Identification and Characterization of Metastasis-Associated Individualized Gene Expression Signature in Osteosarcoma

Guofeng Zhang (gfzhangknu@gmail.com)
Yantai Affiliated Hospital of Binzhou medical University

Yonglin Zhu
Yantai Affiliated Hospital of Binzhou medical University

Chengzhen Jin
Yantai Affiliated Hospital of Binzhou medical University

Baisui Zhou
Yantai Affiliated Hospital of Binzhou medical University

Shanrun Zhao
Yantai Affiliated Hospital of Binzhou medical University

Chuanbo Li
Yantai Affiliated Hospital of Binzhou medical University

Fangang Fu
Yantai Affiliated Hospital of Binzhou medical University

Shanhui Li
Yantai Affiliated Hospital of Binzhou medical University

Primary research

Keywords: metastasis, gene expression, survival, biomarker, osteosarcoma

Posted Date: June 25th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-36107/v1

Abstract

Background: Osteosarcoma (OS) patients with surgical resection still relapse with poor prognosis due to the inability to detect distant metastasis. It's essential to identify metastasis-related biomarkers for OS.

Methods: Here, a computational pipeline follow relative expression orderings (REOs) using gene expression was constructed in metastases and non-metastases OS patients.

Results: 138 metastasis-associated gene pair signatures (MGPSs) were identified follow two independent datasets. A metastases-specific co-expressed MGPS network was constructed for extracting biomarker for clinical application. MGPS such as MYL5 and RPL27A showed strong positive correlation (Cor =0.75, P <0.001). There were thirteen prognostic MGPSs in above network. These prognostic MGPSs could become as a specific classifier to distinguish metastases and non-metastases OS patients. MGPSs were associated with cancer metastasis-related functions. Drug and MGPS network could provide some drug candidates for treatment of OS.

Conclusions: Collectively, the roles of the MGPSs in OS were elucidated, which could be beneficial for understanding OS pathogenesis and treatment.

Introduction

The American Cancer Society reported that the incidence of childhood and adolescent cancers is increasing rapidly. Among these, osteosarcoma (OS) is one of the cancer with highest mortality rate in the pediatric population [1]. In the last 35 years, there was no significant improvement in the five-year survival rate if metastases are present during diagnosis [2, 3]. Similar to the observation in other cancers, metastasis is a major factor for most death (75%) of OS patients [4]. For non-metastatic OS patients, tumor resection together with chemotherapy is a successful treatment plan. However, planning appropriate treatment for metastatic OS is still a great challenge for researchers and clinicians.

Cancer metastasis include two process: (1) cancer cells disseminate from primary tumors to whole body. (2) establish secondary lesions in distant organs [5]. Identifying metastases-associated biomarker was a key insight of diagnosis and treatment for OS. Recent years, some attempt has been contributed to identify metastases-associated biomarkers using gene expression profiles for many kinds of cancers [6–8]. However, the number of metastases-associated biomarkers which could be applicated for clinic is very small. Usually, most of the suggested metastases-associated biomarkers, classify patients into different risk groups by differential expression or clustering. The thresholds of differential expression which generated from training data set couldn't be directly applied to another data set. One of major reason is the gene expression level is sensitive based on diverse batch effect and platform of microarray. Furthermore, this means that the generated data of all the samples detected together should be normalized when only deal with a single sample. In other words, risk composition of other samples which are normalized together determines the risk prediction of a single sample [9].

One of insensitive approach for systematic biases of microarray is within-sample relative expression orderings (REOs). REOs are invariant and robust for monotonic data normalization and against differences within biological individuals [9, 10]. In past years, REOs have been proposed to identify prognostic signatures [11, 12]. However, the performance of REOs in predicting metastases-associated signatures for OS has been not reported. Therefore, it is worth trying to identify robust metastases-associated biomarkers for clinical application in OS based on the rank-based approach.

In present study, a computational pipeline based on REOs using gene expression profiles was constructed in metastases and non-metastases OS patients. Follow two independent datasets, 138 common metastasis-associated gene pair signature (MGPS) were identified in OS. A metastases-specific co-expressed MGPS network and its topological characteristics was constructed and analyzed. Some of MGPS such as MYL5 and RPL27A showed strong positive correlation (Cor = 0.75, P < 0.001) in metastatic OS patients. Thirteen prognostic MGPSs in co-expressed MGPS network were discovered. These prognostic MGPSs could better classify metastases and non-metastases OS patients. Functional analyses suggested that MGPSs were related with regulation of cancer and cancer-related pathways. Drug and MGPS network could provide some drug candidates for treatment of OS. Collectively, this study could provide assistance for diagnosis and treatment of OS.

Materials And Method

Clinical and gene expression profile datasets of OS

There were two main independent datasets in present study. First dataset including gene expression and related clinical information of 92 OS patients containing 23 and 69 metastatic and non-metastatic patients were obtained from the Therapeutically Applicable Research To Generate Effective Treatments (TARGET, https://ocg.cancer.gov/programs/target) data portal, which included 17070 mRNAs. TARGET program provides a comprehensive genomic landscape to explore molecular characteristics of childhood cancers. Providing a novel guide for developing effective therapeutic plans based on generated data is the major goal of TARGET program. The case selection criteria and sample details could be obtained at https://ocg.cancer.gov/programs/target/projects/osteosarcoma. Then, the second gene expression profile was obtained from gene expression omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE33382). Dataset GSE33382 gene expression profile included 23 and 69 metastatic and non-metastatic patients. The detailed information of patients could be found in previous study [13,14]. Corresponding platform files (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL10295) was used to map the gene Ids to genes. The averaged expression values were calculated and applied when multiple probes corresponding to a same gene for each sample. The probes which couldn't match any gene or matched not only a gene would be removed.

