ORIGINAL ARTICLE

Integrated Network Analysis to Determine *CNN1*, *MYL9*, *TAGLN*, and *SORBS1* as Potential Key Genes Associated with Prostate Cancer

Changtao Li [1], Lijuan Pang [1], Fangfang Jin [1], Yongcheng Song [1], Da Zhou [1], Yuxuan Song [1], Yimin Li [1], Shan Jin [1], Lu Zhang [1], Weihua Liang [1], Xihua Shen [1], Jun Li [1,3], Bingyang She [1], Chengyan Wang [1], Luping Ma [1,2]

[1] NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases (First Affiliated Hospital, School of Medicine, Shihezi University) / Department of Pathology and Key Laboratory for Xinjiang Endemic and Ethnic Diseases, Shihezi University School of Medicine, Shihezi, China
[2] Department of Urology, the First Affiliated Hospital of Medical College, Shihezi University, Shihezi, China
[3] Department of Ultrasound, the First Affiliated Hospital of Medical College, Shihezi University, Shihezi, China

SUMMARY

Background: Prostate cancer (PCa) is challenging to treat. It is necessary to screen for related biological markers to accurately predict the prognosis and recurrence of prostate cancer.
Methods: Three data sets, GSE28204, GSE30521, and GSE69223, from the Gene Expression Omnibus (GEO) database were integrated into this study. After the identification of differentially expressed genes (DEGs) between PCa and normal prostate tissues, network analyses including protein-protein interaction (PPI) network, and weighted gene co-expression network analysis (WGCNA) were used to select hub genes. Gene Ontology (GO) term analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed to annotate the functions of DEGs and hub modules of the networks. Survival analysis was performed to validate the correlation between the key genes and PCa relapse.
Results: In total, 867 DEGs were identified, including 201 upregulated and 666 downregulated genes. Three hub modules of the PPI network and one hub module of the weighted gene co-expression network were determined. Moreover, four key genes (*CNN1, MYL9, TAGLN,* and *SORBS1*) were significantly associated with PCa relapse ($p < 0.05$).
Conclusions: CNN1, MYL9, TAGLN, and SORBS1 may be potential biomarkers for PCa development.
(Clin. Lab. 2023;69:1-5. DOI: 10.7754/Clin.Lab.2023.220224)

Correspondence:
Dr. Chengyan Wang
NHC Key Laboratory of Prevention and Treatment of
Central Asia High Incidence Diseases
(First Affiliated Hospital, School of Medicine
Shihezi University)
Department of Pathology and Key Laboratory for
Xinjiang Endemic and Ethnic Diseases
Shihezi University School of Medicine
Shihezi
China
Email: wangchengyanmlp@163.com

Luping Ma
NHC Key Laboratory of Prevention and Treatment of
Central Asia High Incidence Diseases
(First Affiliated Hospital, School of Medicine
Shihezi University)
Department of Pathology and Key Laboratory for
Xinjiang Endemic and Ethnic Diseases
Shihezi University School of Medicine
Shihezi
China
Email: 196930107@qq.com

_____

Supplementary Data

Table S1. Patient characteristics.

