Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Toward genetics-informed control of Border disease in goats: a cross-model machine learning consensus screen for candidate SNPs

**Yalçın YAMAN**

yalcinyaman@gmail.com

Siirt Üniversitesi

**Burcu TOKGÖZ**

Siirt Üniversitesi

**Semih YAZICI**

Tarımsal Araştırmalar ve Politikalar Genel Müdürlüğü

**Research Article**

**Keywords:**

**Posted Date:** September 5th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10642811/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

**Additional Declarations:** No competing interests reported.

## Abstract

Border disease virus (BDV) causes substantial reproductive and economic losses in small ruminants; however, host genetic factors influencing susceptibility have not been examined in goats, nor in any other farm animal species. We investigated whether a detectable host genomic signal for BDV serostatus exists in goats and developed a leakage-safe machine learning framework to prioritize candidate SNPs. Serostatus was determined by ELISA in 442 goats (117 seropositive, 325 seronegative) genotyped at 44,375 quality-controlled SNPs. Within a nested cross-validation design with strictly in-fold feature selection and hyperparameter tuning, eight machine learning models spanning penalized-linear, kernel, tree-ensemble, genomic best linear unbiased prediction, and neural-network families were trained and combined through a performance-weighted Super Learner. Out-of-fold discrimination was modest but consistent across learners (AUC 0.817–0.854), and the stacked ensemble did not exceed the best individual learner. A cross-model consensus score, aggregating importance across learners with performance-based weights, prioritized candidate SNPs for downstream testing; permutation-based significance testing of the top-ranked candidates identified 39 SNPs meeting a false-discovery-rate threshold of $q < 0.05$. These results indicate a genuine, broadly distributed genomic signal for BDV serostatus in goats and provide a methodologically conservative, ranked set of candidate genomic regions. The findings are hypothesis-generating rather than diagnostic or causal, and warrant functional validation and replication in independent, larger cohorts.

## 1. Introduction

Border disease (BD) was first recognized as a distinct clinicopathological entity in sheep in the border counties of England and Wales in 1959, from which the disease takes its name [1], and is also known as "hairy shaker disease" for the excessively hairy birth-coat and rhythmic tremor typical of affected neonates [2, 3]. Its causative agent, border disease virus (BDV), is a single-stranded, positive-sense RNA virus of the genus Pestivirus, family Flaviviridae [1, 4, 5] closely related to bovine viral diarrhea virus (BVDV-1 and − 2) of cattle and classical swine fever virus (CSFV) of pigs [6–8]; the resulting antigenic and genetic overlap among these pestiviruses has direct consequences for diagnosis, surveillance, and control [9]. Sheep are the principal host [8, 10], showing the full clinical spectrum of disease most commonly, whereas clinical disease and virus recovery are both reported less often in goats, where presentation is dominated by abortion and poor kid viability [10]. BDV is nonetheless not strictly host-specific: it has also been isolated from cattle, pigs, deer, chamois, and wild boar [9, 11], and interspecies transmission among sheep, cattle, and pigs occurs regularly [9]. Clinical outcome depends critically on the gestational timing of exposure: infection of immunocompetent adults is generally mild or subclinical, whereas infection of naive pregnant females before fetal immunocompetence is established can cause embryonic death, abortion, stillbirth, or the birth of persistently infected (PI) offspring [12–14]. The PI state arises because the fetus, exposed to the virus before its immune system matures, becomes immunotolerant rather than mounting a protective response [10]; such animals remain seronegative while shedding virus continuously throughout life, making them the principal reservoir sustaining transmission within and between flocks [9, 10, 14].

In goats specifically, BDV infection has now been confirmed in herds in Saudi Arabia, Italy, China, Turkey, and Iraq [4, 7, 15–18], consistent with the virus's broader confirmed distribution across Europe, Australia, New Zealand, North America, India, Turkey, Japan, and China [7, 9]. Reported prevalence of BDV in sheep populations varies widely, from 5% to 90% depending on the region surveyed [8], and herd- and species-level estimates illustrate this heterogeneity: 8.3% herd-level and 50% within-herd individual prevalence in Denmark [19]; 16.1% in sheep and 25.4% in goats in Switzerland [20]; 11.5% individual and 31.3% flock-level prevalence in Austrian goats [21]; 53.7% seropositivity for BVDV/BDV antibodies among sheep and goats combined in Kazakhstan [22]; and, in Mosul, Iraq, an RT-PCR prevalence of 15.9% in sheep and 3% in goats [4], against a substantially higher indirect-ELISA seroprevalence of 46.9% in sheep and 16% in goats [23]. This variability reflects identifiable risk factors rather than chance. The introduction of undetected PI animals is considered the single most important source of new infection on previously unaffected farms [3, 10]; horizontal contact spreads the virus within a flock, while vertical transmission across the placenta remains the principal route by which new PI animals are generated [7, 9]. Co-location with cattle substantially elevates risk: herd prevalence in Denmark reached 24% on farms keeping cattle, versus 4.2% on farms without [19], and communal alpine grazing has independently been linked to pestivirus spread among sheep and goats in Austria and Switzerland [21, 24]. Uncontrolled animal movement, untested purchases, shared grazing land, and the absence of structured vaccination or surveillance programs further compound the risk of introduction and onward spread in naive flocks [4, 14, 25].

Despite this wide and expanding distribution, BD remains difficult to diagnose, treat, and prevent, and these difficulties carry direct practical consequences. Clinical signs are not pathognomonic and are readily confused with other causes of reproductive failure and neonatal loss in small ruminants [6, 26], while the close antigenic relationship among BDV, BVDV, and CSFV means that standard serological and molecular assays can fail to reliably distinguish between them, creating a persistent risk of misdiagnosis [9, 11]. PI animals, which drive transmission, are frequently indistinguishable from healthy stock on clinical grounds alone and can only be reliably identified through laboratory testing [14], a diagnostic burden further compounded by the ongoing emergence of novel, atypical ovine pestiviruses [7, 27, 28]. No specific antiviral treatment exists for BDV infection, so control depends almost entirely on the identification and removal of PI animals and on shielding susceptible pregnant females from exposure, particularly during the first half of gestation, through closed-flock management and pre-

purchase testing [3, 6]. Vaccination, in principle the most direct means of protecting pregnant females, remains an inadequate solution in practice: commercial BDV vaccines are few and licensing is limited [6], BVDV vaccines developed for cattle are not recommended for sheep because circulating BDV strains are antigenically distinct [29], and even BVDV vaccination has been shown not to protect against fetal infection following experimental BDV challenge [30]. The absence of a proven, effective vaccine remains a major unresolved gap in the BD control toolkit [10, 26]. These limitations extend beyond small-ruminant health: because sheep and goats can act as a BDV reservoir for co-grazed or co-housed cattle, unresolved BD in small-ruminant flocks can directly undermine national BVDV eradication programs, with shared housing of cattle and small ruminants identified as the most significant risk factor for BDV introduction into Swiss cattle herds [31], giving surveillance and control of BDV in sheep and goats a role that extends beyond the small-ruminant sector itself [3, 31].

These diagnostic and control limitations translate directly into substantial economic losses for sheep and goat production systems worldwide [4, 9]. Costs accrue from multiple, compounding sources, including infertility and barren females, abortion and stillbirth, high mortality among weak or malformed lambs and kids, reduced growth performance in surviving PI animals, and the recurring costs of diagnostic testing and biosecurity [7, 26]. Because losses arise through both prenatal and postnatal infection pathways, their cumulative impact can be severe at the individual flock level, at times approaching disaster proportions for affected producers [9, 32]. Beyond its direct costs to small-ruminant production, BD imposes an indirect economic burden on the cattle sector through its capacity to interfere with BVDV eradication programs, a concern that has renewed international interest in BD surveillance and control [3, 31].

Despite this extensive characterization of BDV epidemiology, transmission, and the practical limitations of current diagnosis and control, host genetic factors that might explain differential susceptibility or resistance to BDV infection have not, to our knowledge, been examined in sheep or goats, and goat-specific data are particularly absent given that clinical presentation in this species already appears to differ from that documented in sheep [10]. Given the persistent gaps in vaccination efficacy and clinical detection detailed above, host genomic information represents a plausible complementary strategy for disease control — one that, in principle, does not require case-by-case pathogen detection and could offer a heritable, cumulative reduction in disease burden across generations through selective breeding. In the present study, we addressed this gap using a cohort of 442 goats with BDV serostatus determined serologically by ELISA (117 seropositive, 325 seronegative; full serological characterization reported in a companion manuscript, Yaman et al., in preparation) and genome-wide SNP genotyping (44,375 markers after quality control). We developed a leakage-safe, nested cross-validation framework spanning eight machine-learning models from distinct methodological families — penalized-linear, kernel-based, tree-ensemble, genomic best linear unbiased prediction, and neural-network approaches — combined through a performance-weighted Super Learner, and used a cross-model consensus scoring procedure together with permutation-based significance testing to prioritize candidate SNPs whose importance for BDV serostatus was supported convergently across this heterogeneous model panel. Rather than aiming to build a diagnostic or individual-risk-prediction tool, our objective was to establish whether a detectable host genomic signal for BDV serostatus exists in goats and, if so, to generate a methodologically conservative, ranked set of candidate genomic regions to guide future functional validation and genetics-informed approaches to BD control in this species.

## 2. Materials and Methods

# 2.1 Study population and sample composition

The study population comprised 442 goats sampled from seven breeds maintained in Türkiye: Kilis (n = 78), Colored Angora (n = 21), Damascus (n = 62), Hairy Goat (Kıl; n = 62), Honamlı (n = 63), Turkish Saanen (n = 76), and Maltese (n = 80). Each animal was classified as seropositive or seronegative for Border disease virus (BDV) on the basis of the ELISA described in Section 2.2, yielding 117 seropositive (case) and 325 seronegative (control) animals overall. Breed-stratified seroprevalence estimates and their epidemiological interpretation are reported separately in a companion manuscript (Yaman et al., in preparation) and are not analyzed here, where breed composition is used only to characterize the analysis sample.

# 2.2 Serological diagnosis (ELISA)

BDV serostatus was determined using a commercial antigen-capture enzyme-linked immunosorbent assay (ID Screen® BVD p80 Antigen Capture kit; IDvet, Grabels, France) according to the manufacturer's instructions. After the prescribed incubation and washing steps, optical density (OD) was read at 450 nm. For each sample, the sample-to-positive percentage was calculated as

S/P% = (OD_sample − OD_negative control) / (OD_positive control − OD_negative control) × 100.

Following the manufacturer's recommendation, samples with S/P% < 35 were classified as negative and samples with S/P% ≥ 35 as positive.

