

Cell Line	Model	AUROC Mean	AUROC Standard Deviation	AUPRC Mean	AUPRC Standard Deviation
E116	GC-MERGE	0.905	0.001	0.880	0.002
	Shuffled	0.903	0.000	0.879	0.002
	CNN	0.896	0.002	0.865	0.004
	MLP	0.899	0.001	0.870	0.004
	RF-class	0.901	0.002	0.864	0.003
	Logistic	0.900	0.000	0.900	0.000
E123	GC-MERGE	0.919	0.001	0.892	0.001
	Shuffled	0.914	0.001	0.886	0.002
	CNN	0.907	0.003	0.867	0.006
	MLP	0.917	0.002	0.889	0.004
	RF-class	0.929	0.001	0.901	0.003
	Logistic	0.921	0.000	0.921	0.000
E122	GC-MERGE	0.903	0.001	0.876	0.002
	Shuffled	0.899	0.001	0.866	0.002
	CNN	0.889	0.002	0.847	0.008
	MLP	0.899	0.001	0.864	0.003
	RF-class	0.890	0.001	0.852	0.005
	Logistic	0.879	0.000	0.879	0.000

Table S4: **Model comparisons by AUROC and AUPRC for the classification task.** GC-MERGE as well as the alternative deep learning baselines - shuffled, CNN, and MLP - are described in the main text. In addition, we include two other machine learning baselines: a random forest classifier (RF-class) and a logistic regression model (Logistic). Note that with the exception of E123, in which the random forest classifier's performance exceeds that of GC-MERGE, our model consistently achieves higher scores than all other baselines.