Figure S1. Increased RNA count per cell increases probability of correct classification for all models. Logistic regression of ventral tegmental area test partition cell UMI (unique molecular index) count and model classification (correct: 1, incorrect: 0) for chromosome Y (**a**), Xist (**b**), logistic regression (**c**), support vector machine (**d**), random forest (**e**), and multilayer-perceptron (**f**) models. Dashed red lines indicate the RNA count corresponding to a 95% probability of correct classification by a model.

Table S1. Increased UMI count improves the likelihood of accurate cell sex classification. The relationship between transcript UMI (unique molecular index) count and accuracy of model classification (incorrect: 0, correct: 1) in the ventral tegmental area test partition data was modeled using logistic regression for all cell sex classification models. Model: model classifications used to fit logistic regression; Intercept: fit logistic regression model intercept; UMI: fit logistic regression model coefficient for UMI count; Estimate: model estimate for either intercept or UMI terms; StdErr: standard error of either Intercept or UMI terms; Z-score: Estimate / Std Error; pval: p-value associated with the value Z-score column; 95% Probability Intercept: UMI value of where probability of correct classification reaches 95% as predicted by fit model.

Model	Intercept				UMI count				95% Probability Intercept
	Estimate	StdErr	Z-score	pval	Estimate	StdErr	Z-score	pval	
Chr Y	0.969	0.074	13.121	2.49E-39	3.17E-04	2.21E-05	14.364	8.68E-47	6228
Xist	0.194	0.044	4.462	8.14E-06	1.33E-04	8.48E-06	15.635	4.20E-55	20747
LR	1.245	0.083	14.919	2.49E-50	3.29E-04	2.58E-05	12.770	2.43E-37	5157
SVM	1.328	0.084	15.852	1.36E-56	3.12E-04	2.54E-05	12.307	8.26E-35	5175
RF	1.029	0.085	12.161	5.02E-34	3.90E-04	2.79E-05	13.964	2.59E-44	4914
MLP	1.393	0.074	18.747	2.04E-78	2.31E-04	1.98E-05	11.662	2.00E-31	6706