# 2.3 Genotyping and quality control

Genotype data were processed in PLINK v1.9 [33]. Quality control retained samples and variants using the following filters: per-sample missingness < 0.1 (--mind 0.1), minor allele frequency ≥ 0.05 (--maf 0.05), per-variant missingness < 0.1 (--geno 0.1), and Hardy–Weinberg equilibrium $p \geq 1 \times 10^{-6}$ (--hwe 1e-6). After quality control, 44,375 autosomal SNPs remained with no missing calls and no monomorphic markers. Missing genotypes were imputed with Beagle v5 [34]. All QC-passed SNPs were carried forward to the modelling stage without any prior dimensionality reduction or feature pre-screening; the full QC-passed panel was supplied directly to the nested cross-validation pipeline.

## 2.4 Nested cross-validation and model training

Model evaluation and hyperparameter optimization were embedded in a nested cross-validation (CV) design to obtain performance and importance estimates that are not contaminated by information leakage from held-out samples. The outer loop used 5-fold stratified CV to partition animals into training and test sets, and within each outer training set a 4-fold stratified inner loop was used for hyperparameter tuning. To reduce the influence of a particular random partitioning, the entire nested procedure was repeated across 3 independent random seeds, and all downstream summaries were aggregated over the resulting 15 outer folds per model.

Two principles were enforced to avoid the selection bias that arises when any data-driven step is performed on the full data set before evaluation. First, all preprocessing (including standardization and any in-fold feature filtering) was fitted only on the outer training partition and then applied to the corresponding test partition. Second, hyperparameters were selected using only the inner CV within each outer training set, so that the outer test partition was never used during model selection. Both requirements follow directly from the finding that feature selection and hyperparameter tuning must be external to the performance estimate; otherwise cross-validated error is optimistically biased, sometimes severely so [35, 36]. Because class labels were imbalanced, class imbalance was addressed during training using a focal-loss objective for the neural models, which down-weights the loss contributed by well-classified majority examples and focuses learning on harder minority examples [37], together with class-balanced resampling. Hyperparameters were tuned with the Tree-structured Parzen Estimator as implemented in Optuna [38].

A deliberately heterogeneous set of base learners spanning distinct model families was used, so that the downstream cross-model agreement reflects convergence across different inductive biases rather than repetition of a single algorithm. The base learners comprised three gradient-boosted decision-tree implementations — XGBoost [39], LightGBM [40] and CatBoost [41] — a random forest [42], an L1-penalized logistic regression, a genomic best linear unbiased prediction (GBLUP) model [43] with restricted maximum-likelihood variance-component estimation [44], a kernel ridge regression, and a residual multilayer perceptron. Standard estimators and utilities were used through scikit-learn [45].

## 2.5 Base-learner variable importance

Within every outer fold, a per-SNP variable-importance vector was computed for each base learner using a model-appropriate measure: impurity- and split-gain-based importance for the tree ensembles, absolute standardized coefficients for the penalized linear model, and coefficient-magnitude importance for the kernel- and BLUP-based models; permutation-based importance was used for the neural model and for any learner where a native importance measure was not directly available. Importance was therefore estimated only on training data within each fold, consistent with the leakage-safe design in Section 2.4.

## 2.6 Cross-model importance aggregation

To combine importance signals across learners that operate on different scales, each fold-level importance vector was first rank-transformed to a uniform [0, 1] scale (average ranks divided by the number of ranked features), which makes the distributions comparable across models without assuming a common parametric form. For each base learner, the fold-level normalized vectors were then summarized into five per-SNP "views": (i) the mean normalized native importance across folds; (ii) the mean normalized permutation importance across folds (set equal to the native view when a separate permutation estimate was unavailable); (iii) a selection-stability view, defined as the proportion of folds in which the SNP ranked among that learner's top 100 features, in the spirit of stability selection [46]; (iv) a rank view given by the mean normalized Borda position across folds; and (v) an uncertainty view given by the standard deviation of the normalized importance across the seed-by-fold repeats. Because SNPs may be filtered differently across folds, views were assembled over the union of all features seen in any fold, with folds that did not include a given SNP contributing no value for that SNP. We note that the uncertainty view is computed from cross-validation repeat variability rather than a literal bootstrap resampling procedure, and it was reported separately rather than entering the consensus score.

## 2.7 Ensemble integration and performance-weighted consensus scoring

The out-of-fold prediction matrix (samples × base learners) was combined with a Super Learner [47]: a non-negative, L2-regularized logistic combiner fitted against the observed serostatus, with the regularization strength selected by inner cross-validation and the resulting weights normalized to sum to one. The Super Learner served two distinct purposes. First, it yielded a single stacked out-of-fold predictive score, whose discrimination was evaluated on the same footing as the individual base learners (reported in Section 3.4); because this score is a

leakage-safe combination of out-of-fold predictions, it estimates the discrimination attainable by pooling the learners rather than the performance of any single one. Second, the normalized combiner weights were used as performance-based reliability weights in the consensus scoring described below, so that learners predicting serostatus more accurately contributed more to the aggregated SNP importance.

A single consensus importance score was then computed for each SNP as the weighted geometric mean, across learners, of that learner's rank-normalized importance views (native and permutation importance, a selection-stability view, and a rank view), with fixed view weights (1.0, 1.0, 0.5 and 0.5, respectively) each multiplied by the corresponding learner's Super Learner weight; when the combiner could not be fitted, learners were weighted equally. We emphasize that the per-SNP signal entering the consensus derives from the importance views, not from the Super Learner's predictions: the Super Learner contributes only the relative weighting of learners, and the consensus is defined even when it is replaced by equal weights.

This consensus score is intended as a cross-model prioritization statistic for ranking candidate SNPs and is not a calibrated test of association. Two properties should be kept in mind. First, because the weighting is performance-based, a SNP influential in the better-predicting learners is favoured. Second, all learners are trained on the same genotype matrix, so agreement across learners is not equivalent to agreement across statistically independent lines of evidence and may partly reflect shared structure (for example, linkage disequilibrium) in the data [48]. To make the degree of concentration explicit, the dispersion of the Super Learner weights was quantified with a Herfindahl–Hirschman concentration index and a normalized entropy, and a strong concentration onto a single learner was reported as such, because in that situation both the stacked predictor and the consensus reduce towards that single learner.

## 2.8 Significance assessment

Because the consensus score is a ranking statistic, formal significance was assessed separately, and only for the most highly ranked candidates, by a phenotype-label permutation procedure [49]. The top 100 consensus SNPs were evaluated using a single, fast-refitting reference learner, chosen by availability in a fixed priority order (L1-penalized logistic regression, kernel ridge regression, GBLUP, LightGBM, XGBoost, CatBoost, random forest); a neural model was never used for this step. As an independent confirmatory analysis on the full sample, the reference learner was refitted on all animals — with fold-safe preprocessing and the hyperparameters already selected during nested CV held fixed — and its native importance recorded as the observed statistic. The serostatus labels were then permuted 10,000 times; for each permutation the reference learner was refitted with fixed hyperparameters and its native importance recomputed. Empirical p-values were obtained as $(b + 1) / (m + 1)$, where $b$ is the number of permutations whose importance was greater than or equal to the observed value and $m$ is the number of valid permutations, so that p-values remain strictly positive [50]. These p-values were converted to q-values to control the false-discovery rate [51, 52], and an inflation factor $\lambda$ was computed from the permutation p-values as a genomic-control diagnostic [53]. A SNP was flagged as significant at $q < 0.05$. SNPs outside the tested top 100 were not assigned a significance value and are reported by consensus rank only.

# 3. Results

## 3.1 Study sample and genotyping overview

The analysis dataset comprised 442 goats whose Border disease virus (BDV) serostatus was determined serologically by ELISA. The study cohort and its serological characterization are described in a companion manuscript (Yaman et al., in preparation). For the present analysis, animals were classified into two groups on the basis of serostatus: 117 seropositive animals (cases; 26.5%) and 325 seronegative animals (controls; 73.5%), corresponding to a case-to-control ratio of approximately 1:2.8. After quality control, genotypes were available for 44,375 autosomal SNPs with no missing calls and no monomorphic markers, and a single covariate (breeds) was retained for adjustment. The composition of the analysis sample is summarized in Table 1. The pronounced class imbalance was addressed throughout by imbalance-aware training and is reflected in the choice of threshold-independent performance metrics reported below.

**Table 1.** Composition of the analysis sample.

| Characteristic | Value |
|---|---|
| Total animals, n | 442 |
| Seropositive (cases), n (%) | 117 (26.5) |
| Seronegative (controls), n (%) | 325 (73.5) |
| Case-to-control ratio | 1 : 2.8 |
| Autosomal SNPs after QC, n | 44,375 |
| Covariates, n | 1 |
| Missing genotype rate | 0.00% |
| Monomorphic SNPs after QC, n | 0 |

## 3.2 Cross-validated discrimination of the models

All eight base learners and the Super Learner discriminated seropositive from seronegative animals above chance under the repeated nested cross-validation, and their out-of-fold (OOF) discrimination was modest and, notably, similar across learners: OOF AUC spanned a narrow band from 0.817 to 0.854 (a spread of only 0.037; Table 2, Figure 1). The genomic mixed model (GBLUP) attained the highest AUC (0.854), marginally ahead of the Super Learner (0.853), kernel ridge regression (0.849) and the residual multilayer perceptron (ResMLP, 0.846); the L1-penalized logistic regression was lowest (0.817). Because the AUC differences among the leading learners are well within the sampling uncertainty expected at this sample size, no single learner can be regarded as clearly superior on ranking discrimination alone.

The relative ordering depended on the metric considered (Figure 1). Summarized across the full metric panel by a direction-oriented composite z-score (Figure 3), ResMLP ranked highest (composite z = 1.11), reflecting the best Matthews correlation coefficient (MCC = 0.524), Brier score (0.137), Youden's J (0.578) and expected calibration error (ECE = 0.044) among all models, whereas GBLUP and the Super Learner were next (composite z = 0.64 and 0.62). Because the composite is standardized within the compared model set, it summarizes relative standing across metrics rather than an absolute quality; ResMLP's leading composite reflects its consistently favourable profile (best MCC, Brier, Youden's J and ECE) rather than superiority on any single headline metric such as AUC. Average precision (AUPRC), which is informative under the 26.5% case prevalence, ranged from 0.537 to 0.656 — in every case above the prevalence baseline of 0.265 — and was highest for kernel ridge regression (0.656). Sensitivity and specificity showed a clear trade-off across learners: random forest and kernel ridge regression reached the highest sensitivities (0.932 and 0.915) at the cost of lower specificity (0.582 and 0.634), while GBLUP and ResMLP were more balanced (sensitivity 0.769 and 0.812; specificity 0.772 and 0.766).

