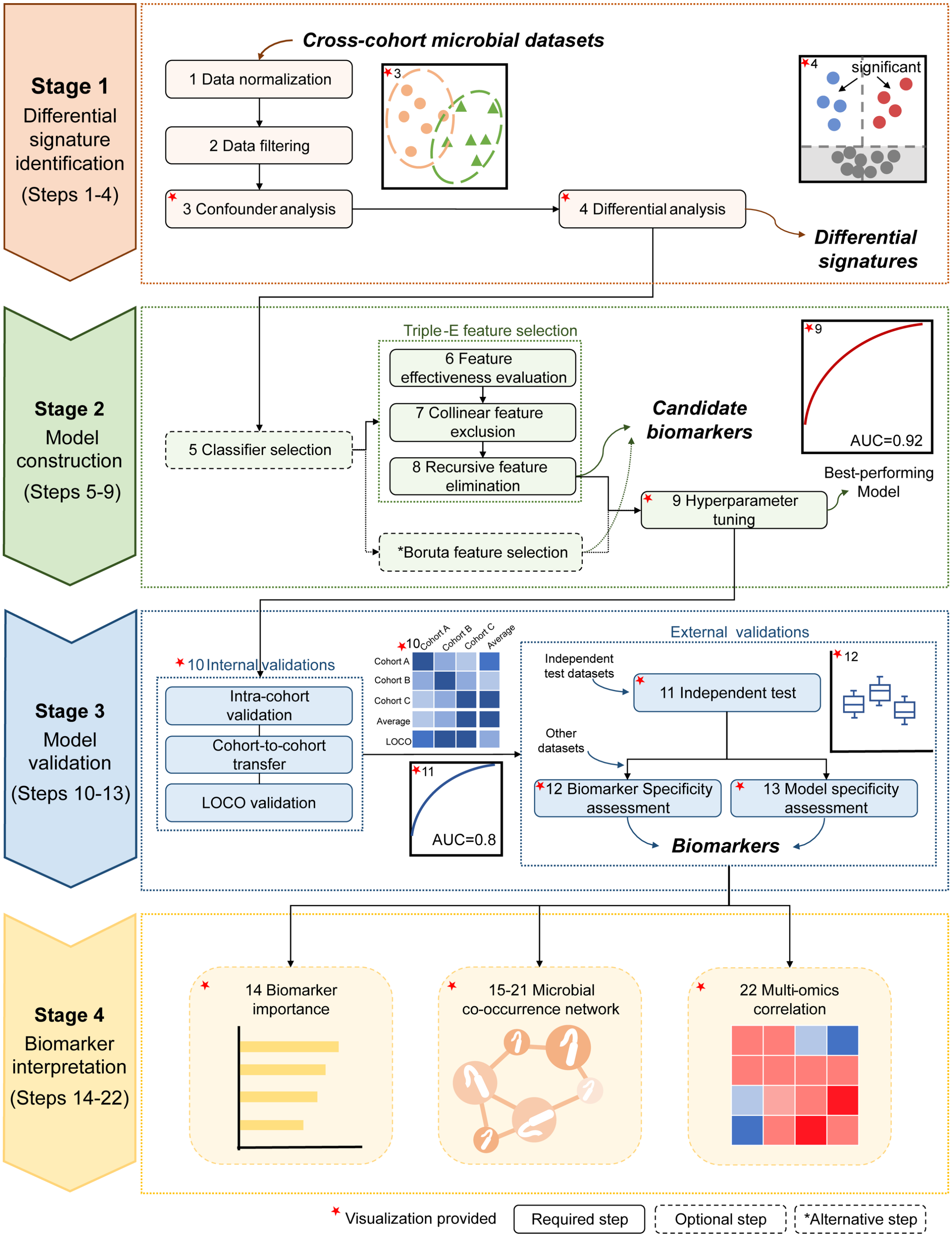
nature protocols

https://doi.org/10.1038/s41596-024-00999-9

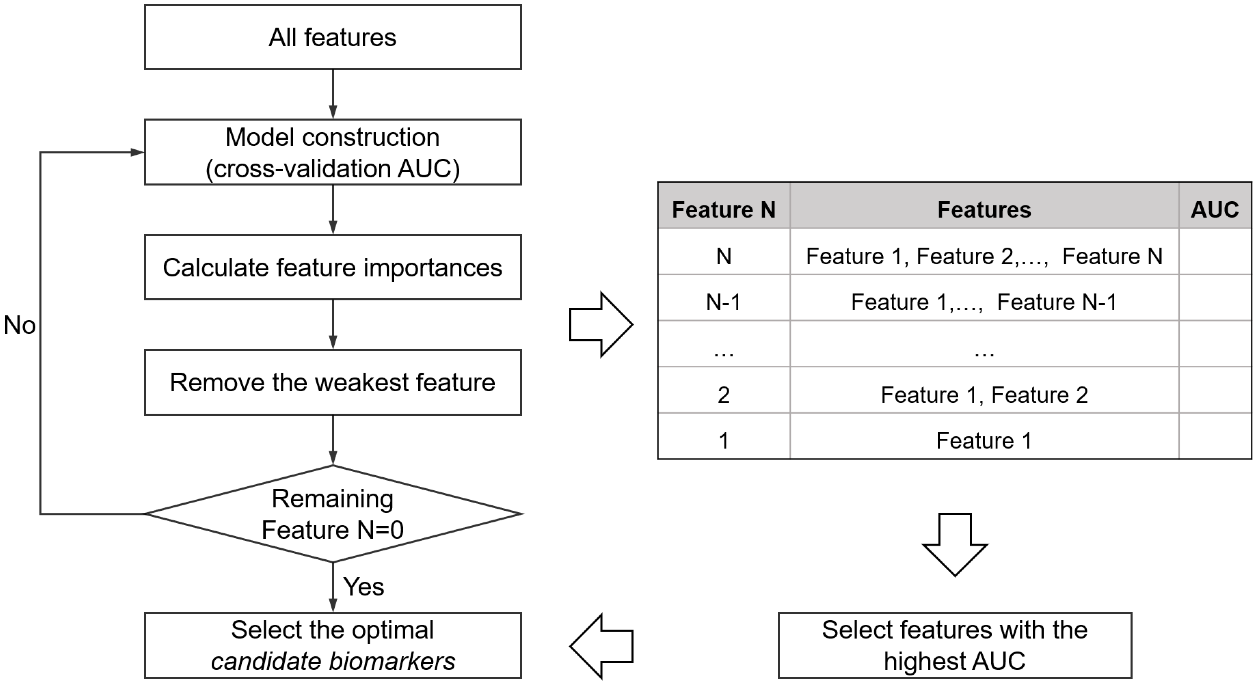
Protocol

Check for updates

Identification and validation of microbial biomarkers from cross-cohort datasets using xMarkerFinder



Supplementary Figure 1. Overview of the xMarkerFinder workflow. Four main stages of xMarkerFinder: differential signature identification (steps 1-4), model construction (steps 5-9), model validation (Steps 10-13), and biomarker interpretation (steps 14-22).



Feature N	Features	AUC
N	Feature 1, Feature 2,…, Feature N	
N-1	Feature 1,…, Feature N-1	
…	…	
2	Feature 1, Feature 2	
1	Feature 1	

Supplementary Figure 2. Recursive feature elimination. The third step of Triple-E feature selection procedure aiming at determining the minimal panel of candidate biomarkers, achieved by recursively removing the weakest feature for model construction per loop to obtain the best panel of features with the highest cross-validation AUC value.

Supplemenraty Table 1. Comparison between xMarkerFinder and other tools

		xMarkerFinder	**LEfSe**	**NetMoss**	**SIAMCAT**	**MIIDL**
Adenoma vs Control	Feature No.	15	72	428	10	1389
	Cross-validation AUC	0.75	0.66	0.61	0.64	0.68
EBW-CRC vs EBW-Control	Feature No.	15	104	313	444	606
	Cross-validation AUC	0.9	0.86	0.73	0.82	0.52
OSCC vs Control	Feature No.	17	60	248	6	60
	Cross-validation AUC	0.85	0.83	0.77	0.75	0.42