Sample GEO accession	Sample type	Sample organism	Gender	Age	PSA	Gleason score	Tumor staging	Series GEO accession
GSM698416	prostate cancer tissue	homo sapiens	male	54	15,87	5	NA	GSE28204
GSM698417	prostate cancer tissue	homo sapiens	male	77	8,26	8	NA	GSE28204
GSM698418	prostate cancer tissue	homo sapiens	male	70	19,56	6	NA	GSE28204
GSM698419	prostate cancer tissue	homo sapiens	male	80	399,1	8	NA	GSE28204
GSM698420	benign prostate tissue of prostate cancer	homo sapiens	male	54	15,87	normal human prostate tissue	NA	GSE28204
GSM698421	benign prostate tissue of prostate cancer	homo sapiens	male	77	8,26	normal human prostate tissue	NA	GSE28204
GSM698422	benign prostate tissue of prostate cancer	homo sapiens	male	60	NA	normal human prostate tissue	NA	GSE28204
GSM698423	benign prostatic hyperplasia tissue	homo sapiens	male	75	17,26	normal human prostate tissue	NA	GSE28204
GSM756911	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 3	NA	GSE30521
GSM756912	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 3	NA	GSE30521
GSM756913	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 4	NA	GSE30521
GSM756914	prostate cancer tissue	homo sapiens	male	NA	NA	5 + 4	NA	GSE30521
GSM756915	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 3	NA	GSE30521
GSM756916	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 3	NA	GSE30521
GSM756917	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 3	NA	GSE30521
GSM756918	normal human prostate tissue	homo sapiens	male	NA	NA	normal human prostate tissue	NA	GSE30521
GSM756919	normal human prostate tissue	homo sapiens	male	NA	NA	normal human prostate tissue	NA	GSE30521
GSM756920	normal human prostate tissue	homo sapiens	male	NA	NA	normal human prostate tissue	NA	GSE30521
GSM756921	normal human prostate tissue	homo sapiens	male	NA	NA	normal human prostate tissue	NA	GSE30521
GSM756922	prostate cancer tissue	homo sapiens	male	NA	NA	4 + 3	NA	GSE30521
GSM756923	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 3	NA	GSE30521
GSM756924	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 3	NA	GSE30521
GSM756925	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 4	NA	GSE30521
GSM756926	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 4	NA	GSE30521
GSM756927	prostate cancer tissue	homo sapiens	male	NA	PSA	3 + 4	NA	GSE30521
GSM756928	prostate cancer tissue	homo sapiens	male	NA	NA	4 + 5	NA	GSE30521
GSM756929	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 4	NA	GSE30521
GSM756930	prostate cancer tissue	homo sapiens	male	NA	NA	4 + 5	NA	GSE30521
GSM756931	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 4	NA	GSE30521
GSM756932	normal human prostate tissue	homo sapiens	male	NA	NA	normal human prostate tissue	NA	GSE30521
GSM756933	prostate cancer tissue	homo sapiens	male	NA	NA	3 + 4	NA	GSE30521
GSM1695583	PCa	homo sapiens	male	63	NA	NA	pT3	GSE69223
GSM1695584	adj. normal	homo sapiens	male	63	NA	NA	pT3	GSE69223
GSM1695585	PCa	homo sapiens	male	53	NA	NA	pT2	GSE69223
GSM1695586	adj. normal	homo sapiens	male	53	NA	NA	pT2	GSE69223
GSM1695587	PCa	homo sapiens	male	57	NA	NA	pT2	GSE69223
GSM1695588	adj.normal	homo sapiens	male	57	NA	NA	pT2	GSE69223
GSM1695589	PCa	homo sapiens	male	60	NA	NA	pT2	GSE69223

Table S1. Patient characteristics (continued).