Relative to GBLUP as a common reference (Figure 2), the remaining learners showed small, predominantly unfavourable metric-wise differences: all learners were at or slightly below GBLUP on AUC (within ≈0.04), and most were below it on AUPRC (with kernel ridge regression the exception, +0.027), while only ResMLP and the Super Learner modestly exceeded GBLUP on MCC (+0.029 and +0.010) and Youden's J (+0.037 and +0.021). Taken together, these results indicate a genuine but modest genomic signal for BDV serostatus that is captured comparably by learners spanning penalized-linear, kernel, tree-ensemble, genomic-BLUP and neural families; no learner dominated across the metric panel, and the choice between them reflects a sensitivity–specificity and calibration trade-off rather than a decisive difference in discrimination. Consistent with the modest and imbalanced signal, all subsequent interpretation is framed in terms of ranking discrimination rather than calibrated individual-level classification.

**Table 2.** Cross-validated (out-of-fold) performance of the base learners and the Super Learner, ordered by AUC. Composite z is a direction-oriented z-score averaged across the metric panel (higher is better).

| Model | composite_z | AUC | AUPRC | MCC | Brier | F1 | Sens | Spec | Youden_J | ECE |
|---|---|---|---|---|---|---|---|---|---|---|
| ResMLP | 1.115 | 0.846 | 0.601 | 0.524 | 0.137 | 0.660 | 0.812 | 0.766 | 0.578 | 0.044 |
| GBLUP | 0.641 | 0.854 | 0.629 | 0.495 | 0.139 | 0.641 | 0.769 | 0.772 | 0.542 | 0.076 |
| SuperLearner | 0.623 | 0.853 | 0.608 | 0.505 | 0.142 | 0.646 | 0.821 | 0.742 | 0.562 | 0.090 |
| KernelRidge | 0.208 | 0.849 | 0.656 | 0.484 | 0.151 | 0.624 | 0.915 | 0.634 | 0.548 | 0.104 |
| LightGBM | -0.052 | 0.820 | 0.537 | 0.486 | 0.150 | 0.631 | 0.863 | 0.686 | 0.549 | 0.057 |
| CatBoost | -0.415 | 0.826 | 0.553 | 0.474 | 0.152 | 0.622 | 0.872 | 0.665 | 0.536 | 0.102 |
| RandomForest | -0.525 | 0.838 | 0.576 | 0.455 | 0.151 | 0.602 | 0.932 | 0.582 | 0.513 | 0.099 |
| XGBoost | -0.676 | 0.831 | 0.555 | 0.438 | 0.148 | 0.601 | 0.838 | 0.658 | 0.496 | 0.086 |
| LogisticRegression_L1 | -0.919 | 0.817 | 0.560 | 0.431 | 0.154 | 0.599 | 0.726 | 0.748 | 0.474 | 0.063 |

*Composite z is a post-hoc multi-metric summary: for each metric, model values were oriented so that higher is better (the error metrics Brier and ECE were sign-inverted), standardized to a z-score across the nine compared models (zero mean and unit population standard deviation within each metric), and averaged with equal weight across the nine metrics (AUC, AUPRC, MCC, Brier, F1, sensitivity, specificity, Youden's J, ECE). It is a relative ranking within the compared model set and carries no absolute scale.*

## 3.3 Overfitting and probability calibration

Every learner exhibited a substantial gap between training and test performance. Training AUC was consistently near ceiling — for example, GBLUP training AUC ranged from 0.996 to 1.000 and kernel ridge regression from roughly 0.97 to 0.99 across folds — whereas the corresponding test AUC was moderate (approximately 0.72–0.89), yielding mean train–test AUC gaps of about 0.12–0.18 (largest for LightGBM, $\Delta \approx 0.18$). These gaps indicate that, even with the in-fold-selected feature set, the fitted models substantially memorized the training data.

Probability calibration was correspondingly imperfect. The mean expected calibration error (ECE) — which summarizes the average magnitude of miscalibration across probability bins [54] — was low to moderate across all learners (0.044–0.104; Table 2), indicating that the average bin-wise discrepancy between predicted and observed frequencies was modest. Taken together, the near-ceiling training performance, the consistent train–test gaps, and the imperfect calibration indicate that the models discriminate moderately but do not produce probabilities that are well calibrated for individual risk estimation. Accordingly, all downstream interpretation relies on rank-based discrimination (AUC and related rank metrics) rather than on the predicted probabilities themselves.

## 3.4 Ensemble integration via the Super Learner

The Super Learner combined the eight base learners' out-of-fold predictions into a stacked predictor with a selected L2 penalty ($\lambda = 0.01$). The stacked predictor achieved an out-of-fold AUC of 0.853, statistically indistinguishable from the best single base learner (GBLUP, OOF AUC 0.854); at this sample size the 0.001 difference lies well within sampling variation, so stacking did not improve discrimination over the strongest individual learner. Importantly, the combiner did not degenerate onto a single learner: the learned weights were distributed rather than concentrated (Figure 3; Herfindahl–Hirschman index = 0.133; normalized entropy = 0.986), indicating no weight collapse. The same picture held across the multi-metric composite (Figure 4): the Super Learner ranked third (composite z = 0.62), just below GBLUP (0.64) and well behind ResMLP (1.11), and it did not rank first on any headline metric. This is the expected behaviour of stacking when the base learners are strongly correlated and one learner is already dominant [55]. The performance-based weights used in the consensus scoring (Section 2.7) therefore reflect a genuine multi-learner combination rather than the importance ranking of one model. In summary, the Super Learner is retained for its methodological role — providing collapse-checked, performance-based reliability weights for the cross-model consensus — while its stacked predictive performance was equivalent to, and did not exceed, that of the best individual learner.

3.5 Cross-model consensus prioritization of SNPs

Of the 44,375 SNPs in the panel, 34,110 (76.9%) received a cross-model consensus score, the remaining SNPs were never selected within any fold and are therefore not represented in the consensus ranking (Section 2.6). Among scored SNPs, the distribution of consensus scores was strongly right-skewed: the great majority clustered at low values, consistent with markers that were selected inconsistently or scored low on the native, permutation, stability and rank views, while a comparatively small number of candidates separated clearly from the bulk of the panel. The highest-ranked SNP, snp56437 (chromosome 11), reached a consensus score of 0.778, followed by snp13272 (chromosome 22) at 0.682; both examples illustrate the upper end of a ranking in which most of the genome-wide panel scored well below this range.

Highly-ranked candidates were not confined to a single chromosome or genomic region, consistent with a polygenic rather than a single-locus pattern of association. Each SNP's consensus score was accompanied by an uncertainty estimate (the standard deviation of the normalized importance across seed × fold repeats; e.g., 0.155 for the top-ranked SNP), which is reported alongside the ranking but, as described in Section 2.6, does not enter the consensus score itself. The 100 highest-ranked SNPs by consensus score were carried forward for formal significance testing, described next.

3.6 Permutation-based significance of prioritized SNPs

Applying the label-permutation test to the 100 highest-ranked consensus SNPs (Section 2.8; single reference learner, 10,000 permutations) identified 39 SNPs whose cross-model consensus importance exceeded the permutation null at a false-discovery rate of $q < 0.05$ (Table 3). These significant SNPs were distributed across 20 autosomes, with no single chromosome contributing more than four, a pattern consistent with a broadly distributed rather than a single-locus signal. Their consensus scores ranged from 0.29 to 0.78 and their q-values were uniformly small ($q \leq 0.005$). The strongest signal, snp56437 on chromosome 11 (consensus score 0.778, $q = 0.001$), was located within DAB2IP; the next most highly ranked significant SNPs lay on chromosomes 22, 16 and 20 (consensus scores 0.63–0.68). With respect to gene proximity, 19 of the 39 SNPs fell within an annotated gene, 12 lay within approximately 23–185 kb of the nearest gene, and 8 had no annotated gene within 200 kb.

**Table 3. SNPs passing the permutation-based false-discovery-rate threshold ($q < 0.05$).** For each SNP, the table gives its chromosomal location, cross-model consensus score and associated uncertainty, permutation q-value, and the nearest annotated gene within 200 kb together with the SNP's position relative to that gene.