Identification of MGPS in OS

First, some metastases-related genes were obtained from CancerSEA (http://biocc.hrbmu.edu.cn/CancerSEA/), which is a database describes the functions of cancer cells [15]. We got 245 metastases-related genes and they would be used to follow analysis. The gene pair must have at least one metastases-related gene. Then, we built a 0-1 matrix for a pair of genes (G_m and G_n) in metastases and non-metastases OS patients. Lastly, if the frequency of metastases OS patients accord with a particular REO pattern was measured by fisher's exact test compared with non-metastases controls. The particular REO pattern was expression $G_m > G_n$ or $G_m < G_n$. The significant gene pairs evaluated with $P < 0.05$ were considered as MGPSs.

Construction and analysis of metastases-specific co-expressed MGPS network and its topological features

In order to construct a metastases-specific co-expressed MGPS network, pearson's correlation coefficients (PCCs) are calculated for each MGPS in metastases and non-metastases patients, respectively. The MGPSs were considered as metastases-specific co-expressed MGPSs, if absolute values of difference between PCCs in metastases and non-metastases patients were bigger than 0.2. Then a metastases-specific co-expressed MGPS network was constructed using Cytoscape 3.3.0 (http://www.cytoscape.org/). The degree analysis of the network was also using Cytoscape 3.3.0.

Survival analysis of MGPSs for OS patients

In order to evaluate the performance about the MGPS for prognosis in OS, we performed survival analysis for genes in each MGPS. Follow the median value of expression level for each gene, the OS patients were divided into two risk groups. And then, Kaplan-Meier (K-M) survival analysis was used for the two groups. $P < 0.05$ was consider as prognostic gene for OS.

Classification power of the prognostic MGPS in OS

Consensus clustering approach was used to classify metastases and non-metastases OS patients based on expression data of genes [16]. A R package named ConsensusClusterPlus (https://www.r-project.org/) was performed to this process. Best category number was select when the areas of Cumulative distribution function (CDF) curves were smallest. Chi-square test was applied to evaluate if metastases and non-metastases OS patients could be classified using this method ($P<0.01$).

Functional and drug enrichment analysis for MGPSs in OS

Online Enrichr (http://amp.pharm.mssm.edu/Enrichr/) tool was applied with default parameters to functional enrichment analysis for genes in MGPSs [17]. The enriched GO terms ($P<0.01$) and KEGG pathways ($P<0.05$) were extracted and considered as MGPS-associated functions. The gene-drug interaction data are download from DrugBank (https://www.drugbank.ca/) [18]. Then a drug-related MGPS network was constructed and analyzed to identify drug repurposing candidates for OS.

Results

Extract GMPSs in OS for metastases and non-metastases patients

We built an integrated and computational pipeline based on REO approach using gene expression profiles (Fig. 1A). First, 245 metastases-related genes were identified and they would be used to follow analysis. Second, a 0–1 matrix for a pair of genes (G_m and G_n) in metastases and non-metastases OS patients according to expression pattern of the gene pair. Then, if the frequency of metastases OS patients accord with a particular REO pattern was measured by fisher's exact test compared with non-metastases controls. The particular REO pattern was expression Gm > Gn or Gm < Gn. Lastly, MGPSs were identified if the P values of Fisher's exact test were smaller than 0.05. We performed this pipeline based on dataset from TARGET. The distribution of P values of MGPSs with fluctuation is between 0.01 and 0.05 (Fig. 1B). We divide all P values into five regions: 0 to 0.01, 0.01 to 0.02, 0.02 to 0.03, 0.03 to 0.04 and 0.04 to 0.05. The amount in each region had no great difference (Fig. 1C). In order to ensure the accuracy and stability of MGPSs for OS, another dataset GSE33382 from GEO was used for identifying MGPSs. The P values of GSE33382 showed similar distribution to TARGET dataset (Fig. 1D, 1E). Lastly, we got 138 common MGPSs in both dataset TARGET and GSE33382 for OS (Fig. 1F).

Some MGPSs showed strong correlation in metastases-specific co-expressed network for OS

The correlation values of MGPSs would showed great difference between metastases and non-metastases OS patients. According to the inference, a metastases-specific co-expressed network was constructed (Fig. 2A). The network contained 71 nodes and 57 edges. More than half MGPSs showed great difference (absolute values of difference for PCCs > 0.25) between metastases and non-metastases OS patients. There were 50.88% and 5.26% MGPSs were both positive and negative correlated in metastases and non-metastases OS patients (Fig. 2B). The correlated direction of 43.86% MGPSs were reversed in metastases and non-metastases OS patients. The global network showed scale-free distribution (R square = 0.75) which is a specific topological feature of transcriptional regulatory network (Fig. 2C). We found metastases-related gene RPL27A had highest degree in this metastases-specific co-expressed network (degree = 19). Most of the interacted strength between RPL27A and other genes had great difference in metastases and non-metastases OS patients. For example, gene RPL27A and MYL5 had strong positive correlations (Cor = 0.75, P < 0.001) in metastatic OS patients (Fig. 2D). However, gene RPL27A and MYL5 were not correlated in non-metastatic OS patients (Cor = 0.02, P = 0.86, Fig. 2E). The results indicated that these MGPSs maybe play key and specific roles in metastases process for OS patients.

Some genes in MGPSs were associated with prognosis for OS patients

As we known, time to metastases and number of lesions are the most important prognostic factors for OS patients [19]. Thus it's essential to consider the prognostic roles of MGPSs in OS patients. 13 genes

were associated with survival in above metastases-specific co-expressed network for OS (Fig. 3A). For example, prognosis-related MGPSs CYBB and P4HA1 showed strong negative correlations in metastases OS patients (Cor= -0.35, Fig. 3B). Prognosis-related MGPSs CYBB and P4HA1 also showed strong negative correlations in metastases OS patients (Cor= -0.25, Fig. 3C). Gene P4HA1, CYBB and CD53 were all related with survival (Fig. 3D, E, F). In addition, samples with high expression for these genes usually showed worse survival. All the results indicated that some genes in MGPSs were associated with prognosis for OS patients.