| Sample GEO accession	Sample type	Sample organism	Gender	Age	PSA	Gleason score	Tumor staging	Series GEO accession
GSM1695590	adj. normal	homo sapiens	male	60	NA	NA	pT2	GSE69223
GSM1695591	PCa	homo sapiens	male	50	NA	NA	pT2	GSE69223
GSM1695592	adj. normal	homo sapiens	male	50	NA	NA	pT2	GSE69223
GSM1695593	PCa	homo sapiens	male	69	NA	NA	pT2	GSE69223
GSM1695594	adj. normal	homo sapiens	male	69	NA	NA	pT2	GSE69223
GSM1695595	PCa	homo sapiens	male	60	NA	NA	pT2	GSE69223
GSM1695596	adj. normal	homo sapiens	male	60	NA	NA	pT2	GSE69223
GSM1695597	PCa	homo sapiens	male	62	NA	NA	pT3	GSE69223
GSM1695598	adj. normal	homo sapiens	male	62	NA	NA	pT3	GSE69223
GSM1695599	PCa	homo sapiens	male	67	NA	NA	pT2	GSE69223
GSM1695600	adj. normal	homo sapiens	male	67	NA	NA	pT2	GSE69223
GSM1695601	PCa	homo sapiens	male	54	NA	NA	pT2	GSE69223
GSM1695602	adj. normal	homo sapiens	male	54	NA	NA	pT2	GSE69223
GSM1695603	PCa	homo sapiens	male	60	NA	NA	pT2	GSE69223
GSM1695604	adj. normal	homo sapiens	male	60	NA	NA	pT2	GSE69223
GSM1695605	PCa	homo sapiens	male	47	NA	NA	pT3	GSE69223
GSM1695606	adj. normal	homo sapiens	male	47	NA	NA	pT3	GSE69223
GSM1695607	PCa	homo sapiens	male	61	NA	NA	pT3	GSE69223
GSM1695608	adj. normal	homo sapiens	male	61	NA	NA	pT3	GSE69223
GSM1695609	PCa	homo sapiens	male	66	NA	NA	pT3	GSE69223
GSM1695610	adj. normal	homo sapiens	male	66	NA	NA	pT3	GSE69223
GSM1695611	PCa	homo sapiens	male	50	NA	NA	pT2	GSE69223
GSM1695612	adj. normal	homo sapiens	male	50	NA	NA	pT2	GSE69223
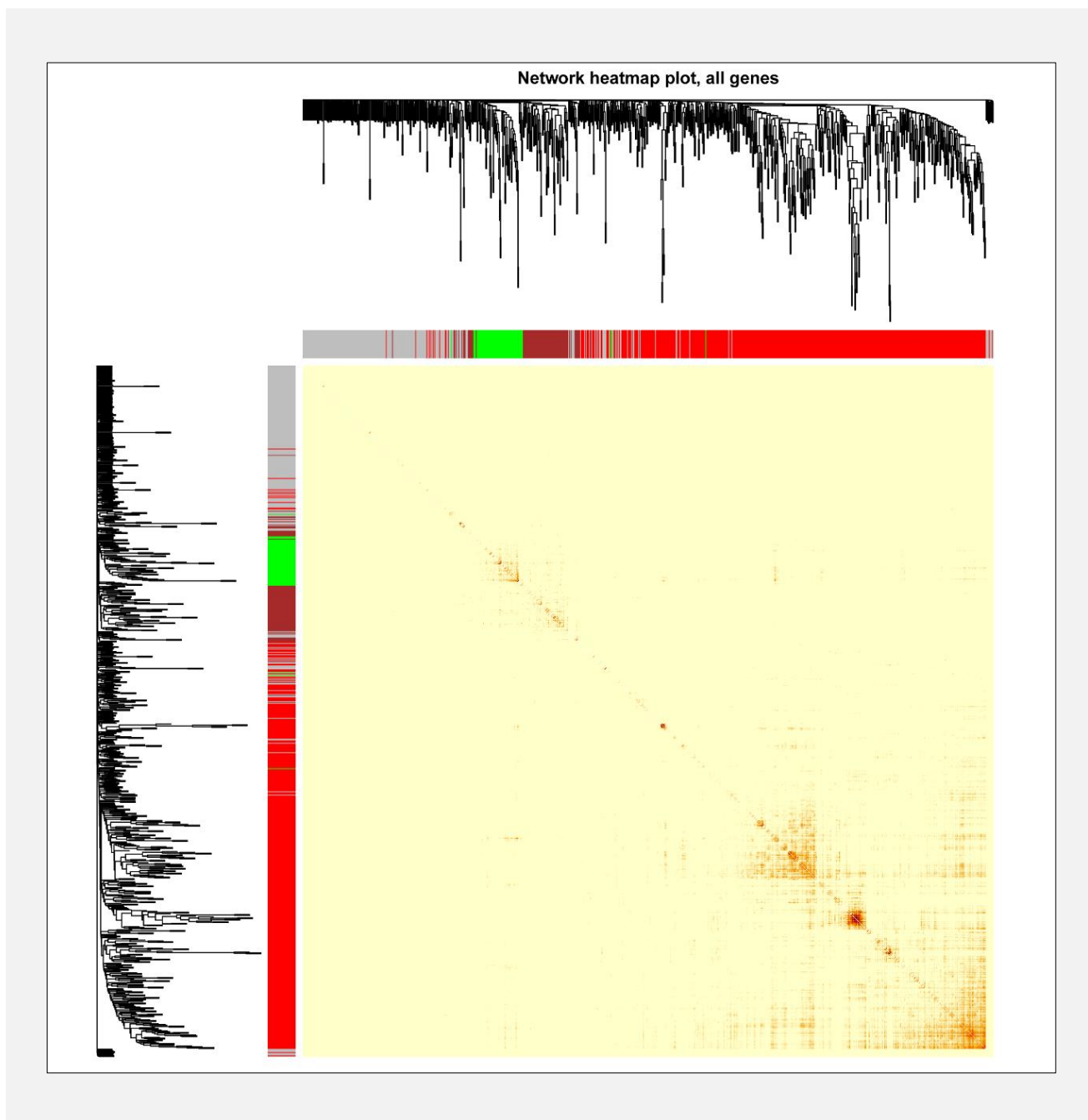
Figure S1. The weighted gene co-expression network visualized by a heatmap plot. The heatmap depicts the Topological Overlap Matrix (TOM) among all genes. Light color denotes low overlap and progressively darker red color denotes higher overlap. The left side and the top show genes dendrogram and modules assignment.
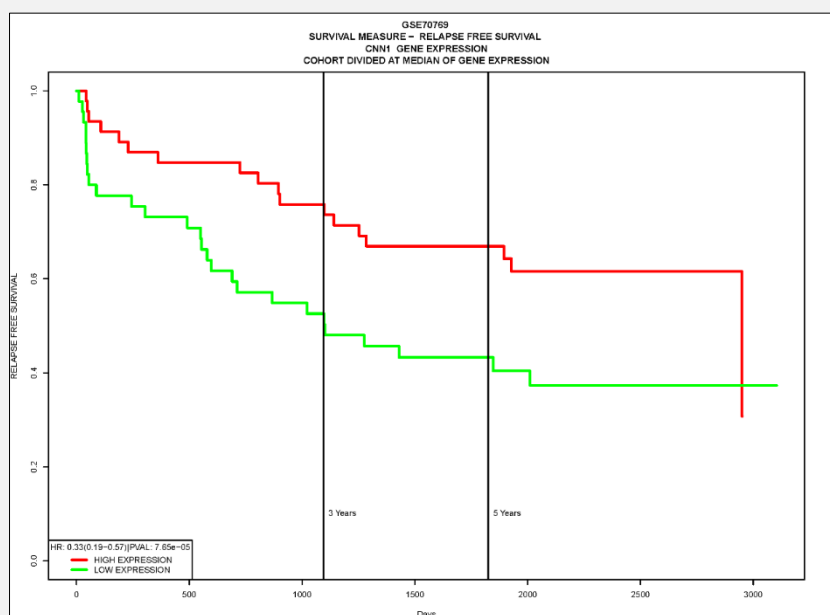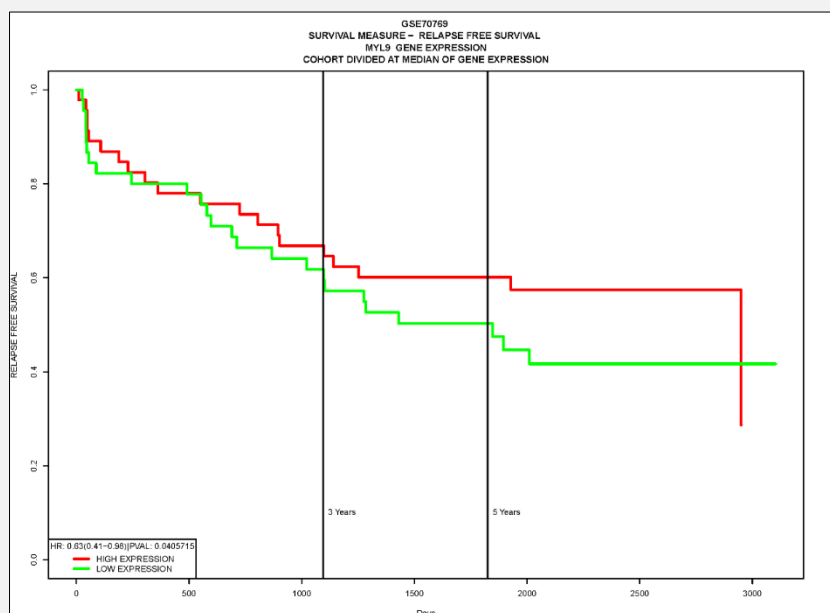
Figure S2. Kaplan–Meier survival curve of relapse-free survival between the high and low expression group in *CNN1*. Cohort is divided at median of gene expression.