| SNP ID | Chr | Pos | Consensus Score | Uncertainty | q_value | Nearest gene within 200KB range | Location |
|---|---|---|---|---|---|---|---|
| snp29912-scaffold325-754182 | 1 | 20990231 | 0.43186 | 0.25122 | 0.001 | NRIP1 | 185 kb |
| snp48171-scaffold682-595023 | 1 | 72766906 | 0.43947 | 0.20906 | 0.003 | ATP13A3 | 23 kb |
| snp37704-scaffold464-2337223 | 2 | 34816242 | 0.59697 | 0.21789 | 0.001 | IKZF2 | in the gene |
| 3_7699659_AF-PAKI | 3 | 21 | 0.37235 | 0.22989 | 0.002 | ENSCHIG00000010191 | in the gene |
| snp17461-scaffold1806-250263 | 3 | 1999895 | 0.29657 | 0.1181 | 0.004 | na | na |
| snp26001-scaffold2691-207930 | 5 | 30658025 | 0.37381 | 0.2398 | 0.002 | SPMIP11 | in the gene |
| snp40714-scaffold521-537618 | 5 | 28389180 | 0.39704 | 0.24069 | 0.002 | LETMD1 | 23 kb |
| snp48904-scaffold698-1426669 | 5 | 110207656 | 0.40509 | 0.2329 | 0.002 | GRAP2 | in the gene |
| snp18678-scaffold1887-353922 | 6 | 15814449 | 0.29579 | 0.24204 | 0.003 | LRIT3 | 41 kb |
| snp23892-scaffold2408-266677 | 6 | 85328015 | 0.41454 | 0.26319 | 0.002 | ENSCHIG0000009526 | in the gene |
| snp4748-scaffold1158-260266 | 7 | 26701629 | 0.33085 | 0.17273 | 0.002 | EDIL3 | in the gene |
| snp29392-scaffold318-731121 | 9 | 53517297 | 0.38508 | 0.22 | 0.003 | PTPRK | in the gene |
| snp2637-scaffold1076-2966283 | 10 | 11322885 | 0.3027 | 0.2136 | 0.003 | CEP128 | in the gene |
| snp31048-scaffold343-1829065 | 10 | 73040878 | 0.51413 | 0.25594 | 0.002 | SCG5 | in the gene |
| snp38023-scaffold473-309097 | 10 | 100043923 | 0.32001 | 0.2285 | 0.002 | APC | in the gene |
| snp15198-scaffold1621-656418 | 11 | 60405792 | 0.41483 | 0.26085 | 0.002 | B3GNT2; COMMD1 | 37 kb; 53 kb |
| snp3235-scaffold1098-108633 | 11 | 4005472 | 0.29326 | 0.13862 | 0.002 | CRACDL | in the gene |
| snp56437-scaffold887-473135 | 11 | 92190859 | 0.77756 | 0.15478 | 0.001 | DAB2IP | in the gene |
| snp42739-scaffold568-2285187 | 12 | 19749203 | 0.29572 | 0.198 | 0.002 | na | na |
| snp50196-scaffold717-5358187 | 12 | 59798648 | 0.39839 | 0.25204 | 0.001 | na | na |
| snp32522-scaffold371-1217074 | 13 | 40301728 | 0.31119 | 0.20477 | 0.003 | PAX1 | 27 kb |
| snp48684-scaffold691-2078262 | 13 | 59821510 | 0.3111 | 0.21982 | 0.002 | TCF15 | 44 kb |
| snp48694-scaffold691-2541731 | 13 | 59347221 | 0.4027 | 0.23521 | 0.001 | PSMF1 | 35 kb |
| snp56113-scaffold875-2628829 | 13 | 13266240 | 0.45864 | 0.22506 | 0.002 | na | na |
| snp43138-scaffold574-63323 | 14 | 68461471 | 0.36814 | 0.24635 | 0.003 | na | na |
| GoatD01.032764 | 16 | 54 | 0.66946 | 0.21883 | 0.002 | ENSCHIG00000014347 | 29 kb |
| snp17321-scaffold18-4191820 | 16 | 29024322 | 0.67818 | 0.15457 | 0.002 | AHCTF1 | in the gene |
| snp27892-scaffold299-3503242 | 16 | 76906018 | 0.32122 | 0.19341 | 0.004 | na | na |
| snp8977-scaffold1327-155697 | 16 | 22609933 | 0.34812 | 0.19948 | 0.005 | MARK1 | in the gene |
| snp47246-scaffold663-905216 | 17 | 66576050 | 0.42466 | 0.24793 | 0.004 | MND1 | in the gene |
| snp14620-scaffold1589-2865953 | 18 | 28950082 | 0.36036 | 0.16023 | 0.004 | na | na |
| snp7176-scaffold1266-811094 | 18 | 38376664 | 0.34682 | 0.17972 | 0.005 | PSMD7 | 116 kb |
| 19_8765032_AF-PAKI | 19 | 27 | 0.31262 | 0.23372 | 0.002 | na | na |
| snp10605-scaffold1377-108280 | 19 | 26115396 | 0.46852 | 0.23212 | 0.002 | ZNF232 | 23 kb |
| snp14659-scaffold1591-277491 | 20 | 10572806 | 0.62565 | 0.17192 | 0.002 | SLC30A5 | in the gene |
| snp13272-scaffold1509-2105048 | 22 | 48179057 | 0.68195 | 0.23099 | 0.001 | PBRM1 | in the gene |
| snp38338-scaffold485-648989 | 25 | 16611582 | 0.4034 | 0.25195 | 0.003 | SYT17 | in the gene |
| snp43018-scaffold570-937873 | 25 | 6531622 | 0.48003 | 0.17422 | 0.001 | RBFOX1 | in the gene |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| snp37867-scaffold466-3649788 | 27 | 24711317 | 0.38297 | 0.2417 | 0.003 | MSR1 | 51 kb |

*Significance was assessed by permuting the phenotype label 10,000 times and refitting a single reference learner (L1-penalized logistic regression) on the full sample, applied to the 100 highest-ranked consensus SNPs; q-values were obtained by the Storey procedure (Section 2.8). Consensus score and uncertainty derive from the cross-model consensus (Section 2.7); uncertainty is the standard deviation of the normalized importance across seed × fold repeats. Chr, chromosome; Pos, position (genome assembly [ARS1/CHIR — specify]). "In the gene" indicates that the SNP lies within the gene body; otherwise the distance to the nearest gene boundary is given. "na" indicates no annotated gene within 200 kb. SNP identifiers follow the genotyping array's naming.*

### 3.7 Genomic-control diagnostic

The genomic-control inflation factor, computed from the permutation *p*-values of the 100 consensus candidates (Section 2.8), was λ = 0.000. This near-zero value does not indicate deflation of a genome-wide test; rather, it is an artifact of computing the median-based statistic on an enriched, pre-selected panel characterized by a strongly bimodal *p*-value distribution. Accordingly, λ is reported for completeness but is not interpreted as a measure of test calibration, nor should it be compared to genomic-control values from unselected genome-wide association studies.

## 4. Discussion

# 4.1 Principal findings

Across a heterogeneous panel of eight base learners and a stacked ensemble, out-of-fold discrimination of BDV serostatus from genome-wide SNP genotypes was modest but consistent, spanning a narrow AUC range of 0.817–0.854. No single learner or model family — penalized-linear, kernel, tree-ensemble, genomic-BLUP or neural — clearly dominated the others, and the relative ordering of learners depended on which metric was examined. This pattern indicates a genuine, broadly distributed genomic signal for host response to Border disease virus, of a magnitude consistent with a polygenic trait rather than a small number of large-effect loci. The cross-model consensus framework developed here is best understood as a prioritization tool for generating a ranked list of candidate SNPs from this signal, not as a mechanism for establishing causal variants; the scope and limits of that framework are discussed below.

# 4.2 Predictive performance in context

The stacked Super Learner predictor achieved essentially the same out-of-fold discrimination as the best individual base learner (GBLUP; AUC 0.853 vs 0.854), and did not lead on any headline metric of the composite panel. Given the sample size, a difference of this magnitude is well within sampling variation and should not be interpreted as evidence that stacking failed; rather, it is the expected outcome when base learners are strongly correlated — as heterogeneous learners trained on the same genotype matrix and the same modest signal necessarily are — and one learner is already close to the ceiling attainable from that signal. This observation is informative in its own right: it suggests that, for this trait and sample size, most of the exploitable genomic information was already captured by the simplest well-specified model (GBLUP), with the more flexible learners neither adding nor losing appreciable discrimination.

# 4.3 Consensus scoring as a cross-model prioritization framework

The consensus score combines each learner's importance views using weights derived from a Super Learner fitted to out-of-fold predictions, so that learners predicting serostatus more accurately contribute more to the aggregated ranking. Two properties of this construction bound what can be claimed. First, because the weighting is performance-based rather than an independent line of evidence, a SNP that is influential in the better-predicting learners is favoured by design; a high consensus score therefore reflects convergence of importance across models that share the same training data, not agreement among statistically independent analyses. Second, because all learners were trained on the same genotype matrix and the same modest, imbalanced phenotype, shared structure in the data — including linkage disequilibrium and, in principle, unmodelled confounding — can in part underlie apparent cross-model agreement. The Super Learner weights did not collapse onto a single learner (Herfindahl–Hirschman index = 0.133; normalized entropy = 0.986), so the consensus reported here reflects a genuine multi-learner combination rather than the importance ranking of one dominant model; had collapse occurred, the consensus would have reduced to that single learner's ranking, and we would have reported it as such. Within these bounds, the consensus score is appropriately used to prioritize candidates for downstream confirmation and functional follow-up, and it is this prioritization role — not a claim of causal or independently replicated association — that we intend throughout.

# 4.4 Methodological strengths

Several design choices reduce the risk that the reported findings are artifacts of the analysis pipeline rather than of the data. All hyperparameter tuning and feature filtering were performed strictly within each training fold, avoiding the well-documented optimistic bias

that arises when any data-driven step precedes cross-validated evaluation [35, 36]; this is a stricter standard than some comparable genomic-prediction pipelines apply in practice — for example, a recent mutual-information-based prediction pipeline reported results computed on the full dataset without an explicit in-fold repetition of feature selection — and it should reduce, though it does not eliminate, upward bias in both the predictive metrics and the SNP importance scores reported here. The deliberate use of a heterogeneous panel of learners means that the reported cross-model agreement, where it exists, cannot be attributed to a shared inductive bias of a single algorithm family. Finally, full data fingerprinting, seed control, and per-fold checkpointing were used throughout, so that all reported results can be traced to a specific, reproducible model and data state.

## 4.5 Biological plausibility of prioritized candidate genes

The candidate genes discussed below were assigned to prioritized SNPs by physical proximity (nearest annotated gene within 200 kb) rather than by direct functional evidence; a substantial fraction of the significant SNPs are intergenic or only proximity-linked (Table 3). These gene–SNP links are therefore hypothetical and require functional validation (e.g., eQTL mapping, chromatin-conformation capture, or targeted assays); none of the associations below has been experimentally tested in a pestivirus context.

Innate pattern recognition and inflammatory regulation

MSR1 (macrophage scavenger receptor 1) has the most direct, internally consistent support among the genes reviewed. In a murine quantitative-trait-locus study, Msr1 was identified as the candidate gene underlying strain differences in lipopolysaccharide (LPS)-induced interleukin-10 (IL-10) production; Msr1-knockout mice showed a selective reduction in LPS-induced IL-10, with no corresponding change in TNF-α or IL-6, indicating a role specific to the anti-inflammatory arm of the innate response [56]. Independently, targeted disruption of the same receptor increased susceptibility to Listeria monocytogenes and herpes simplex virus type-1 infection in mice, indicating a broader contribution to antimicrobial host defense [57]. In an unrelated viral system, MSR1 acted as a cofactor for vesicular stomatitis virus (VSV) cellular entry, with Msr1-deficient mice showing reduced viral load and mortality after central-nervous-system challenge [58]. Together, these three independent, gene-specific studies support a dual, context-dependent role: protective in endotoxemia-type innate signaling and in bacterial/herpesvirus defense, proviral in at least one neurotropic infection.

COMMD1 is a validated negative regulator of NF-κB signaling. A first-in-human phase I trial of CIGB-552, a synthetic peptide that binds COMMD1 and raises its protein levels, produced measurable immunological effects (shifts in circulating CD4+/CD8 + ratios) in patients with advanced solid tumors, demonstrating that the COMMD1–NF-κB axis is pharmacologically tractable in vivo [59]. This supports COMMD1's plausibility as a regulator of inflammatory tone relevant to infection outcomes; the cited trial was conducted in an oncology rather than infectious-disease context, and that distinction should be retained.