Prognosis-related genes in MGPSs could classify metastases and non-metastases OS patients

We constructed a prognosis-related MGPS network which extracted from metastases-specific co-expressed network. This network contained prognosis-related genes and their direct neighbored genes in metastases-specific co-expressed network (Fig. 4A). There were 25 nodes (13 prognosis-related genes) and 18 edges in this prognosis-related MGPS network. We found three MGPSs including P4HA1-CYBB, P4HA1-EVI2B and P4HA1-CD53 were all associated with survival. In order to validate if these prognosis genes in MGPSs could become biomarkers for metastases of OS patients, we classify metastases and non-metastases OS patients using gene expression profile follow a consensus clustering method. The prognosis-related genes in MGPSs could distinguish all samples into diverse groups. Final number of group was 4 based on area under CDF curve plot (Fig. 4B **and C**). Each OS patients group had a consensus expression pattern and could distinguish clearly (Fig. 4D). Most OS patients could be classified accurately (Chi-square test, $P < 0.001$) and especially the last group (C4) were all matched with the non-metastases OS patients (Fig. 4E). Collectively, all the above results indicated that the expression of prognosis-related genes in MGPSs could be served as specific biomarkers to distinguish metastases and non-metastases OS patients.

Functional characterizations and drug repurposing candidates of MGPSs in OS

Functional enrichment analyses were performed to characterize the functions of MGPSs in OS. For GO enrichment analyses, these MGPSs were enrichment in some critical biological functions such as regulation of microtubule polymerization, negative regulation of autophagy and negative regulation of smooth muscle cell migration (Fig. 5A). Dynamics of acting filament cooperated with microtubules could drive cell motility process. Many studies revealed that microtubule dynamics were necessary for promoting epithelial-mesenchymal transition (EMT) [20–22]. Expansion of dormant tumor cells into metastases or anti-tumor inflammatory responses would be restricted and promoted by autophagy. On the contrary, metastasis would be promoted by self-eating based on strengthening fitness of tumor cell environmental stresses response including anoikis during metastatic progression [23]. For KEGG pathway enrichment analyses, these MGPSs were enrichment in some key pathways associated with cancer or metastasis such as Natural killer cell mediated cytotoxicity, B cell receptor signaling pathway and MAPK signaling pathway (Fig. 5B). Natural killer cells participate to immune response against metastasis [24]. These functional enrichment analysis showed these GMPSs were associated with cancer metastasis.

We also constructed a drug-related MGPS network to explore drug repurposing candidates for OS (Fig. 5C). We found Proline (DB00172) was a drug repurposing candidate and its target gene were P4HA1 and PYCR2. Proline is one of the twenty amino acids in organism, which is a component of protein. Normal functions of jonts and tendons and maintain and strengthen all reply on proline at a great extent [25]. Azacytidine is clinically used to treat myelodysplastic syndrome, a group of heterogeneous bone marrow stem cell diseases [26]. In our analysis, Azacitidine was considered as a drug candidate for OS. Azacitidine is a cytidine nucleoside analogue, which has the clinical activity of myelodysplastic syndrome and the potential activity of solid tumor.

Discussion

In this study, we suggested that the MGPSs using REO approach follow expression level of genes play specific roles in metastasis process of OS patients. These MGPSs could be as metastasis-related biomarkers for OS. Differential expression analysis for gene expression profiles between metastases and non-metastases samples was considered as a common approach. The adoption of normalization for other samples contribute to differential expression of patients and it would generate a lot of uncertainty for risk classification of patient. Especially, the uncertainty would increase if the sample size was small and couldn't represent all the disease patients [27]. Specially, qualitative REOs-based MGPSs would identify more accurated personalized information of individual patient than traditional differential expression in clinical application.

A hypothesis that the OS patients with poor prognosis might harbor metastases was advanced. Thus we performed survival analysis and identified 13 prognosis-related genes in MGPSs. The problems of insufficient power and too complex to evaluate metastasis would be present based on all the MGPSs in OS patients. According the clinical needs, we proposed prognosis-related MGPSs based a strict voting criterion for survival analysis and proved that these prognosis-related MGPSs performed better than all the MGPSs. The results indicated that prognosis-related MGPSs maybe more suitable for clinical application of predicting metastasis. In order to explore if these prognosis-related genes in MGPSs could be specific biomarkers of metastasis in OS patients, we used them to classify metastases and non-metastases patients based on gene expression. The 13 prognosis-related genes could classify metastases and non-metastases patients ($P < 0.001$).

In prognosis-related GMPSs network, gene P4HA1 was a key node with highest degree. The previous study has reported that P4HA1 was the active catalytic component of prolyl 4-hydroxylase in cancers [28], glioma [29], prostate cancer [30] and pancreatic cancer [31], P4HA1 can promote chemoresistance, tumor growth and metastasis. The prognosis-related gene CD53 could form a MGPS with P4HA1 in OS patients. CD53 is essential for CD2 signal transduction, growth regulation and cell survival of cancer [32]. P4HA1 was also a drug target in drug-related MGPS network. The roles of P4HA1 in OS should be explored and validated.

Conclusion

In this study, some MGPSs were identified and characterized for OS patients based on gene expression profiles. Correlations of some MGPSs showed obvious difference between metastases and non-metastases samples. Prognosis-related genes in metastases-specific co-expressed network could become as specific biomarkers to classify metastases and non-metastases OS patients. The functional analysis showed the association between MGPSs and cancer metastasis in OS patients. Collectively, our study provides novel insights into the mechanisms underlying the roles of MGPSs in metastasis process for OS.

Declarations

Funding

The work is supported by the research initiation fund of Yantai Affiliated Hospital of Binzhou Medical College (BY2019KYQD50).

Disclosure of interest

The authors declare that they have no competing interest.

Acknowledgments

Not applicable.

Availability of data and material

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Authors' contributions

ZGF conceived and designed the experiments, ZSR, LCB, FFG and LSH analysed the data, and ZYL, JCZ and ZBS wrote the manuscript.

