

Table S1. The primer sequences used to amplify the 5 genes.

Gene Symbol	Forward primer (5'->3')	Reverse primer (5'->3')	Product length (bp)
ZNF695	ACAGCCAGACACTCAGTTTTGT	TCCACTCACCCACAATTTCCC	160
CENPA	TAGGCGCTTCCTCCCATCAA	CACATATTTCTCTTGCCAGGCG	129
TORAP	TGTCCATCACCCCTTTGGACG	GGATACCGATGCGCCAGG	214
BIRC5	TGACGACCCCATAGAGGAAC	CGCACTTTCTCCGCAGTTTC	186
KIF20A	GTGGCACGTCTTCGGACCTA	TTCTTGCGTACCACAGACCC	149

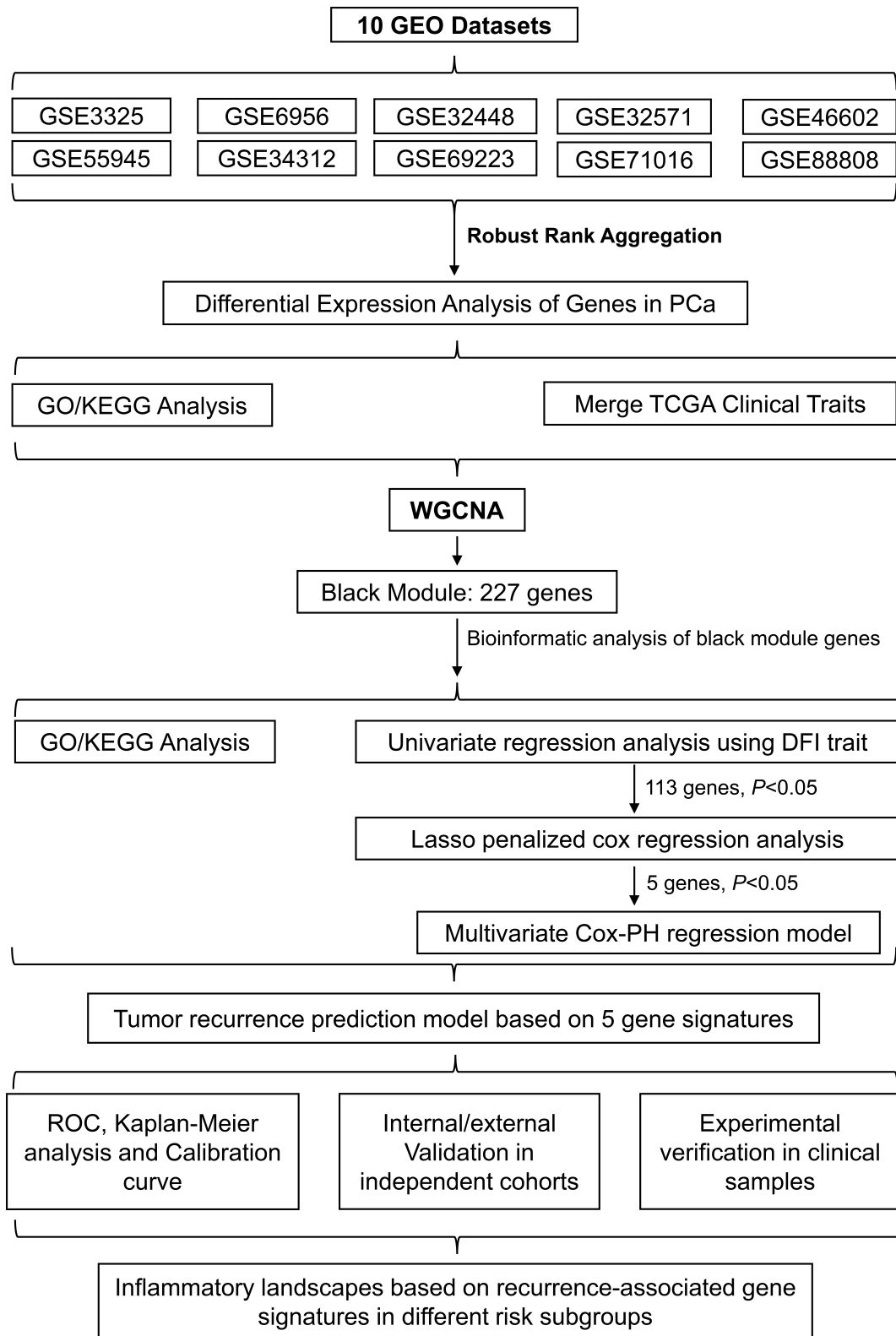


Figure S1. The flow chart of the analysis procedure.

