

Additional File 4: A snippet of Robust rule patterns discovered for breast and colon cancers. Each row contain rule pattern information, including the learned models of the rule, the rule identifier, the source of rule origin, and the rule patterns. Pr denotes the patterns origin from prior rules, and Nr denotes new rules.

Breast	Rule No.	Source	Patterns		
GEO15852	01.	Pr	IF (KRT18 = High)	THEN	(Class = Case)
	02.	Pr	IF (KRT18 = Low)	THEN	(Class = Control)
Colon	Rule No.	Source	Patterns		
GEO20916, GEO24514	01.	Pr	IF (MMP7 = High)	THEN	(Class = Case)
	02.	Pr	IF (MMP7 = Low)	THEN	(Class = Control)
	03.	Nr	IF (CDH3 = High)	THEN	(Class = Case)
	04.	Nr	IF (CDH3 = Low)	THEN	(Class = Control)
GEO23878	05.	Nr	IF ((ABCA8 = Low)	THEN	(Class = Case)
	06.	Nr	IF (ABCA8 = High)	THEN	(Class = Control)
	07.	Pr	IF (CXCL12 = Low)	THEN	(Class = Case)
	08.	Pr	IF (CXCL12 = High)	THEN	(Class = Control)
	09.	Pr	IF (LPAR1 = Low)	THEN	(Class = Case)
	10.	Pr	IF (LPAR1 = Low)	THEN	(Class = Control)
GEO24514	11.	Pr	IF (MCM2 = Low)	THEN	(Class = Control)
	12.	Pr	IF (MCM2 = High)	THEN	(Class = Case)
	13.	Pr	IF (PMAIP1 = Low)	THEN	(Class = Control)
	14.	Pr	IF (PMAIP1 = High)	THEN	(Class = Case)
	15.	Nr	IF (TEX10 = High)	THEN	(Class = Case)
	16.	Nr	IF (TEX10 = Low)	THEN	(Class = Control)
GEO9348	17.	Pr	IF (INHBA = High)	THEN	(Class = Case)
	18.	Pr	IF (INHBA = Low)	THEN	(Class = Control)
GEO10715	19.	Pr	IF (LILRB4 = High)	THEN	(Class = Control)
	20.	Pr	IF (LILRB4 = Low)	THEN	(Class = Control)