References

1. Gonzalez-Fernandez Y, Brown HK, Patino-Garcia A, Heymann D, Blanco-Prieto MJ. Oral administration of edelfosine encapsulated lipid nanoparticles causes regression of lung metastases in pre-clinical models of osteosarcoma. Cancer Lett. 2018;430:193–200.
2. Negri GL, Grande BM, Delaidelli A, El-Naggar A, Cochrane D, Lau CC, Triche TJ, Moore RA, Jones SJ, Montpetit A, Marra MA, Malkin D, Morin RD, Sorensen PH. Integrative genomic analysis of matched primary and metastatic pediatric osteosarcoma. J Pathol. 2019;249:319–31.
3. Isakoff MS, Bielack SS, Meltzer P, Gorlick R. Osteosarcoma: Current Treatment and a Collaborative Pathway to Success. J Clin Oncol. 2015;33:3029–35.
4. Huang YM, Hou CH, Hou SM, Yang RS. The metastasectomy and timing of pulmonary metastases on the outcome of osteosarcoma patients. Clin Med Oncol. 2009;3:99–105.
5. Jin K, Gao W, Lu Y, Lan H, Teng L, Cao F. Mechanisms regulating colorectal cancer cell metastasis into liver (Review). Oncol Lett. 2012;3:11–5.
6. Liang J, Chen M, Hughes D, Chumanevich AA, Altilia S, Kaza V, Lim CU, Kiaris H, Mythreye K, Pena MM, Broude EV, Roninson IB. CDK8 Selectively Promotes the Growth of Colon Cancer Metastases in the Liver by Regulating Gene Expression of TIMP3 and Matrix Metalloproteinases. Cancer Res. 2018;78:6594–606.
7. Hon CC, Ramilowski JA, Harshbarger J, Bertin N, Rackham OJ, Gough J, Denisenko E, Schmeier S, Poulsen TM, Severin J, Lizio M, Kawaji H, Kasukawa T, Itoh M, Burroughs AM, Noma S, Djebali S, Alam T, Medvedeva YA, Testa AC, Lipovich L, Yip CW, Abugessaisa I, Mendez M, Hasegawa A, Tang D, Lassmann T, Heutink P, Babina M, Wells CA, Kojima S, Nakamura Y, Suzuki H, Daub CO. **de** Hoon MJ, Arner E, Hayashizaki Y, Carninci P, Forrest AR. An atlas of human long non-coding RNAs with accurate 5' ends. Nature. 2017; 543: 199–204.
8. Tobin NP, Harrell JC, Lovrot J, Egyhazi Brage S, Frostvik Stolt M, Carlsson L, Einbeigi Z, Linderholm B, Loman N, Malmberg M, Walz T, Ferno M, Perou CM, Bergh J, Hatschek T, Lindstrom LS. Molecular subtype and tumor characteristics of breast cancer metastases as assessed by gene expression significantly influence patient post-relapse survival. Ann Oncol. 2015;26:81–8.
9. Wang H, Sun Q, Zhao W, Qi L, Gu Y, Li P, Zhang M, Li Y, Liu SL, Guo Z. Individual-level analysis of differential expression of genes and pathways for personalized medicine. Bioinformatics. 2015;31:62–8.
10. Geman D, d'Avignon C, Naiman DQ, Winslow RL. Classifying gene expression profiles from pairwise mRNA comparisons. Stat Appl Genet Mol Biol. 2004;3:Article19.
11. Zhou X, Li B, Zhang Y, Gu Y, Chen B, Shi T, Ao L, Li P, Li S, Liu C, Guo Z. A relative ordering-based predictor for tamoxifen-treated estrogen receptor-positive breast cancer patients: multi-laboratory

cohort validation. Breast Cancer Res Treat. 2013;142:505–14.

12. Xu L, Tan AC, Winslow RL, Geman D. Merging microarray data from separate breast cancer studies provides a robust prognostic test. BMC Bioinformatics. 2008;9:125.
13. Kuijjer ML. **van den** Akker BE, Hilhorst R, Mommersteeg M, Buddingh EP, Serra M, Burger H, **Hogendoorn PC, Cleton-Jansen AM.** Kinome and mRNA expression profiling of high-grade osteosarcoma cell lines implies Akt signaling as possible target for therapy. BMC Med Genomics. 2014; 7: 4.
14. Kuijjer ML, Peterse EF. **van den** Akker BE, Briaire-de Bruijn IH, Serra M, Meza-Zepeda LA, Myklebost O, Hassan AB, **Hogendoorn PC, Cleton-Jansen AM.** IR/IGF1R signaling as potential target for treatment of high-grade osteosarcoma. BMC Cancer. 2013; 13: 245.
15. Yuan H, Yan M, Zhang G, Liu W, Deng C, Liao G, Xu L, Luo T, Yan H, Long Z, Shi A, Zhao T, Xiao Y, Li X. CancerSEA: a cancer single-cell state atlas. Nucleic Acids Res. 2019;47:D900-D8.
16. Wilkerson MD. **Hayes DN.** ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26:1572–3.
17. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG. **Monteiro CD, Gundersen GW, Ma'ayan A.** Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic acids research. 2016;44:W90-7.
18. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018;46:D1074-D82.
19. Aljubran AH, Griffin A, Pintilie M, Blackstein M. Osteosarcoma in adolescents and adults: survival analysis with and without lung metastases. Ann Oncol. 2009;20:1136–41.
20. Li N, Jiang P, Du W, Wu Z, Li C, Qiao M, Yang X, Wu M. Siva1 suppresses epithelial-mesenchymal transition and metastasis of tumor cells by inhibiting stathmin and stabilizing microtubules. Proc Natl Acad Sci U S A. 2011;108:12851–6.
21. Nakaya Y, Sukowati EW, Wu Y, Sheng G. RhoA and microtubule dynamics control cell-basement membrane interaction in EMT during gastrulation. Nat Cell Biol. 2008;10:765–75.
22. Rodriguez OC, Schaefer AW, Mandato CA, Forscher P, Bement WM, Waterman-Storer CM. Conserved microtubule-actin interactions in cell movement and morphogenesis. Nat Cell Biol. 2003;5:599–609.
23. Kenific CM, Thorburn A, Debnath J. Autophagy and metastasis: another double-edged sword. Curr Opin Cell Biol. 2010;22:241–5.
24. Langers I, Renoux VM, Thiry M, Delvenne P, Jacobs N. Natural killer cells: role in local tumor growth and metastasis. Biologics. 2012;6:73–82.
25. Phang JM, Liu W, Hancock CN, Fischer JW. Proline metabolism and cancer: emerging links to glutamine and collagen. Curr Opin Clin Nutr Metab Care. 2015;18:71–7.

