Supplementary

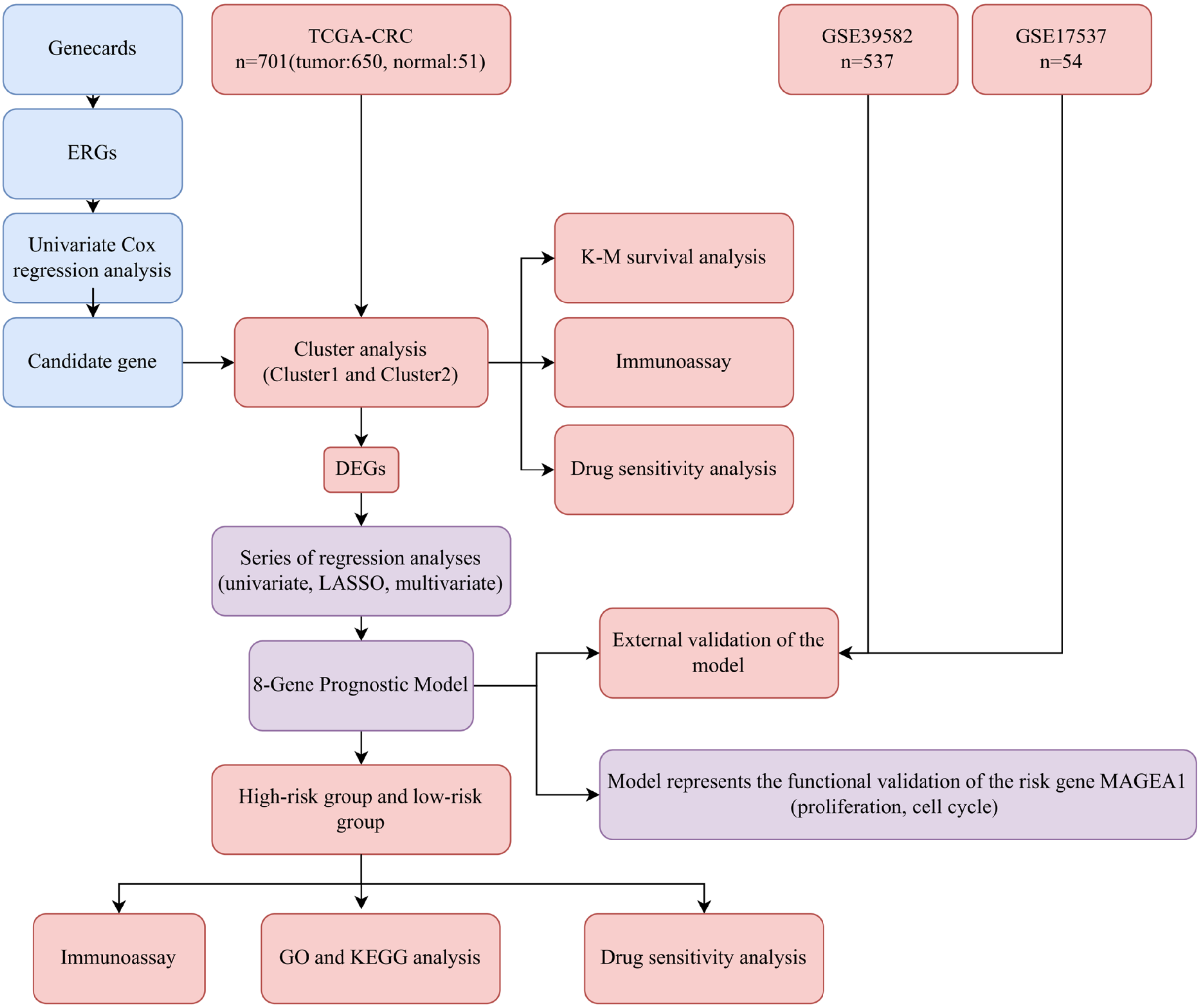

Figure S1 Flow chart of the study.

