Table 1. Statistical Analysis Fram

Complete analytical workflow from laboratory development through field va

Analysis Phase	Research Question	Model ID	Statistical Method
Laboratory Development			
Discovery	Can immune genes predict infection costs?	**DISC-1**	Linear regression (PC1, PC2 → weig loss)
Optimization	Can machine learning improve prediction?	**DISC-2**	Random forest (19 genes → weight loss)
Validation	Is the model reliable?	**DISC-3**	Train-test cross-validation
Cross-Population Translation			
Gene Validation	Which genes show consistent responses across populations?	**TRANS-1**	Linear regression per gene (lab vs field)
Field Translation			
Detection	Does the model work in wild populations?	**FIELD-1**	Predicted vs. observed infection status
Discrimination	Can it distinguish parasite species?	**FIELD-2**	Predicted loss by species identity
Scaling	Does it correlate with infection severity?	**FIELD-3**	Predicted loss vs. parasite load
Biological Validation			
Physiological relevance	Does it capture real health impacts?	**PROOF-1**	Predicted loss vs. body condition
Specificity	Is the response Eimeria-specific?	**PROOF-2**	Predicted loss vs. parasite communi

Framework demonstrates progression from basic linear prediction (R^2 = 0.106) through machine learning optir relevance. Cross-population translation validates 3/19 genes as conserved biomarkers.

Table 1. Statistical Analysis Fram

Complete analytical workflow from laboratory development through field va

Analysis Phase	Research Question	Model ID	Statistical Method

[1] Significance levels: *p < 0.05, **p < 0.01, ***p < 0.001

[2] Train-test validation used 70% training, 30% testing from full dataset

[3] Cross-validated genes: CXCL9 (both species), TICAM1, PRF1 (E. falciformis)

[4] E.f: Eimeria falciformis; E.r: E. ferrisi

Framework demonstrates progression from basic linear prediction (R^2 = 0.106) through machine learning optir relevance. Cross-population translation validates 3/19 genes as conserved biomarkers.