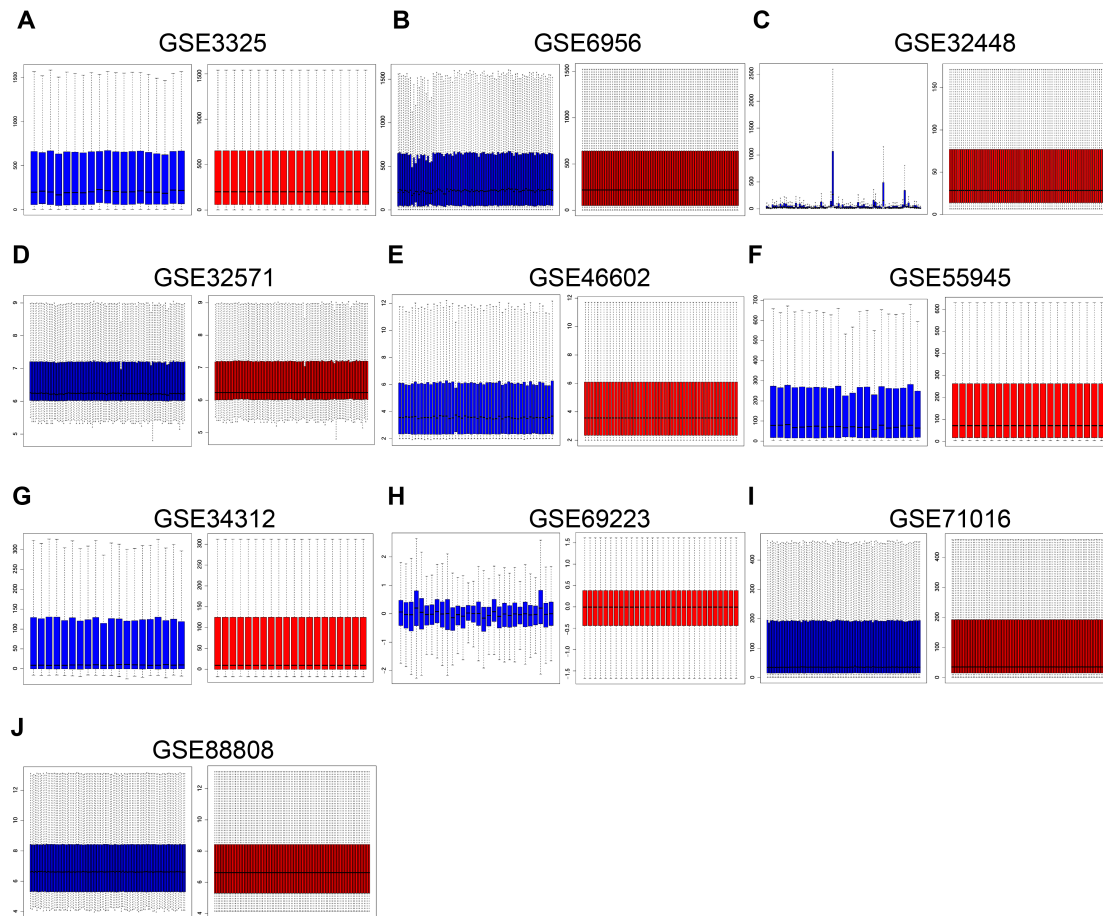


Figure S2. (A-J) RawBox graph shows normalized processing on GEO datasets.

