

Table S1: Performance of Extra tree classifier on training and validation dataset developed using 17 types of composition-based features for gram-positive ABPs

Feature Type	Training set						Validation set					
	Sn	Sp	Acc	AUC	AUPRC	MCC	Sn	Sp	Acc	AUC	AUPRC	MCC
AAC	85.1	85.8	85.4	0.93	0.93	0.71	83.9	90.9	87.4	0.94	0.94	0.75
DPC	83.9	84.0	83.9	0.91	0.92	0.68	83.3	88.7	86.0	0.93	0.92	0.72
ATC	75.9	75.5	75.7	0.84	0.84	0.52	67.7	82.3	75.0	0.84	0.83	0.51
BTC	67.9	67.7	67.8	0.75	0.74	0.36	65.1	63.4	64.3	0.70	0.70	0.29
CTC	80.0	79.0	79.5	0.88	0.88	0.59	81.7	86.6	84.1	0.90	0.90	0.68
PCP	81.7	82.0	81.9	0.91	0.91	0.64	82.8	89.3	86.0	0.94	0.93	0.72
AAI	83.5	83.7	83.6	0.91	0.91	0.67	81.2	89.3	85.2	0.92	0.92	0.71
RRI	80.9	79.8	80.4	0.89	0.90	0.61	80.1	85.5	82.8	0.91	0.91	0.66
PRI	78.8	78.9	78.8	0.87	0.88	0.58	0.6	76.3	81.5	0.88	0.87	0.63
DDR	82.4	82.1	82.3	0.91	0.91	0.65	79.6	87.6	83.6	0.93	0.91	0.67
SEP	62.0	55.8	58.9	0.61	0.58	0.18	52.2	64.0	58.1	0.60	0.59	0.16
SER	83.9	83.3	83.6	0.92	0.92	0.67	78.5	90.3	84.4	0.92	0.92	0.69
SPC	82.3	81.5	81.9	0.90	0.90	0.64	81.2	89.3	85.2	0.92	0.91	0.71
PAAC	84.5	85.2	84.9	0.93	0.93	0.70	81.7	90.3	86.0	0.93	0.93	0.72
APAAC	85.5	84.7	85.1	0.93	0.93	0.70	83.9	88.2	86.0	0.93	0.92	0.72

QSO	84.7	85.1	84.9	0.92	0.92	0.70	82.8	90.3	86.6	0.93	0.92	0.73
SOC	60.0	60.0	60.0	0.66	0.67	0.20	58.6	67.7	63.2	0.64	0.63	0.27

Sn: Sensitivity, **Sp**: Specificity, **Acc**: Accuracy, **MCC**: Matthews Correlation Coefficient, **AUC**: Area Under the Receiver Operating Characteristic curve, **AUPRC**:Area Under the Precision-Recall Curve, **AAC**: Amino acid composition, **APAAC**: Amphiphilic pseudo amino acid composition, **DDR**: Distance distribution of residue, **DPC**: Di-peptide composition, **QSO**: Quasi-sequence order, **PCP**: Physico-chemical properties composition, **PAAC**: Pseudo amino acid composition, **RRI**: Residue repeat Information, **SPC**: Shannon entropy of physicochemical properties, **ATC**: Atomic composition , **BTC**: Bond type composition, **CTC**: Conjoint triad descriptors, **AAI**: Amino Acid index, **PRI**: Property repeats index, **SEP**: Shannon entropy of a protein, **SER**: Shannon entropy of a residue, **SOC**: Sequence order coupling number