

IMPROVED CHARACTERIZATION OF SOYBEAN MEAL-INDUCED ENTERITIS USING MACHINE LEARNING AUTOMATION AND STANDARDIZATION TO SCORE DISTAL INTESTINAL HISTOLOGY SLIDES

Principal Investigator: Jacob Bledsoe, Ph.D.
University of Idaho, Aquaculture Research Institute
Hagerman Fish Culture Experiment Station

PURPOSE OF THE STUDY

Trout are economically important finfish. Soybean-derived feed ingredients can induce inflammatory responses in the distal intestine of salmonids such as rainbow trout (*Oncorhynchus mykiss*), called soybean meal-induced enteritis (SBMIE).

Distal intestinal histopathology is the gold standard for assessing enteritis severity. This assessment is labor-intensive and difficult to scale, as it has traditionally required specialist time and expense to perform. But recent advances in digital pathology and machine learning have created new opportunities to automate and streamline these image interpretation tasks.

OBJECTIVES

In an effort to more rapidly and accurately evaluate intestinal health related to soy inclusion rates in commercial feeds, a machine-learning framework for automated analysis of rainbow trout distal intestinal histology was developed by:

- ▶ Classifying enteritis status from digitized whole-slide images to identify the most strongly associated histological features
- ▶ Evaluating the feasibility of automated prediction of individual traits included in a standardized scoring rubric
- ▶ Building a scalable and repeatable computational workflow that can support future model development for applications in feed ingredient evaluation, selective breeding for plant ingredient tolerance, and fish health assessment

STUDY DESIGN

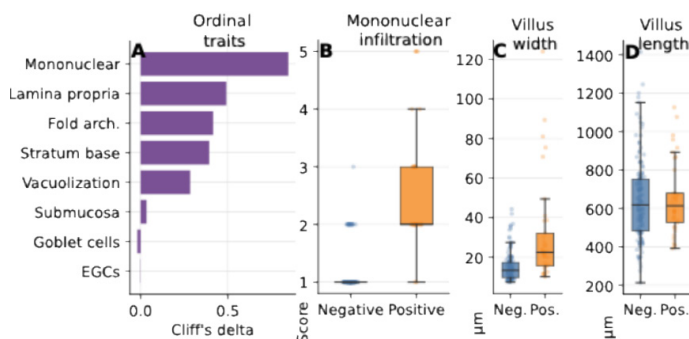
Study design was framed to use 3 independent rainbow trout distal intestinal histology datasets available through internal University of Idaho research collaborations, consisting of 182 hematoxylin and eosin stained whole-slide images. These studies included a range of dietary treatments expected to induce varying degrees of intestinal inflammation and enteropathy.

Whole-slide images were initially reviewed to identify non-representative or technically deficient samples, such as: severe tissue damage, incomplete intestinal cross-sections, excessive folding, poor staining quality, or inadequate

representation of distal intestinal tissue. These were excluded from downstream analyses. Following quality control screening, slides were digitized, scored by an independent aquatic veterinary histopathologist, and assembled into a unified training dataset for machine-learning analysis.

A standardized scoring rubric was developed specifically for this project by integrating and harmonizing features from previously published salmonid enteritis scoring systems. Statistical analyses were conducted to identify histological features most strongly associated with enteritis status.

Mononuclear inflammatory cell infiltration (MICI) demonstrated the strongest association with enteritis status, so it was prioritized as the primary candidate trait for model training, along with secondary additional histological features determined to be closely associated with enteritis.



Histopathological feature screening. (A) Ordinal traits ranked by effect size, led by mononuclear infiltration. (B) Mononuclear infiltration scores were higher in enteritis-positive samples. (C) Quantitative villus width and (D) villus length measurements showed broad overlap between enteritis-negative and -positive samples, suggesting these morphometric features were less discriminating than mononuclear infiltration in this dataset, though villus width did show some separation by enteritis class.