26. Nguyen AN, Hollenbach PW, Richard N, Luna-Moran A, Brady H, Heise C, MacBeth KJ. Azacitidine and decitabine have different mechanisms of action in non-small cell lung cancer cell lines. Lung Cancer. 2010;1:119–40.
27. Zhang M, Yao C, Guo Z, Zou J, Zhang L, Xiao H, Wang D, Yang D, Gong X, Zhu J, Li Y, Li X. Apparently low reproducibility of true differential expression discoveries in microarray studies. Bioinformatics. 2008;24:2057–63.
28. Xiong G, Stewart RL, Chen J, Gao T, Scott TL, Samayoa LM. O'Connor K, Lane AN, Xu R. Collagen prolyl 4-hydroxylase 1 is essential for HIF-1alpha stabilization and TNBC chemoresistance. *Nat Commun*. 2018; 9: 4456.
29. Zhou Y, Jin G, Mi R, Zhang J, Xu H, Cheng S, Zhang Y, Song W, Liu F. Knockdown of P4HA1 inhibits neovascularization via targeting glioma stem cell-endothelial cell transdifferentiation and disrupting vascular basement membrane. Oncotarget. 2017;8:35877–89.
30. Chakravarthi BV, Pathi SS, Goswami MT, Cieslik M, Zheng H, Nallasivam S, Arekapudi SR, Jing X, Siddiqui J, Athanikar J, Carskadon SL, Lonigro RJ, Kunju LP, Chinnaiyan AM, Palanisamy N, Varambally S. The miR-124-prolyl hydroxylase P4HA1-MMP1 axis plays a critical role in prostate cancer progression. Oncotarget. 2014;5:6654–69.
31. Cao XP, Cao Y, Li WJ, Zhang HH, Zhu ZM. P4HA1/HIF1alpha feedback loop drives the glycolytic and malignant phenotypes of pancreatic cancer. Biochem Biophys Res Commun. 2019;516:606–12.
32. Yunta M, Lazo PA. Apoptosis protection and survival signal by the CD53 tetraspanin antigen. Oncogene. 2003;22:1219–24.

Figures

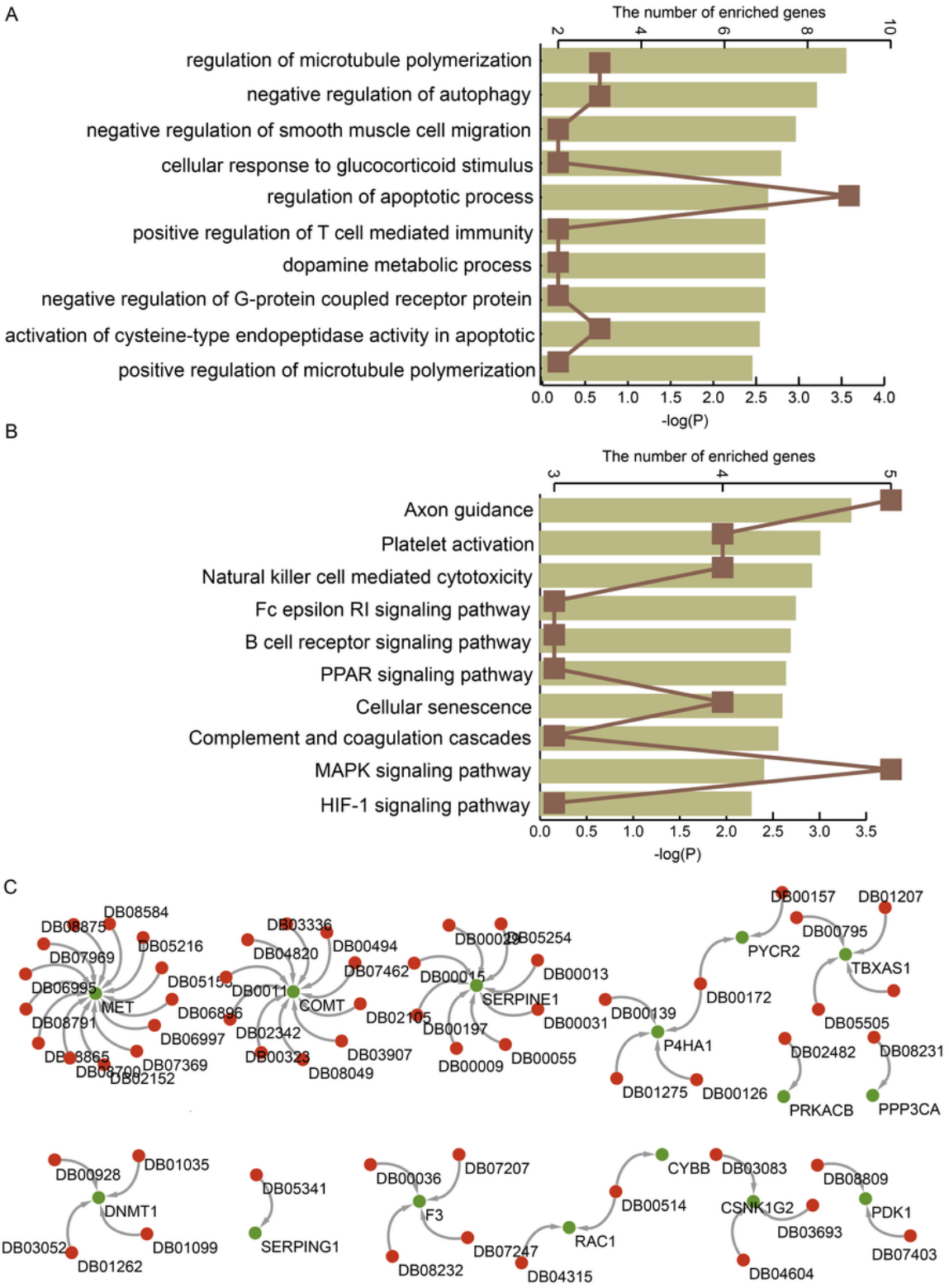

Figure 1

Functional characterizations and drug repurposing candidates of MGPSs in OS. (A) GO terms enriched for genes in MGPSs for OS and bar plots represent −log10(P). The dot line graph shows number of enriched genes in each GO term. (B) KEGG pathway enriched for genes in MGPSs for OS and bar plots represent −log10(P). The dot line graph shows number of enriched genes in each KEGG pathways. (C) The drug and MGPS network. Red and green represent drug and gene, respectively.