Antigen processing and T-cell receptor signaling

*PSMD7* encodes a non-ATPase subunit of the 19S proteasome regulatory particle. The proteasome, including its 19S regulatory components, generates the MHC class I-restricted peptides required for CD8 + T-cell recognition of infected cells, and immunoproteasome-dependent processing specifically has been tied to the early phase of antiviral immune responses [60]. This is pathway-level, not gene-specific, support: no reviewed source studies PSMD7 individually in an infection context, so the claim should stay at the level of "PSMD7 is a structural component of a pathway with documented antiviral relevance," not be presented as a PSMD7-specific finding.

*GRAP2* is a hematopoietic Grb2-family adaptor that bridges the LAT and SLP-76 adaptors downstream of T-cell receptor (TCR) engagement. Genetic ablation studies show it is not strictly required for proximal TCR signaling but acts as a dose-dependent amplifier of TCR responsiveness — particularly to weak stimuli — and is required for FcεRI-mediated mast-cell activation [61].

Monogenic immune-disease classification: *IKZF2* and *PAX1*

Both genes appear in the 2022 IUIS Expert Committee classification of novel monogenic inborn errors of immunity: *IKZF2* is listed with both autosomal-recessive and autosomal-dominant disease-causing variants, and *PAX1* is separately listed among the newly recognized genes in the same update [62]. This is genuine evidence that variation at both loci associates with clinically significant immune dysfunction, though it is classification-level rather than mechanistic evidence. *IKZF2* has an additional, gene-specific mechanistic link: single-cell transcriptomic profiling of stem-like CD8 + T cells during chronic viral infection identified Ikzf2 directly among the genes upregulated by BACH2, the transcriptional repressor that enforces the stem-like, self-renewing fate protecting these cells from terminal exhaustion *[63]*. This gives *IKZF2* the strongest, most directly verified link to chronic-infection CD8 + T-cell biology among the genes reviewed here after *MSR1* and C*OMMD1*; *PAX1* currently has classification-level support only.

Wnt-pathway immune regulation: *APC*

*APC* (adenomatous polyposis coli), best known as the tumor-suppressor component of the Wnt/β-catenin destruction complex (APC–Axin–CK1α–GSK3β), also regulates β-catenin turnover in dendritic cells (DCs). Canonical Wnt/β-catenin signaling in DCs programs a tolerogenic, immunosuppressive DC state that shapes downstream T-cell responses, documented primarily in tumor and autoimmune contexts [64]. This is genuine evidence for a Wnt–immune connection operating through the same destruction complex *APC* participates in; because the cited work addresses antitumor and autoimmune settings rather than infection, any extension to infectious disease should be framed as inference, not established finding.

Viral-genome interaction with host chromatin regulators: *PBRM1*

HPV-encoded microRNAs have predicted cellular targets that include *PBRM1*, a SWI/SNF chromatin-remodeling subunit — categorized among predicted cell-cycle/M-phase targets in the study that identified these viral microRNAs. This supports the general principle that viral genomes interface with host chromatin-regulatory networks. Because PBRM1 is classified as a cell-cycle target rather than part of a distinct immune-regulation gene set, any immune-evasion role should be framed as an inference related to chromatin regulation rather than a direct immune-function discovery [65].

Remaining candidate genes

For other genes, no source identified in this review provides direct, gene-specific evidence of immune or antiviral function. This is stated as a limitation rather than implied through citation: absence of prior literature is itself informative for genes emerging from a statistical discovery pipeline and does not undermine their candidacy.

## 4.6 Limitations

Several limitations qualify the interpretation of these results and should guide how the findings are used.

*Scope of the significance tests.* The single-model permutation test evaluates only one fast-refitting reference learner, applied to the full sample rather than held-out data, and only among the 100 highest-ranked consensus candidates; SNPs outside this set were not tested and carry no significance value. The reference learner is chosen from a fixed priority list of fast-refitting models, excluding the neural model from consideration for this step.

*Sample size and class imbalance.* With 117 seropositive animals among 442 sampled, statistical power — for both discrimination and per-SNP significance testing — is constrained, and the resulting AUC and importance estimates carry non-trivial sampling uncertainty even where point estimates appear stable across folds and seeds.

*Single cohort and study design.* All analyses draw on one cohort of goats characterized by a single serological assay; generalization to other breeds, geographic regions, or diagnostic criteria has not been assessed and should not be assumed.

*Functional literature coverage.* For a subset of the top consensus candidate genes, no prior literature provides direct, gene-specific evidence regarding their immune or antiviral functions. While this absence of empirical data is expected for a statistical discovery pipeline uncovering novel candidates, these targets currently lack functional validation.

## 4.7 Future directions

The candidate SNPs prioritized here motivate several concrete next steps rather than standing as a terminal result. Functional follow-up—including expression quantitative trait locus (eQTL) analysis, regulatory-element mapping, and targeted functional assays—is required to move from statistical association to mechanistic evidence, especially given that most prioritized SNPs reside in non-coding regions or are merely proximity-linked to annotated genes. Replication in independent, ideally multi-breed and larger cohorts would address both the sample-size limitation and the single-cohort generalizability concern raised above.

## 5. Conclusion

This study provides evidence of a modest, broadly distributed host genomic signal associated with Border disease virus serostatus in goats — captured consistently across machine-learning models spanning several methodological families, rather than by any single algorithm. Combining leakage-safe importance estimates with performance-based model weighting, the cross-model consensus framework developed here prioritizes a ranked set of candidate SNPs, of which 39 withstand single-model permutation testing at a controlled false-discovery rate. These candidates are best treated as a starting point rather than an endpoint: a hypothesis-generating shortlist for functional validation and independent replication, not a validated set of causal resistance loci.

## Abbreviations

### Methodology & Machine Learning

- **AUC:** Area Under the Receiver Operating Characteristic Curve
- **AUPRC:** Area Under the Precision-Recall Curve
- **BLUP:** Best Linear Unbiased Prediction
- **CatBoost:** Categorical Boosting
- **CV:** Cross-Validation
- **ECE:** Expected Calibration Error
- **GBLUP:** Genomic Best Linear Unbiased Prediction
- **LightGBM:** Light Gradient Boosting Machine
- **MCC:** Matthews Correlation Coefficient
- **MLP:** Multilayer Perceptron
- **OOF:** Out-Of-Fold
- **RF:** Random Forest
- **ResMLP:** Residual Multilayer Perceptron
- **ROC:** Receiver Operating Characteristic
- **TPE:** Tree-structured Parzen Estimator
- **XGBoost:** Extreme Gradient Boosting

### Genomics & Molecular Biology

- **APC:** Adenomatous Polyposis Coli
- **eQTL:** Expression Quantitative Trait Locus
- **FDR:** False Discovery Rate
- **MHC:** Major Histocompatibility Complex
- **QC:** Quality Control
- **RT-PCR:** Reverse Transcription Polymerase Chain Reaction
- **SNP:** Single Nucleotide Polymorphism

### Immunology & Virology

- **BD / BDV:** Border Disease / Border Disease Virus
- **BVDV:** Bovine Viral Diarrhea Virus
- **CSFV:** Classical Swine Fever Virus
- **DC:** Dendritic Cell
- **ELISA:** Enzyme-Linked Immunosorbent Assay
- **HPV:** Human Papillomavirus
- **IL:** Interleukin
- **LPS:** Lipopolysaccharide
- **NF-κB:** Nuclear Factor Kappa B
- **OD:** Optical Density
- **PI:** Persistently Infected / Persistent Infection
- **S/P%:** Sample-to-Positive Percentage
- **TCR:** T-cell Receptor
- **TNF-α:** Tumor Necrosis Factor Alpha
- **VSV:** Vesicular Stomatitis Virus

## Declarations

**Ethics approval and consent to participate**

This study was approved by the Ethics Committee for the Use of Animals in Research and Experimentation of the Sheep Breeding and Research Institute, Ministry of Agriculture and Forestry, Türkiye (Approval No: 066/14.02.2023) and was conducted in accordance with the national guidelines for animal experimentation. Prior to the study, official permission and informed consent were obtained from the Siirt Provincial Directorate of Agriculture, Ministry of Agriculture and Forestry, Türkiye (Approval No: E-64380313-325.01-8848896/08.02.2023). The authors also complied with the ARRIVE guidelines.

**Consent for publication**

Not applicable.

**Availability of data and material**

The data used in this study have been deposited in the figshare.com database under https://doi.org/10.6084/m9.figshare.30845045

**Competing interests**

The authors declare no competing interests

**Funding**

None

**Authors' contributions**

Y.Y. conceptualized and designed the study, performed the machine learning analyses, interpreted the results, and wrote the manuscript. B.T. and S.Y. curated the data and performed the preliminary statistical analyses.

**Acknowledgements**

The genotype data utilized in this study were obtained from a previously conducted and successfully completed TÜBİTAK project (TOVAG-222O107), and we extend our sincere gratitude to TÜBİTAK for their support. The corresponding author also expresses his most profound gratitude to Gazi Mustafa Kemal Atatürk, the founder of the modern Republic of Turkey and our supreme leader, who made science his guide.

**Corresponding author**

yalcinyaman@gmail.com

## References

1. Nettleton PF, Gilray JA, Russo P, Dlissi E. Border disease of sheep and goats. Vet Res. 1998;29:327–40.
2. Cantile C, Youssef S. Chapter 4 - Nervous System. In: Maxie MG, editor. Jubb, Kennedy & Palmer's Pathology of Domestic Animals: Volume 1 (Sixth Edition). W.B. Saunders; 2016. p. 250–406. https://doi.org/10.1016/B978-0-7020-5317-7.00004-7.
3. More S, Bøtner A, Butterworth A, Calistri P, Depner K, Edwards S, et al. Assessment of listing and categorisation of animal diseases within the framework of the Animal Health Law (Regulation (EU) No 2016/429): Border disease. EFSA Journal. 2017;15.
4. Hassan SD. Prevalence of border disease virus in sheep and goats in Mosul, Iraq. Iraqi Journal of Veterinary Sciences. 2021;35:257–62. https://doi.org/10.33899/ijvs.2020.126758.1372.
5. Khordadmehr M, Sadri H, Shirazi J, Babazadeh S, Akbari H, Jigari-Asl F, et al. Investigating Viral Causes of Abortions: A Comprehensive Molecular and Pathological Study on Small Ruminants in East Azerbaijan Province, Iran. Vet Med Sci. 2025;11:e70308. https://doi.org/10.1002/vms3.70308.
6. Cecco BS de, Santos IR dos, Molossi FA, Canal CW, Barros CSL de, Driemeier D, et al. Viral diseases of sheep in Brazil: a review and current status. Cienc Rural. 2023;53:e20220218. https://doi.org/https://doi.org/10.1590/0103-8478cr20220218.
7. Mao L, Li W, Hao F, Yang L, Li J, Sun M, et al. Research Progress on Emerging Viral Pathogens of Small Ruminants in China during the Last Decade. Viruses. 2022;14:1288. https://doi.org/10.3390/v14061288.
8. Navarro-Lopez R, Perez-de La Rosa J, Villarreall-Silva M, Solis-Hernandez M, Rojas-Torres E, Lemus Y Sanchez J, et al. Identification of border disease virus in naturally infected pigs in Mexico. JSHAP. 2023;31:72–6. https://doi.org/10.54846/jshap/1296.
9. Righi C, Petrini S, Pierini I, Giammarioli M, De Mia GM. Global Distribution and Genetic Heterogeneity of Border Disease Virus. Viruses. 2021;13:950. https://doi.org/10.3390/v13060950.

