

Supplementary Table S1. Full hyperparameter grid and software versions.

All feature selection and model training used the outer training partition exclusively within each of the 15 outer nested CV folds. The held-out outer test fold was not accessed at any prior step.

Component	Parameter	Value(s)
Variance pre-filter	Top-K features retained	min(10 x n_train, 2000)
	Applied to	Outer training partition only
Elastic Net	Library	scikit-learn 1.4
	Solver	SAGA
	Penalty	elasticnet
	C (regularisation)	(0.01, 0.03, 0.10, 0.32, 1.0, 3.2, 10.0, 100.0)
	l1_ratio	(0.1, 0.9)
	Inner CV	5-fold stratified
	max_iter	2000
	class_weight	balanced
	Genes selected	Top-30 by absolute coefficient
XGBoost	Library	xgboost 2.0
	n_estimators	100
	max_depth	4
	learning_rate	0.05
	subsample	0.8
	colsample_bytree	0.8
	scale_pos_weight	Per-fold positive:negative class ratio
	Genes selected	Top-30 by gain importance
ENOXGB panel	Base model	Logistic regression
	Features	Union of top-30 EN and top-30 XGB genes
	Trained on	Outer training fold only
	Evaluated on	Held-out outer test fold
Outer CV	Strategy	RepeatedStratifiedKFold
	n_splits	5
	n_repeats	3

	Total outer folds	15
	random_state	42
Consensus panel	Selection threshold	>=33% of outer folds
Software	Python	3.10
	pandas	2.x
	numpy	1.x
	scikit-learn	1.4
	xgboost	2.0
	mygene	3.2
	matplotlib	3.x
	joblib	Outer fold parallelisation