RESULTS

The 182 retained analysis units that passed initial quality control steps were partitioned into 136 training units (10,186 tiles) and 46 validation units (3,443 tiles). Three machine-learning objectives were evaluated:

- ▶ **Binary Enteritis Classification:** Slides were classified as either enteritis-negative or -positive
- ▶ **Ordinal Mononuclear Infiltration Classification:** Separate models were trained to predict ordinal severity scores for individual histological traits, with emphasis on MICI
- ▶ **Binary Mononuclear Infiltration Classification:** MICI scores were sorted into 2 biologically meaningful categories — low (scores 1-2), and moderate-to-severe inflammation (3-5)

Summary of model performance metrics.

MODEL	CLASSIFICATION TASK	SELECTED CHECKPOINT	VALIDATION PERFORMANCE
BINARY ENTERITIS (RESNET18)	Enteritis Negative vs. Positive	Epoch 1	Accuracy = 69.6% Sensitivity = 100.0% Specificity = 62.2% Precision = 39.1% F1 = 0.563 AUROC = 0.859 AUPRC = 0.589
ORDINAL MONONUCLEAR INFILTRATION (RESNET18)	Scores 1-5	Epoch 10	Accuracy = 58.7% Macro F1 = 0.302 Weighted F1 = 0.605
BINARY MONONUCLEAR INFILTRATION (RESNET18)	Scores 1-2 vs. 3-5	Epoch 12	Accuracy = 43.5% Sensitivity = 100.0% Specificity = 38.1% Precision = 13.3% F1 = 0.235 AUROC = 0.899 AUPRC = 0.798

The model correctly identified all enteritis-positive validation samples, and the absence of false-negative predictions demonstrates the utility of the approach as a conservative screening tool. From an aquafeed development perspective, a model that minimizes missed cases of intestinal inflammation is preferable to one that sacrifices sensitivity to reduce false-positive predictions.

Collectively, these results provide proof of principle that weakly supervised deep learning can identify biologically meaningful inflammatory signatures from digitized rainbow trout distal intestinal histology, and support development of scalable computational tools for future enteropathy assessment.

Over decades, histopathological scoring systems have evolved to commonly evaluate traits such as mucosal fold architecture, supranuclear vacuolization, lamina propria expansion, connective tissue accumulation, and more to determine SBMIE. The present study found MICI exhibited substantially stronger associations with enteritis status than any other evaluated variable.

STUDY LIMITATIONS & NEXT STEPS

Some limitations to this study should be considered in future evaluation and refinement of machine-learning models. The most significant was dataset size — logistical barriers with respect to slide availability, digitization, data-sharing restrictions, and study compatibility prevented incorporation of multi-institutional datasets of slides from Atlantic salmon and rainbow trout spanning a broader range of dietary treatments and enteritis severity.

Other limitations included, but were not limited to:

- ▶ Severe class imbalance within the dataset, with relatively few samples exhibiting advanced inflammatory enteritis lesions, substantially limiting the ability of ordinal models to learn distinctions among higher-severity categories
- ▶ That only one histopathologist reviewed scores — multiple independent scorers would provide opportunities to quantify observer agreement and train models more robust to differences in interpretation

Despite these, this study establishes a strong foundation for future development of automated histopathology tools for aquaculture. The ability of the binary enteritis classifier to achieve perfect sensitivity at the analysis-unit level is particularly encouraging for further development and training of the algorithm.

Additional digitized slides representing a wider range of salmonids, diet, and enteritis severity would improve model robustness; emphasis should be placed on increasing representation of moderate and severe inflammatory phenotypes. Inclusion of Atlantic salmon (*Salmo salar*) datasets would also be valuable, given the extensive literature on SBMIE in that species, and would facilitate evaluation of model transferability across salmonids.

Future studies should incorporate multiple histopathologists during rubric development and scoring, and improvements in slide preparation and annotation protocols may also further improve model performance. Computationally, several alternative approaches warrant further investigation as well.

BENEFIT FOR THE SOYBEAN FARMER

This model represents a desirable tool that could serve as a first-pass screen that rapidly identifies samples most likely to contain SBMIE, allowing histopathologists to focus their review efforts on high-priority samples. This can provide a foundation for development of scalable computational tools capable of supporting future aquaculture nutrition.

This has the potential to accelerate evaluation of soy-based ingredients and nutritional strategies that support their greater use in aquafeeds, by reducing the cost and turnaround time required for gold-standard intestinal histopathological assessments of fish in order to make alterations to feeding.

SCAN FOR
PROJECT
DETAILS