10. Bouzalas IG, Gelasakis AI, Chassalevris T, Apostolidi ED, Pappas F, Ekateriniadou L, et al. Circulation of Pestiviruses in Small Ruminants from Greece: First Molecular Identification of Border Disease Virus. Vaccines (Basel). 2023;11:918. https://doi.org/10.3390/vaccines11050918.
11. Dastjerdi A, Strong R, La Rocca SA, Wessels M, Wessels J, Whitaker K, et al. Investigation into an outbreak of Border disease virus in pigs in England. Transbound Emerg Dis. 2022;69:1698–706. https://doi.org/10.1111/tbed.14539.
12. Borel N, Frey CF, Gottstein B, Hilbe M, Pospischil A, Franzoso FD, et al. Laboratory diagnosis of ruminant abortion in Europe. Vet J. 2014;200:218–29. https://doi.org/10.1016/j.tvjl.2014.03.015.
13. Hurtado A, Sanchez I, Bastida F, Minguijón E, Juste RA, García-Pérez AL. Detection and quantification of pestivirus in experimentally infected pregnant ewes and their progeny. Virol J. 2009;6:189. https://doi.org/10.1186/1743-422X-6-189.
14. Kalaiyarasu S, Rajukumar K, Mishra N, Sudhakar SB, Singh VP. Detection and Genetic Characterization of Border Disease Virus (BDV) Isolated from a Persistently Infected Sheep in a Migratory Flock from Rajasthan State, Northwestern India. https://www.mdpi.com/1999-4915/16/9/1390. Accessed 6 Aug 2026.
15. Ciulli S, Purpari G, Agnello S, Di Marco P, Di Bella S, Volpe E, et al. Evidence for Tunisian-Like Pestiviruses Presence in Small Ruminants in Italy Since 2007. Transbound Emerg Dis. 2017;64:1243–53. https://doi.org/10.1111/tbed.12498.
16. De Mia GM, Greiser-Wilke I, Feliziani F, Giammarioli M, De Giuseppe A. Genetic characterization of a caprine pestivirus as the first member of a putative novel pestivirus subgroup. J Vet Med B Infect Dis Vet Public Health. 2005;52:206–10. https://doi.org/10.1111/j.1439-0450.2005.00850.x.
17. Incili CA, Eroksuz Y, Timurkan MO, Karabulut B, Yerlikaya Z, Eroksuz H. Simultaneous Occurrence of Nasal Carcinoma Induced by Enzootic Nasal Tumour Virus-2 (ENTV-2) and Border Disease Virus in Three Goats. Vet Med Sci. 2025;11:e70325. https://doi.org/10.1002/vms3.70325.
18. Sharawi SSA, Yousef MR, Dokhan KZ, Al-Hofufy AN. VIROLOGIC, SEROLOGIC AND PATHOLOGIC STUDIES INTO BORDER DISEASE VIRUS (BDV) INFECTION IN A GOAT HERD IN SAUDIA ARABIA. Assiut Veterinary Medical Journal. 2010. https://doi.org/10.21608/avmj.2010.174238.
19. Tegtmeier C, Stryhn H, Uttenthal Å, Kjeldsen AM, Nielsen TK. Seroprevalence of Border Disease in Danish Sheep and Goat Herds. Acta Vet Scand. 2000;41:339–44. https://doi.org/10.1186/BF03549624.
20. Danuser R, Vogt H-R, Kaufmann T, Peterhans E, Zanoni R. Seroprevalence and characterization of pestivirus infections in small ruminants and new world camelids in Switzerland. Schweiz Arch Tierheilkd. 2009;151:109–17. https://doi.org/10.1024/0036-7281.151.3.109.
21. Krametter-Froetscher R, Loitsch A, Kohler H, Schleiner A, Schiefer P, Moestl K, et al. Prevalence of antibodies to pestiviruses in goats in Austria. J Vet Med B Infect Dis Vet Public Health. 2006;53:48–50. https://doi.org/10.1111/j.1439-0450.2006.00906.x.
22. Zhigailov AV, Perfilyeva YV, Malysheva AA, Cherusheva AS, Berdygulova ZA, Naizabayeva DA, et al. Survey and Associated Risk Factors for the Presence of Ruminant Pestiviruses in Domestic Ovine and Caprine Populations from Kazakhstan. Viruses. 2025;17:676. https://doi.org/10.3390/v17050676.
23. Dhahir HS, Talb A-OQ, Asim MH. Preliminary Study of Seroprevalence of Border Disease Virus (BDV) Among Sheep and Goats in Mosul City, Iraq. Adv Anim Vet Sci. 2019;7. https://doi.org/10.17582/journal.aavs/2019/7.7.566.569.
24. Braun U, Bachofen C, Büchi R, Hässig M, Peterhans E. Infection of cattle with Border disease virus by sheep on communal alpine pastures. Schweizer Archiv für Tierheilkunde. 2013;155:123–8. https://doi.org/10.1024/0036-7281/a000428.
25. Hidayat W, Wuryastuty H, Wasito R. Detection of Pestivirus in small ruminants in Central Java, Indonesia. Vet World. 2021;14:996–1001. https://doi.org/10.14202/vetworld.2021.996-1001.
26. Akkaya F, Evci FZ, Küpeli ZA, Kandemir Ç, Taşkın T, Koşum N. Border Disease In A Lamb: A Case Report. Journal of the Hellenic Veterinary Medical Society. 2022;73:5083–8. https://doi.org/10.12681/jhvms.28216.
27. Postel A, Schmeiser S, Oguzoglu TC, Indenbirken D, Alawi M, Fischer N, et al. Close Relationship of Ruminant Pestiviruses and Classical Swine Fever Virus. Emerg Infect Dis. 2015;21:668–72. https://doi.org/10.3201/eid2104.141441.
28. Sozzi E, Lavazza A, Gaffuri A, Bencetti FC, Prosperi A, Lelli D, et al. Isolation and Full-Length Sequence Analysis of a Pestivirus from Aborted Lamb Fetuses in Italy. Viruses. 2019;11:744. https://doi.org/10.3390/v11080744.
29. Mahmoud MA, Ghazy AA, Shaapan RM. Review of diagnostic procedures and control of some viral diseases causing abortion and infertility in small ruminants in Egypt. Iraqi Journal of Veterinary Sciences. 2021;35:513–21. https://doi.org/10.33899/ijvs.2020.127114.1461.
30. Meyer G, Combes M, Teillaud A, Pouget C, Bethune M-A, Cassard H. Vaccination of Sheep with Bovine Viral Diarrhea Vaccines Does Not Protect against Fetal Infection after Challenge of Pregnant Ewes with Border Disease Virus. Vaccines (Basel). 2021;9:805. https://doi.org/10.3390/vaccines9080805.

31. Schweizer M, Stalder H, Haslebacher A, Grisiger M, Schwermer H, Di Labio E. Eradication of Bovine Viral Diarrhoea (BVD) in Cattle in Switzerland: Lessons Taught by the Complex Biology of the Virus. Front Vet Sci. 2021;8:702730. https://doi.org/10.3389/fvets.2021.702730.
32. Getahun C, Dessie T, Tesema E, Alamirew E. Sero-prevalence of Border Disease Virus Antibodies in Recently Introduced Dorper Sheep Flocks at Debre Birhan Agricultural Research Center, Ethiopia. Journal of Applied Veterinary Sciences. 2022;0:0–0. https://doi.org/10.21608/javs.2022.155122.1171.
33. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7. https://doi.org/10.1186/s13742-015-0047-8.
34. Browning BL, Zhou Y, Browning SR. A One-Penny Imputed Genome from Next-Generation Reference Panels. Am J Hum Genet. 2018;103:338–48. https://doi.org/10.1016/j.ajhg.2018.07.015.
35. Ambroise C, McLachlan GJ. Selection bias in gene extraction on the basis of microarray gene-expression data. Proc Natl Acad Sci U S A. 2002;99:6562–6. https://doi.org/10.1073/pnas.102102699.
36. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics. 2006;7:91. https://doi.org/10.1186/1471-2105-7-91.
37. Lin T-Y, Goyal P, Girshick R, He K, Dollár P. Focal Loss for Dense Object Detection. In: 2017 IEEE International Conference on Computer Vision (ICCV). 2017. p. 2999–3007. https://doi.org/10.1109/ICCV.2017.324.
38. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework. 2019. https://doi.org/10.48550/arXiv.1907.10902.
39. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. p. 785–94. https://doi.org/10.1145/2939672.2939785.
40. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc.; 2017. p. 3149–57.
41. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. 2019. https://doi.org/10.48550/arXiv.1706.09516.
42. Breiman L. Random Forests. Machine Learning. 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.
43. VanRaden PM. Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science. 2008;91:4414–23. https://doi.org/10.3168/jds.2007-0980.
44. Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. Efficient control of population structure in model organism association mapping. Genetics. 2008;178:1709–23. https://doi.org/10.1534/genetics.107.080101.
45. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825–30.
46. Meinshausen N, Bühlmann P. Stability selection - Meinshausen - 2010 - Journal of the Royal Statistical Society: Series B (Statistical Methodology) - Wiley Online Library. 2010. https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1467-9868.2010.00740.x. Accessed 7 Aug 2026.
47. van der Laan MJ, Polley EC, Hubbard AE. Super learner. Stat Appl Genet Mol Biol. 2007;6:Article25. https://doi.org/10.2202/1544-6115.1309.
48. Pudjihartono N, Fadason T, Kempa-Liehr AW, O'Sullivan JM. A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. Front Bioinform. 2022;2:927312. https://doi.org/10.3389/fbinf.2022.927312.
49. Ojala M, Garriga GC. Permutation Tests for Studying Classifier Performance. In: 2009 Ninth IEEE International Conference on Data Mining. 2009. p. 908–13. https://doi.org/10.1109/ICDM.2009.108.
50. Davison AC, Hinkley DV. Bootstrap Methods and their Application. Cambridge: Cambridge University Press; 1997. https://doi.org/10.1017/CBO9780511802843.
51. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing | Journal of the Royal Statistical Society Series B: Statistical Methodology | Oxford Academic. 1995. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
52. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A. 2003;100:9440–5. https://doi.org/10.1073/pnas.1530509100.
53. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999;55:997–1004. https://doi.org/10.1111/j.0006-341x.1999.00997.x.