https://dx.doi.org/10.21037/jgo-2025-593

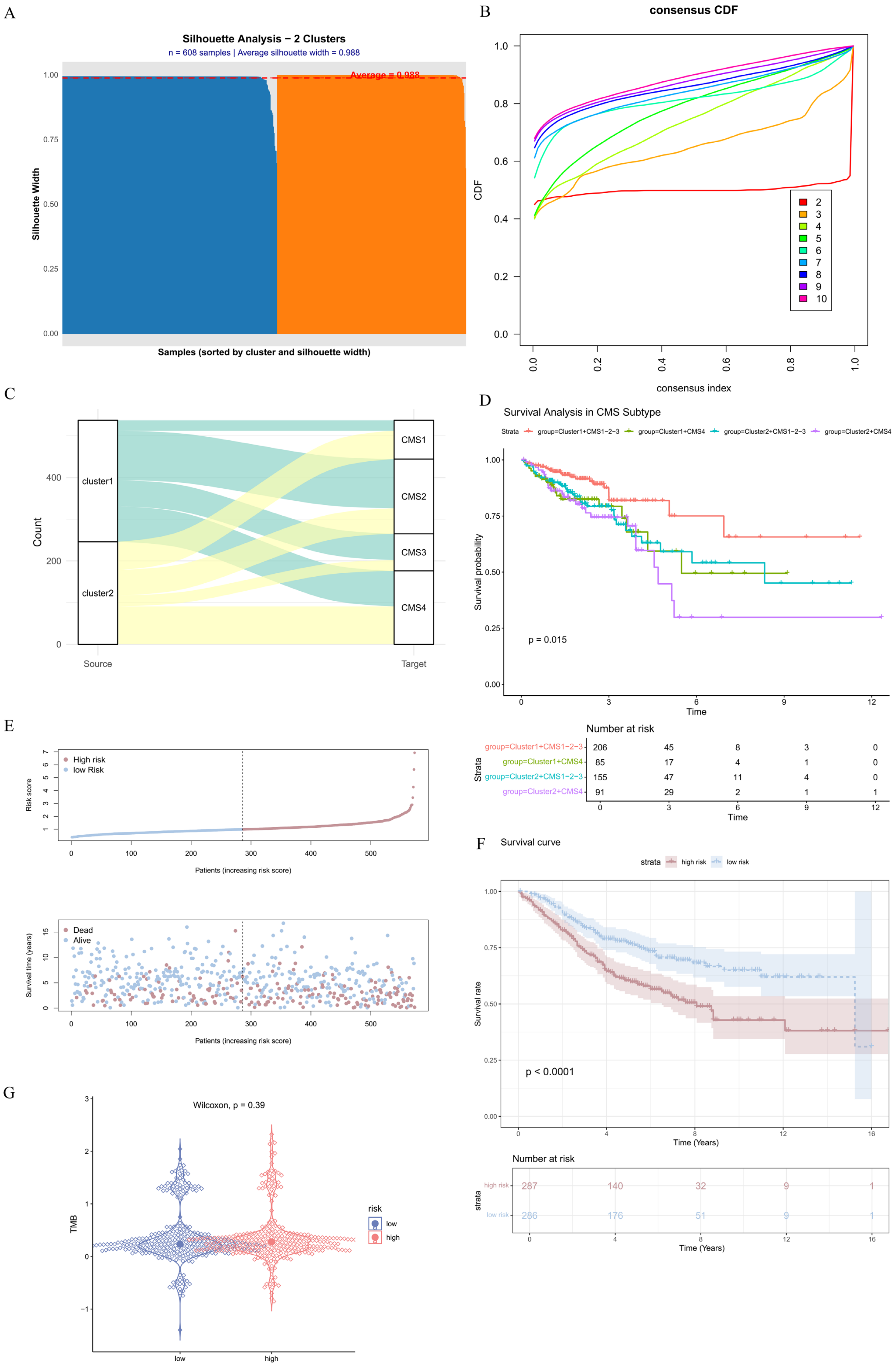

Figure S2 Determination of ERG molecular typing, association with CMS typing, and validation of prognosis and mutation load (A) sample outline width graph when K=2. (B) Consensus cumulative distribution function (CDF) curve. (C) Sankey diagram showing the correspondence between cluster 1 and cluster 2 and CMS1-4. (D) K-M survival curves of patients belonging to CMS1-3 subtypes within cluster 1 and cluster 2 and those belonging to CMS4 subtype. (E) risk score distribution and survival outcome statistics of the validation set GSE39582. A total of 190 deaths occurred, including 116 in the high-risk group and 74 in the low-risk group. (F) Survival curve of the validation set GSE39582. (G) Comparison of TMB (tumor mutation load) between high-risk and low-risk groups.

Table S1 13 key genes with potential prognostic value

Gene	Coef
SALL1	-0.050068
PRG4	0.0838879722652723
FTCD	0.0220505531961425
UPK3B	0.0437714482491519
RBP7	0.0444395203952587
CXCL13	-0.004961
KRT84	0.193342283305005
MAGEA1	0.0410112562331138
MMP1	-0.041344
PCOLCE2	0.10214457570025
ALPP	0.0222404998085306
HOXC6	0.0775678365009155
FOXD1	0.0513109420042971