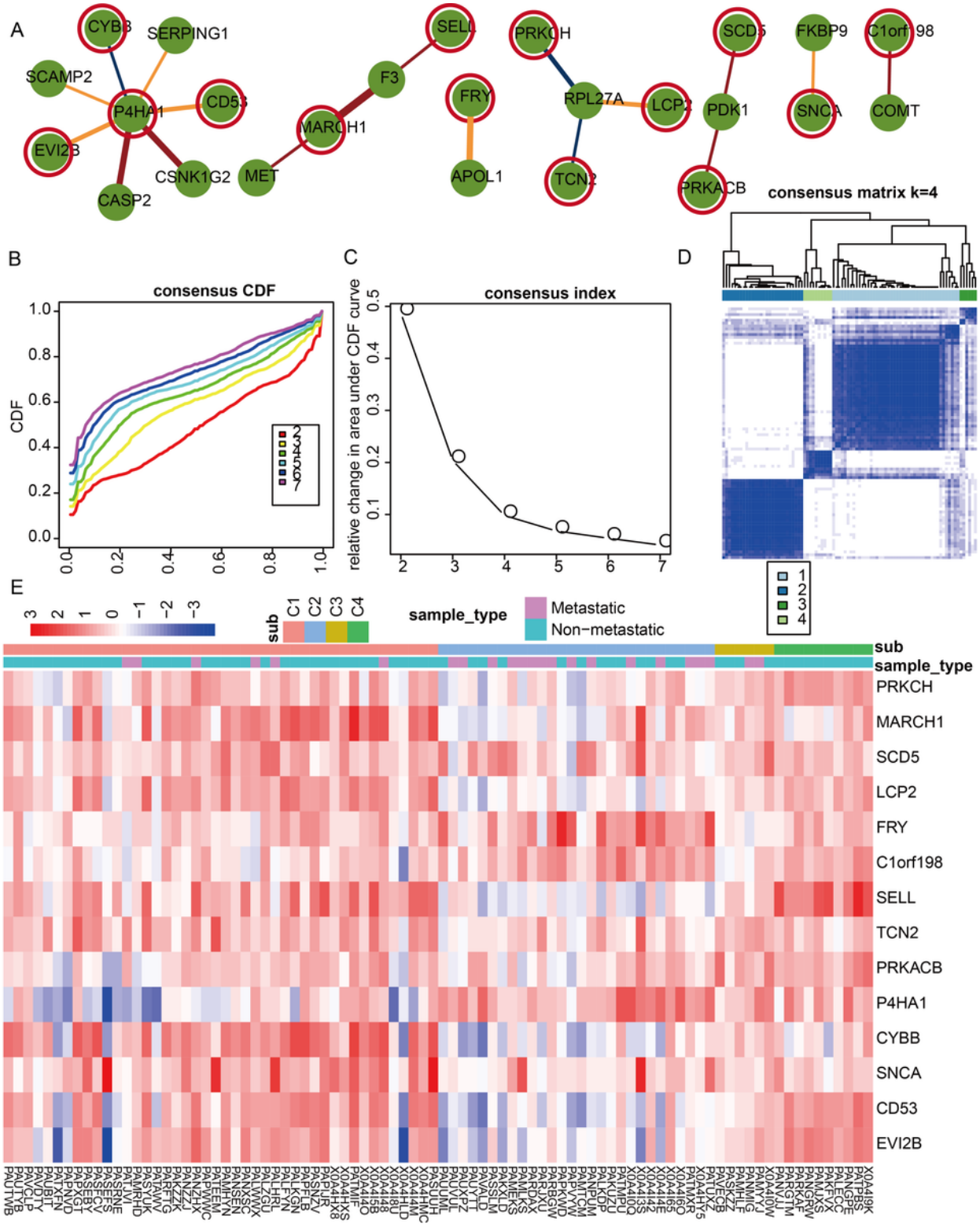

Figure 2

Prognosis-related genes in MGPSs could classify metastases and non-metastases OS patients. (A) Prognosis-related MGPS network for OS patients. The red circle represents prognosis gene. (B) Cumulative distribution function plot of the consensus index. (C) Relative change in area under CDF curve of different group number. (D) Consensus cluster heatmap of OS patients. (E) The gene expression

heatmap, group type classified by consensus cluster is represented by sub label, the metastasis status of the OS patients is represented by sample type.