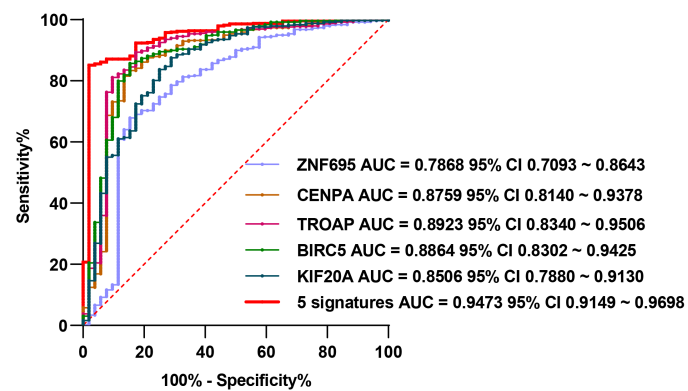


Figure S3. The combined diagnosis model based on 5 gene signatures shows higher diagnosis efficiency in TCGA cohort.

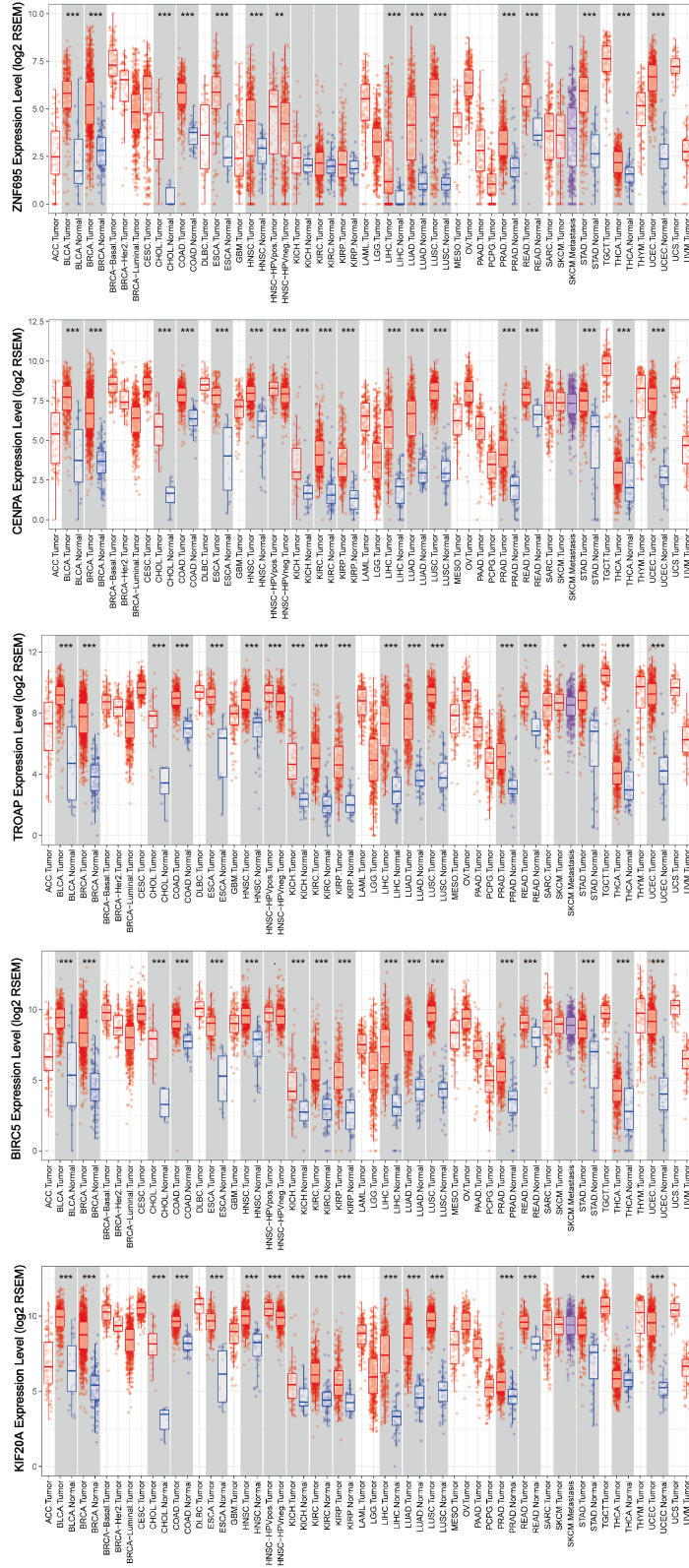


Figure S4. Assessment the imbalance of these 5 gene signatures in other tumors based on TCGA database algorithm.