HR (95% CI) - 0.33 (0.19 - 0.57), p-value < 0.05.

Figure S3. Kaplan–Meier survival curve of relapse-free survival between the high and low expression group in *MYL9*. Cohort is divided at median of gene expression.
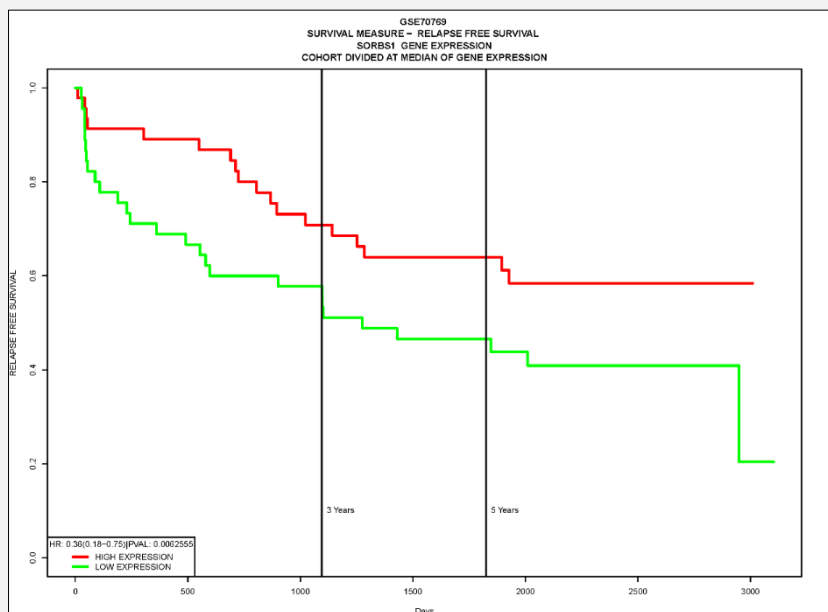
HR (95% CI) - 0.63 (0.41 - 0.98), p-value < 0.05.