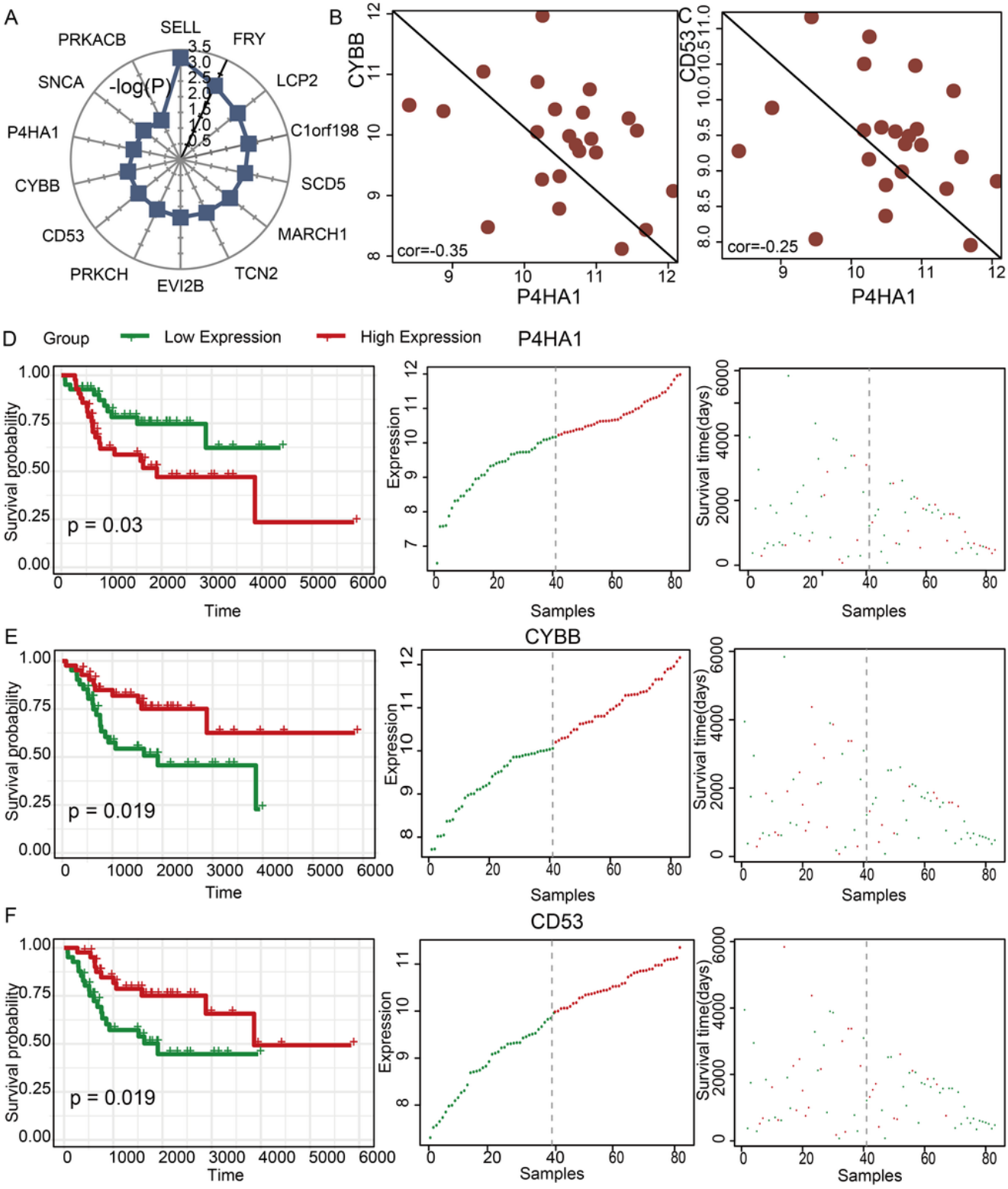

Figure 3

Some MGPSs in metastases-specific co-expressed MGPS network were associated with survival in OS patients. (A) The radar chart shows the P values of 13 prognosis-related MGPSs. (B) The point plot shows the expression correlation of gene P4HA1 and CYBB in metastases samples. (C) The point plot

shows the expression correlation of gene P4HA1 and CD53 in metastases samples. (D-F) The overall survival of two OS groups with high and low expression are showed in K-M curves. Two-sided log-rank test is used to evaluate the difference between two K-M curves. The risk score distribution of the genes based on gene expression. The patient survival status of the genes.

