;128:1030-1038.
11. Chen YL, Su IJ, Cheng HY, et al. BIOMED-2 protocols to detect clonal immunoglobulin and T-cell receptor gene rearrangements in B- and T-cell lymphomas in southern Taiwan. *Leuk Lymphoma*. 2010;51:650-655.
12. Keller SM, Moore PF. A novel clonality assay for the assessment of canine T cell proliferations. *Vet Immunol Immunopathol*. 2012;145:410-419.
13. Kim Y, Choi YD, Choi C, Nam JH. Diagnostic utility of a clonality test for lymphoproliferative diseases in Koreans using the BIOMED-2 PCR assay. *Korean J Pathol*. 2013;47:458-465.
14. Waugh EM, Gallagher A, Haining H, et al. Optimisation and validation of a PCR for antigen receptor rearrangement (PARR) assay to detect clonality in canine lymphoid malignancies. *Vet Immunol Immunopathol*. 2016;182:115-124.
15. Hammer SE, Groiss S, Fuchs-Baumgartinger A, et al. Characterization of a PCR-based lymphocyte clonality assay as a complementary tool for the diagnosis of feline lymphoma. *Vet Comp Oncol*. 2016;15:1354-1369.
16. Langerak AW, Groenen PJ, Bruggemann M, et al. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. *Leukemia*. 2012;26:2159-2171.
17. van Dongen JJ, Langerak AW, Bruggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 concerted action BMH4-CT98-3936. *Leukemia*. 2003;17:2257-2317.
18. Langerak AW. Toward standardization of clonality testing in veterinary medicine. *Vet Pathol*. 2016;53:705-706.
19. Seeley EH, Caprioli RM. MALDI imaging mass spectrometry of human tissue: method challenges and clinical perspectives. *Trends Biotechnol*. 2011;29:136-143.
20. Lazova R, Seeley EH, Kutzner H, et al. Imaging mass spectrometry assists in the classification of diagnostically challenging atypical Spitzoid neoplasms. *J Am Acad Dermatol*. 2016;75:1176-1186.e1174.
21. Kriegsmann J, Kriegsmann M, Casadonte R. MALDI TOF imaging mass spectrometry in clinical pathology: a valuable tool for cancer diagnostics (review). *Int J Oncol*. 2015;46:893-906.
22. Aichler M, Walch A. MALDI imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Lab Invest*. 2015;95:422-431.
23. Washabau RJ, Day MJ, Willard MD, et al. Endoscopic, biopsy, and histopathologic guidelines for the evaluation of gastrointestinal inflammation in companion animals. *J Vet Intern Med*. 2010;24:10-26.
24. Day MJ, Bilzer T, Mansell J, et al. Histopathological standards for the diagnosis of gastrointestinal inflammation in endoscopic biopsy samples from the dog and cat: a report from the World Small Animal Veterinary Association Gastrointestinal Standardization Group. *J Comp Pathol*. 2008;138(Suppl 1):S1-S43.
25. Hou TY, Chiang-Ni C, Teng SH. Current status of MALDI-TOF mass spectrometry in clinical microbiology. *J Food Drug Anal*. 2019;27:404-414.
26. Seeley EH, Schwamborn K, Caprioli RM. Imaging of intact tissue sections: moving beyond the microscope. *J Biol Chem*. 2011;286:25459-25466.
27. Farmer ER, Gonin R, Hanna MP. Discordance in the histopathologic diagnosis of melanoma and melanocytic nevi between expert pathologists. *Hum Pathol*. 1996;27:528-531.

28. Vafaee F, Diakos C, Kirschner MB, et al. A data-driven, knowledge-based approach to biomarker discovery: application to circulating microRNA markers of colorectal cancer prognosis. *NPJ Syst Biol Appl*. 2018;4:20.
29. Swan AL, Mobasheri A, Allaway D, Liddell S, Bacardit J. Application of machine learning to proteomics data: classification and biomarker identification in postgenomics biology. *OMICS*. 2013;17:595-610.
30. NCCN. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) Esophageal and Esophagogastric Junction Cancers. Version 1.2014*. 2014. Available from: https://www.spg.pt/wp-content/uploads/Guidelines/NCCN/esophageal.pdf.
31. Schmidt HM, Roberts JM, Bodnar AM, et al. Thoracic multidisciplinary tumor board routinely impacts therapeutic plans in patients with lung and esophageal cancer: a prospective cohort study. *Ann Thorac Surg*. 2015;99:1719-1724.
32. Groenen PJ, Langerak AW, van Dongen JJ, et al. Pitfalls in TCR gene clonality testing: teaching cases. *J Hematop*. 2008;1:97-109.
33. Voelkl B, Vogt L, Sena ES, Würbel H. Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLoS Biol*. 2018;16:e2003693.
34. Karp NA. Reproducible preclinical research-is embracing variability the answer? *PLoS Biol*. 2018;16:e2005413.
35. Pope KV, Tun AE, McNeill CJ, et al. Outcome and toxicity assessment of feline small cell lymphoma: 56 cases (2000-2010). *Vet Med Sci*. 2015;1:51-62.
36. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med*. 1978;299:926-930.
37. Moore PF, Woo JC, Vernau W, Kosten S, Graham PS. Characterization of feline T cell receptor gamma (TCRG) variable region genes for the molecular diagnosis of feline intestinal T cell lymphoma. *Vet Immunol Immunopathol*. 2005;106:167-178.
38. Watanabe M, Hashida S, Yamamoto H, et al. Estimation of age-related DNA degradation from formalin-fixed and paraffin-embedded tissue according to the extraction methods. *Exp Ther Med*. 2017;14:2683-2688.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Marsilio S, Newman SJ, Estep JS, et al. Differentiation of lymphocytic-plasmacytic enteropathy and small cell lymphoma in cats using histology-guided mass spectrometry. *J Vet Intern Med*. 2020;34:669–677. https://doi.org/10.1111/jvim.15742