Figure S4. Kaplan–Meier survival curve of relapse-free survival between the high and low expression group in *SORBS1*. Cohort is divided at median of gene expression.
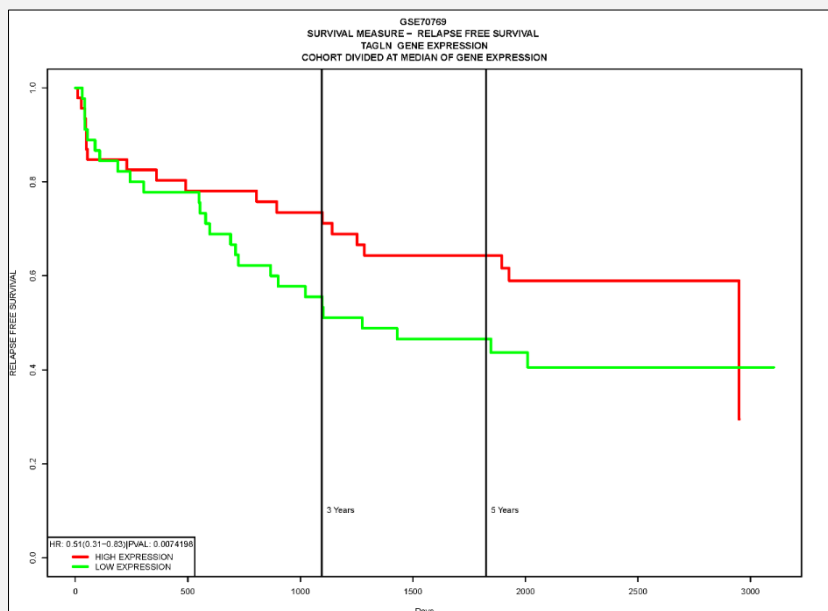
HR (95% CI) - 0.36 (0.18 - 0.75), p-value < 0.05.

Figure S5. Kaplan–Meier survival curve of relapse-free survival between the high and low expression group in *TAGLN*. Cohort is divided at median of gene expression.

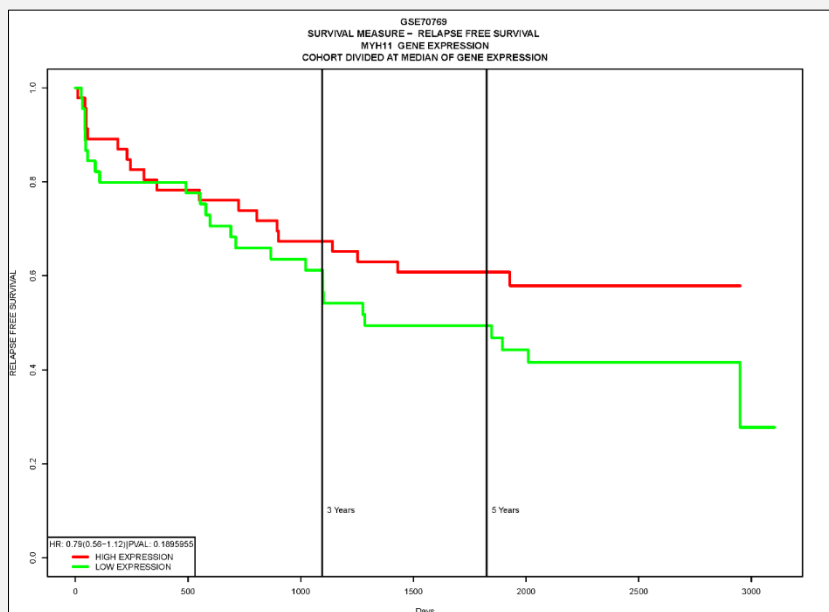**HR (95% CI) - 0.51 (0.31 - 0.83), p-value < 0.05.**

Figure S6. Kaplan–Meier survival curve of relapse-free survival between the high and low expression group in *MYH11*. Cohort is divided at median of gene expression.

HR (95% CI) - 0.79 (0.56 - 1.12), p-value > 0.05.