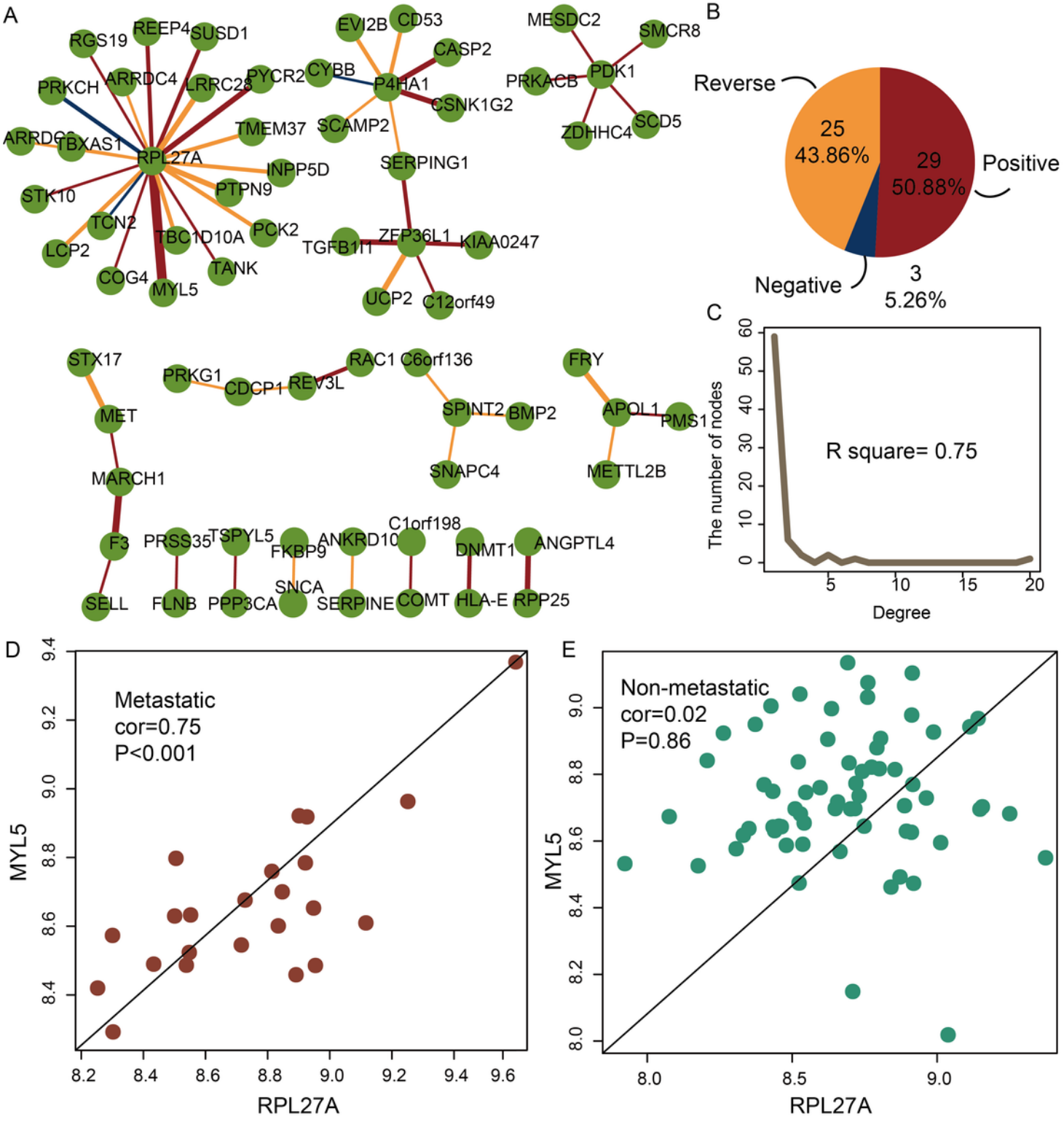

Figure 4

Construction of metastases-specific co-expressed MGPS network for OS. (A) A metastases-specific co-expressed MGPS network. Positive and negative correlations are represented by red and blue, respectively. The thicker edges represent bigger difference of correlations between metastases and non-metastases OS patients. (B) The pie chart shows percents for reverse, negative and positive interactions. (C) The plot shows degree distribution of metastases-specific co-expressed MGPS network. (D) The point plot shows the expression correlation of gene RPL27A and MYL5 in metastases samples. (E) The point plot shows the expression correlation of gene RPL27A and MYL5 in non-metastases samples.

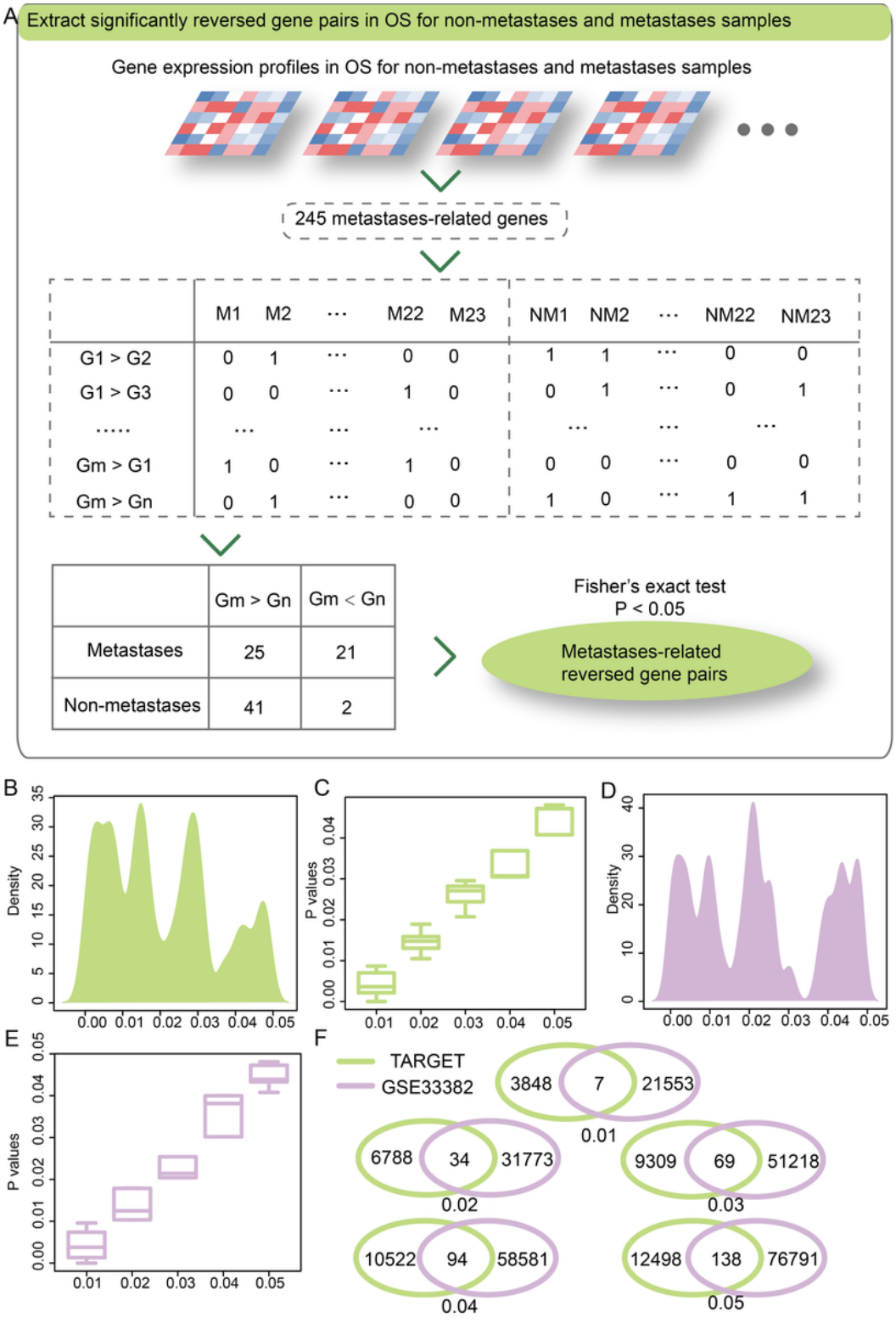

Figure 5

Identification of MGPSs based on gene expression profile in OS patients. (A) The workflow of identifying MGPSs based on gene expression profile in metastases and non-metastases samples. (B) The density distribution curve shows distribution of P values of MGPSs based on TARGET dataset. (C) The box plots show P values of MGPSs based on TARGET dataset in diverse regions. (D) The density distribution curve shows distribution of P values of MGPSs based on GSE33382 dataset. (E) The box plots show P values of MGPSs based on GSE33382 dataset in diverse regions. (F) The venn plots show intersection between TARGET and GSE33382 datasets follow diverse regions of P values.