Am J Cancer Res 2023;13(4):1582-1593
www.ajcr.us /ISSN:2156-6976/ajcr0149574

Brief Communication

Identification and validation of *MSMB* as a critical gene for prostate cancer development in obese people

Xi Chen[1*], Haopeng Li[1*], Boya Liu[2*], Xin'an Wang[1], Weidong Zhou[1], Gang Wu[1], Chengdang Xu[1]

[1]Department of Urology, Tongji Hospital, School of Medicine, Tongji University, Shanghai, The People's Republic of China; [2]Rheumatology and Immunology Department, Hebei Province Cangzhou Hospital of Integrated Traditional Chinese and Western Medicine, Cangzhou, Hebei, The People's Republic of China. [*]Equal contributors and co-first authors.

Received February 8, 2023; Accepted March 9, 2023; Epub April 15, 2023; Published April 30, 2023

Abstract: Prostate cancer (PCA) is one of the most common types of cancer and can seriously endanger the health of older men. Obesity is prevalent all around the world and triggered by lots of factors such as diet, environment and fat metabolism disorder can cause many neoplasms, including PCA. Evidence suggests that genetic changes increase the risk of PCA and obesity. However, the specific obesity-related genes leading to PCA are unknown. Obesity-related genes associated with PCA were identified and analyzed though three public electronic databases: Gene Expression Omnibus, The Cancer Genome Atlas, and Chinese Prostate Cancer Genome and Epigenome Atlas. The effect of obesity-related genes in PCA were analyzed using clinical data from different databases, while associations with immune cells were determined by TIMER web tool. The expression and function of obesity-related genes were verified using clinical samples from obese patients with PCA and PCA cells. We found that four genes, *MSMB*, *BMP5*, *THBS4*, and *POPDC3*, may lead to PCA occurrence in patients with obesity. In Gene Expression Omnibus database, *MSMB* and *BMP5* were downregulated, while *THBS4* and *POPDC3* were upregulated. This trend was mainly preserved in the other electronic databases. We also discovered *MSMB* and *THBS4* can affect PCA progression, and all these genes were risk factors for castration-resistant prostate cancer. Moreover, *MSMB* can impact disease-free survival status of patients with PCA. These obesity-related genes were also correlated with immune cells and immune cell infiltration in PCA. We further uncovered that *MSMB* was downregulated in clinical PCA and castration-resistant prostate cancer samples from patients with obesity and *MSMB* decreased PCA cells proliferation. These results indicate that *MSMB* is essential for PCA development in people with obesity and can be a biomarker for predicting PCA occurrence and progression in obese people.

Keywords: Obesity, prostate cancer, *MSMB*, cancer development

Introduction

In 2021, prostate cancer (PCA) was the first most common cancer diagnosis and the second most common cause of cancer-related mortality in American men [1]. Today, a similar trend is observed in Chinese men since increased age and lifestyle changes accelerate morbidity and mortality rates of PCA [2]. Thus, understanding the factors that promote PCA progression is paramount. Although we know that many factors contribute to the occurrence and development of PCA, including race, environment, smoking, and obesity [3], the reasons for these factors leading to PCA need more study.

The size of a population with obesity increases with a rapidly developing society [4]. Since 1975, excessive body weight has affected more than two-thirds of the American population, with more than one-third of adults and one-fifth of adolescents affected by obesity [5]. Obesity is driven by diverse factors, such as diet, environment, heredity and fat metabolism disorder [6]. It can also cause many other metabolic dysfunctions like hyperglycemia and dyslipidemia, raising the risk of numerous neoplasms [7]. In the case of PCA, obesity may affect its occurrence, development, drug sensitivity, and outcome [8]. Various underlying factors, including heredity and genetic alteration, may contributed to obese people suffered PCA [9, 10].

Although genetic factors are essential contributors to PCA from obesity [8], few obesity-related genes that cause PCA have been studied. We hypothesized that identifying genes associated with obesity and PCA would serve as a starting point for delineating the underlying mechanism that promotes PCA in persons with obesity.

In this study, we used Gene Expression Omnibus (GEO) database to explore the obesity-related genes in PCA. We found four obesity-related genes associated with PCA occurrence: microseminoprotein beta (*MSMB*), bone morphogenetic protein 5 (*BMP5*), thrombospondin 4 (*THBS4*), and popeye domain containing 3 (*POPDC3*). We verified their expression in patients with PCA using the GSE35988 dataset, The Cancer Genome Atlas (TCGA) and the Chinese Prostate Cancer Genome and Epigenome Atlas (CPGEA) databases. All these genes were decisive for PCA occurrence, while *MSMB* and *THBS4* were for PCA progression. Moreover, *MSMB* affects the disease-free survival (DFS) of patients with PCA. Finally, *MSMB* was downregulated in obese PCA and castration-resistant prostate cancer (CRPC) patients and decreased PCA cells proliferation. Together, our study suggests *MSMB* is a critical gene for PCA development in obese people and can be a biomarker predicting PCA progression in obese patients.