54. Fan KL, Pennello G, Liu Q, Petrick N, Samala RK, Samuelson FW, et al. Calzone: A Python package for measuring calibration of probabilistic models for classification. Journal of Open Source Software. 2025;10:8026. https://doi.org/10.21105/joss.08026.
55. Džeroski S, Ženko B. Is Combining Classifiers with Stacking Better than Selecting the Best One? Machine Learning. 2004;54:255–73. https://doi.org/10.1023/B:MACH.0000015881.36452.6e.
56. Fulton WB, Reeves RH, Takeya M, De Maio A. A quantitative trait loci analysis to map genes involved in lipopolysaccharide-induced inflammatory response: identification of macrophage scavenger receptor 1 as a candidate gene. J Immunol. 2006;176:3767–73. https://doi.org/10.4049/jimmunol.176.6.3767.
57. Suzuki H, Kurihara Y, Takeya M, Kamada N, Kataoka M, Jishage K, et al. A role for macrophage scavenger receptors in atherosclerosis and susceptibility to infection. Nature. 1997;386:292–6. https://doi.org/10.1038/386292a0.
58. Yang D, Lin T, Li C, Harrison AG, Geng T, Wang P. A critical role for MSR1 in vesicular stomatitis virus infection of the central nervous system. iScience. 2021;24:102678. https://doi.org/10.1016/j.isci.2021.102678.
59. Vallespi MG, Mestre B, Marrero MA, Uranga R, Rey D, Lugiollo M, et al. A first-in-class, first-in-human, phase I trial of CIGB-552, a synthetic peptide targeting COMMD1 to inhibit the oncogenic activity of NF-κB in patients with advanced solid tumors - Vallespi - 2021 - International Journal of Cancer - Wiley Online Library. 2021. https://doi.org/10.1002/ijc.33695Digital%20Object%20Identifier%20(DOI).
60. Strehl B, Textoris-Taube K, Jäkel S, Voigt A, Henklein P, Steinhoff U, et al. Antitopes define preferential proteasomal cleavage site usage. J Biol Chem. 2008;283:17891–7. https://doi.org/10.1074/jbc.M710042200.
61. Yablonski D. Bridging the Gap: Modulatory Roles of the Grb2-Family Adaptor, Gads, in Cellular and Allergic Immune Responses. Front Immunol. 2019;10:1704. https://doi.org/10.3389/fimmu.2019.01704.
62. Tangye SG, Al-Herz W, Bousfiha A, Cunningham-Rundles C, Franco JL, Holland SM, et al. Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. J Clin Immunol. 2022;42:1473–507. https://doi.org/10.1007/s10875-022-01289-3.
63. Yao C, Lou G, Sun H-W, Zhu Z, Sun Y, Chen Z, et al. BACH2 enforces the transcriptional and epigenetic programs of stem-like CD8+ T cells. Nat Immunol. 2021;22:370–80. https://doi.org/10.1038/s41590-021-00868-7.
64. Suryawanshi A, Hussein MS, Prasad PD, Manicassamy S. Wnt Signaling Cascade in Dendritic Cells and Regulation of Anti-tumor Immunity. Front Immunol. 2020;11:122. https://doi.org/10.3389/fimmu.2020.00122.
65. Qian K, Pietilä T, Rönty M, Michon F, Frilander MJ, Ritari J, et al. Identification and validation of human papillomavirus encoded microRNAs. PLoS One. 2013;8:e70202. https://doi.org/10.1371/journal.pone.0070202.

## Figures

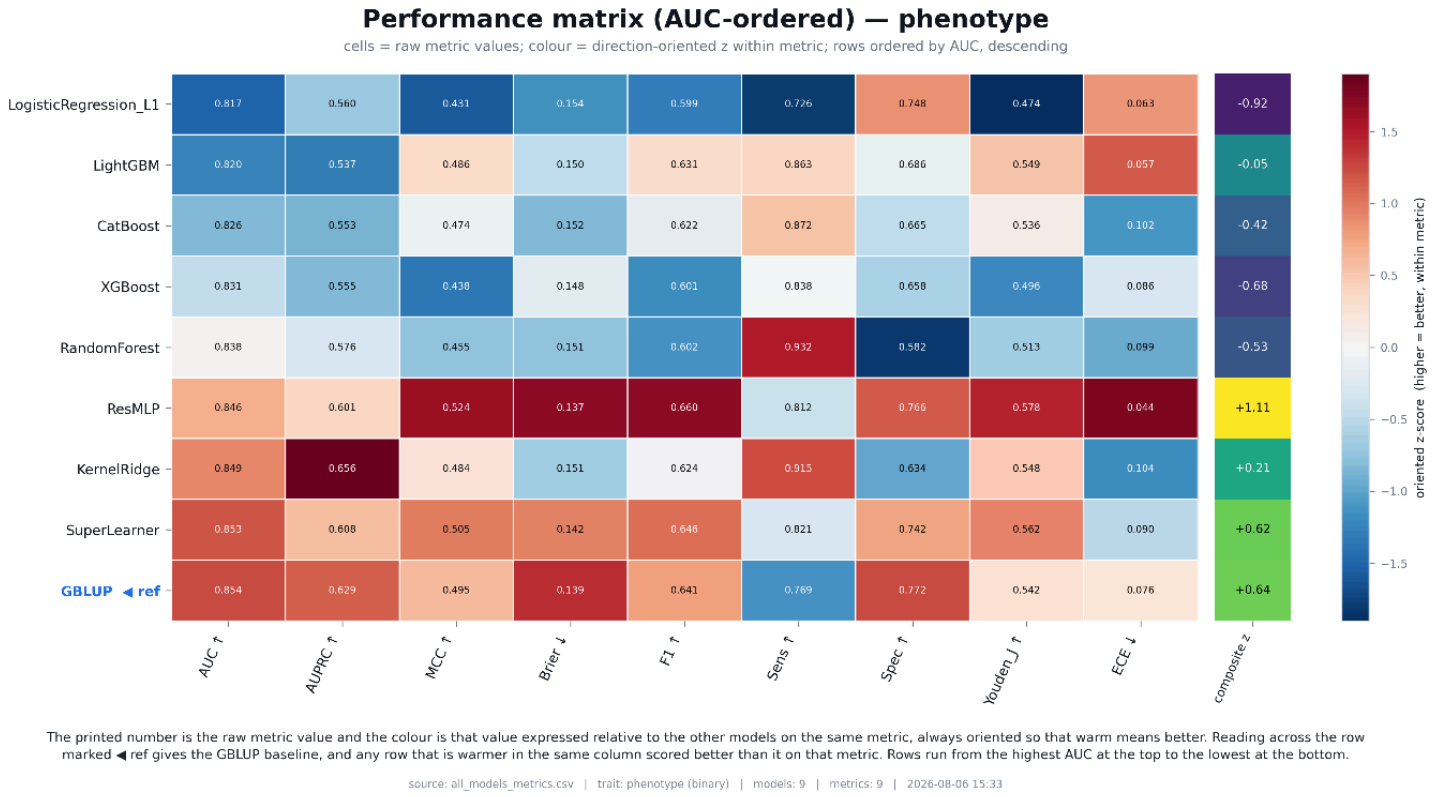



**Figure 1**

**Cross-validated performance matrix (models × metrics), ordered by AUC.**
Out-of-fold performance of the eight base learners and the Super Learner across nine metrics. Each cell shows the raw metric value; the colour encodes the direction-oriented z-score of that value within its metric column (warmer = better, cooler = worse), so colours are comparable down a column but not across columns.

*AUC, area under the ROC curve; AUPRC, area under the precision–recall curve; MCC, Matthews correlation coefficient; F1, F1 score; Sens., sensitivity; Spec., specificity; Youden J, Youden's index; ECE, expected calibration error. Brier and ECE are error metrics (lower is better) and were sign-inverted before orientation. The rightmost column ("composite z") is the unweighted mean of the oriented z-scores across the nine metrics; because it is standardized within the compared model set, it ranks models descriptively and is not a test statistic.*

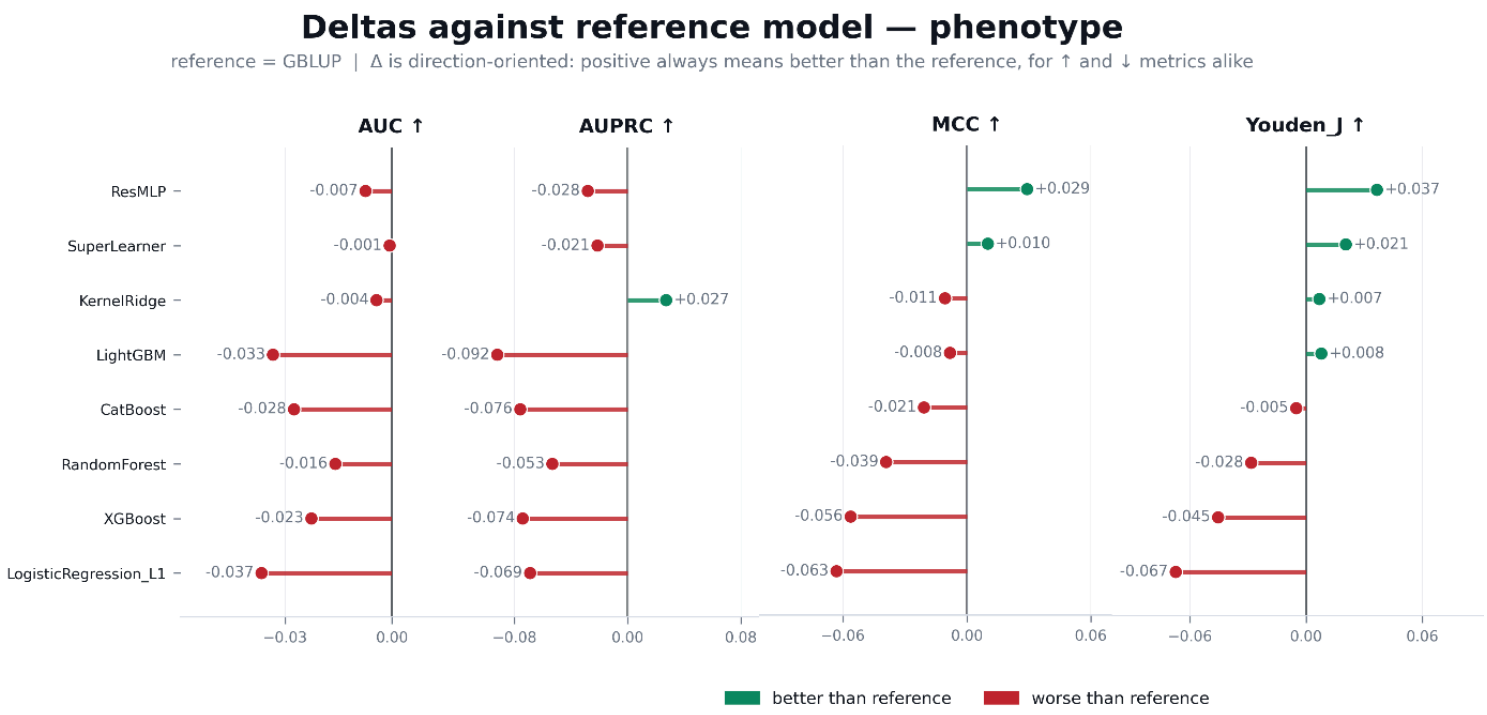



**Figure 2**

**Metric-wise differences of each learner relative to GBLUP.**
Direction-oriented deltas against GBLUP as the reference model for AUC, AUPRC, MCC and Youden's J; positive values (green) indicate performance better than GBLUP and negative values (red) worse. Most learners lie at or slightly below GBLUP, with only ResMLP and the Super Learner modestly exceeding it on MCC and Youden's J.

*Deltas are oriented so that positive always means better than the reference for both higher-is-better (↑) and lower-is-better metrics. Differences are descriptive; no statistical test of the pairwise differences is implied. Abbreviations as in Figure 1.*

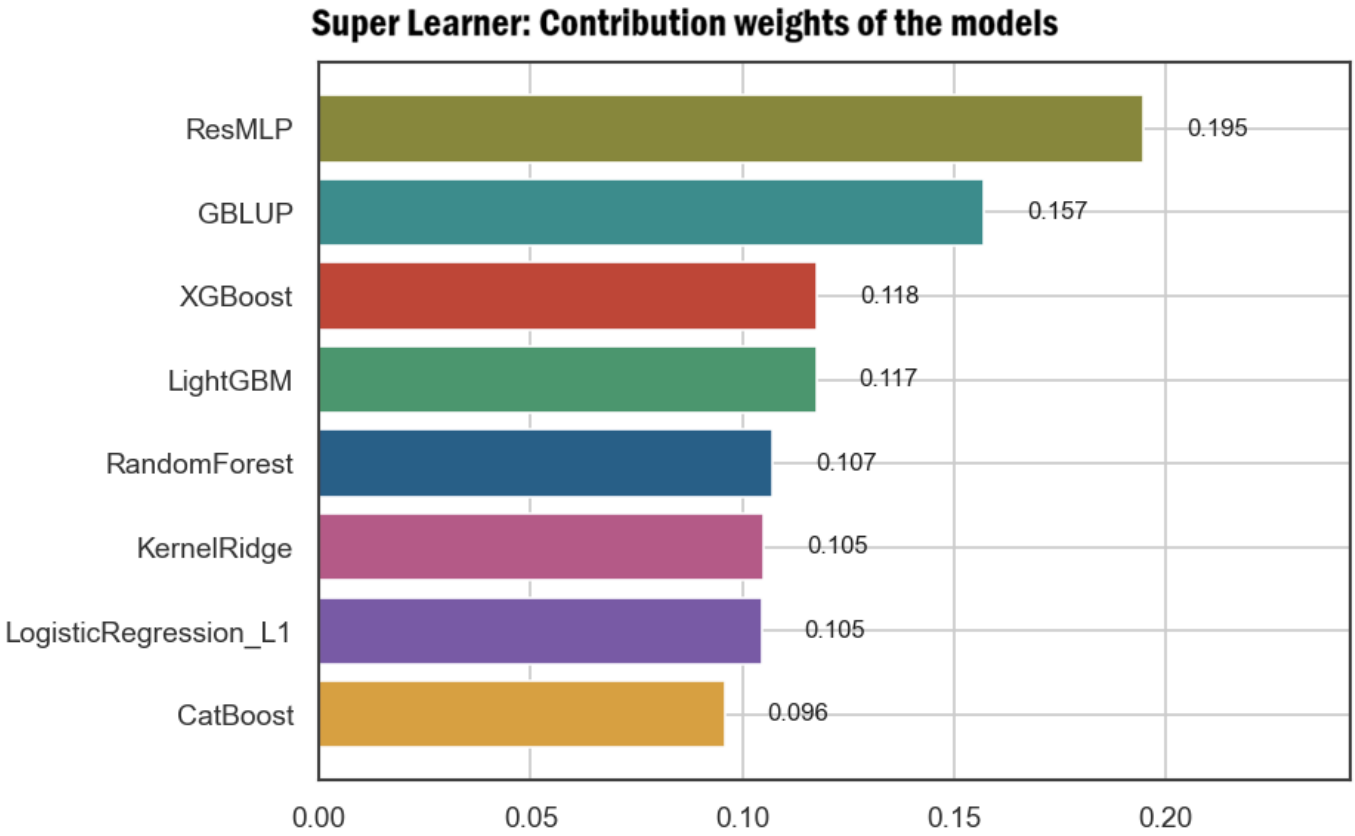



**Figure 3**

**Super Learner: contribution weights of the base learners.**
Normalized L2-regularized logistic combiner weights for each of the eight base learners, summing to one; these are the same weights used as performance-based reliability weights in the cross-model consensus scoring (Section 2.7). ResMLP (0.195) and GBLUP (0.157) receive the largest weights, but the distribution is broad rather than concentrated on a single learner (Herfindahl–Hirschman index = 0.133; normalized entropy = 0.986).

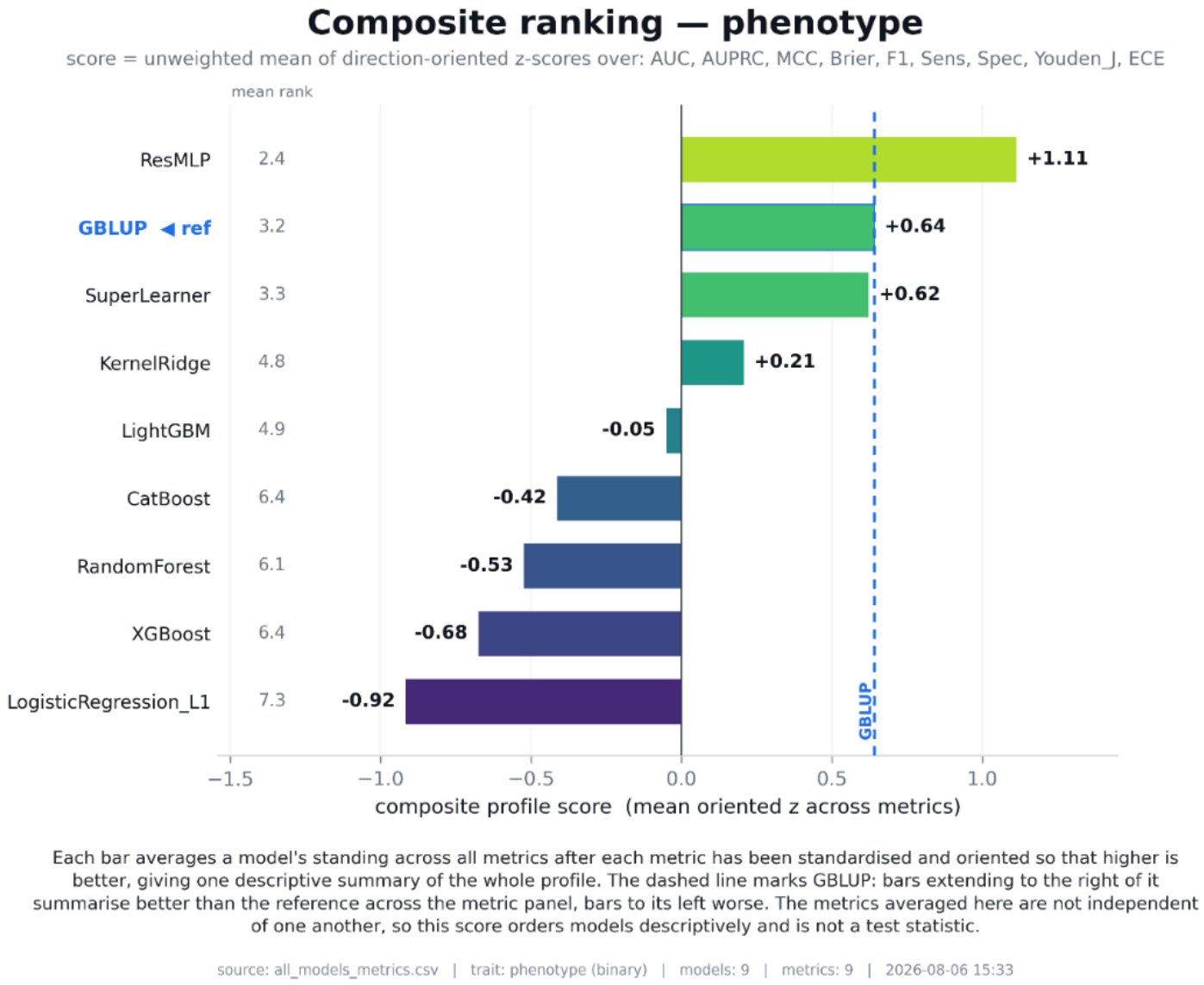



**Figure 4**

**Composite ranking of the models.**

Descriptive multi-metric summary: for each model, the composite profile score is the unweighted mean, across the nine metrics, of that model's direction-oriented z-score within each metric. Bars to the right of the dashed reference line (GBLUP) summarize better than GBLUP across the metric panel, and bars to its left worse; the mean rank across metrics is shown at left. ResMLP has the most favourable overall profile, while the Super Learner (z = 0.62) ranks just below GBLUP (z = 0.64).

*Metrics averaged: AUC, AUPRC, MCC, Brier, F1, sensitivity, specificity, Youden's J and ECE (Brier and ECE sign-inverted). The averaged metrics are not independent, so the composite orders models descriptively and is not a test statistic. The score is standardized within the compared model set and therefore has no absolute scale.*