Materials and methods

Data getting and processing

The GSE79021 dataset containing data of PCA patients with obesity was retrieved from the GEO database (http://www.ncbi.nlm.nih.gov/geo/) and used to identify obesity-related genes in PCA. The patients were assigned into a group with obesity whose BMI > 25 which can be considered as overweight or the one without obesity [11]. Datasets GSE46602 and GSE35988 with microarray and clinical data of patients with PCA and healthy persons were also retrieved to determine PCA-associated genes. The microarray and clinical information of patients with PCA from TCGA (http://cancergenome.nih.gov/) and the CPGEA (http://www.cpgea.com) databases were downloaded. The raw data were processed as reported in previous study [12]. The genes with an adjusted *P*-value < 0.05 and $|\log_2 FC| > 1$ were considered significant. A Venn diagram was generated to recognize the genes associated with obesity and PCA.

Online web tools, pathways and immune infiltration analysis

Online web tools UALCAN (http://ualcan.path.uab.edu/), GEPIA (http://www.gepia.cancer-pku.cn/), and TIMER (https://cistrome.shinyapps.io/timer/) were used in analyzing the data form TCGA. DAVID (https://david.ncifcrf.gov/) was used to find the pathways of obesity-related genes enriched [13]. The effect of obesity-related genes on immune cell infiltration was estimated using TIMER.

Construction of risk prediction model

Logistic regression was performed to quantify the hazard ratios of gene changes leading to PCA. A forest plot was built for a graphical representation of the hazard ratios, and a nomogram was constructed to predict the risk value of the obesity-related genes in causing PCA. A forest map was utilized to show the hazard ratios more intuitively.

Survival analysis

Cox proportional hazards and least absolute shrinkage and selection operator (LASSO) regression models were used to estimate whether the obesity-related genes can affect the disease-free survival (DFS) of patients with PCA. Both models were created with R software. The LASSO risk score was calculated as follows: risk score = $\Sigma jn = 1$ Coef j × Xj, where Coef j indicates the LASSO-calculated coefficient, and Xj, mRNA expression of the obesity-related genes. The effectiveness of the LASSO model was assessed with Log (λ_i). The influence of these genes on DFS and overall survival (OS) of patients was validated by the GEPIA.

Clinical specimen collection

Clinical PCA and CRPC specimens consisting of paired cancerous and para-cancerous normal tissues were obtained from the Tongji Hospital, School of Medicine, Tongji University. The Ethics Committee of Tongji Hospital (SBKT-2021-220) approved the methods for sample collection. All patients donated tissues confirmed obesity with a BMI > 25. They were famil-

iar with the experiments in the study and gave written informed consent before sample collection.

Cell culture

Two human PCA cell lines were purchased from The Cell Bank of Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China): PC-3 and DU145. They were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium (No. R8758, Merck KGaA, Darmstadt, Germany) containing 10% fetal bovine serum (No. 10091, Gibco, Waltham, MA, USA) and humid environment at 37°C with 5% CO_2 and 95% air.

Cell transfection and cell proliferation

Cell transfection was performed with Lipofectamine 2000 (No. 11668019, Invitrogen). Overexpression plasmids carrying *MSMB* (oeMSMB) and plasmids with blank controls (oeControl) were purchased from Youze Biotechnology (Hubei, China). Cell proliferation ability was checked by Cell Counting Kit-8 (CCK-8) (Dojindo Laboratories, Japan). The experimental procedures were performed as described previously [12].

Total RNA extraction and RT-qPCR

Total RNA was extracted from tissues of patients with PCA and CRPC using the TRIzol reagent (No. T9424, Invitrogen, Carlsbad, CA, USA). It was reverse transcribed to cDNA using the Advantage RT-for-PCR Kit (No. 639505, Takara Bio Inc., Kusatsu, Japan). The cDNA was assessed with quantitative PCR with a TB Green Premix Ex Taq™ II kit (No. RR420A, Takara Bio Inc., Kusatsu, Japan) according to the manufacturer's instructions. Primer sequences of MSMB were as follows: forward, 5'-CTTTGCCACCTTCGTGACTTTATGC-3'; reverse, 5'-CTGGGAGCCCTGTGCCTACTAG-3'. GAPDH served as endogenous control and the primers of it were: forward, 5'-GGAGCGAGATCCCTCCAAAAT-3'; reverse, 5'-GGCTGTTGTCATACTTCTCATGG-3'. The $2^{-\Delta\Delta Ct}$ method was used to quantify mRNA expression.

Antibodies

The primary antibodies were rabbit monoclonal antibodies against MSMB (No. DF6720, Affinity company, Suzhou, China) and GAPDH (No. ab9485, Abcam plc, Cambridge, UK). The secondary antibody was HRP AffiniPure Goat Anti-Rabbit IgG (No. A0216, Beyotime Biotechnology, Inc., Shanghai, China).

Western blotting

Tissues and cells were lysed with RIPA lysis buffer, and the lysates were treated with Dual Color Protein Loading Buffer (No. NP0007, Thermo Fisher Scientific, Waltham, MA, USA). Proteins were resolved with 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membranes (No. 71078, Merck KGaA, Darmstadt, Germany). Blocking was performed using protein-free blocking solution (Thermo Fisher Scientific). Overnight incubation at 4°C followed using the primary antibodies against MSMB (1:500) and GAPDH (1:2000). The membranes were washed 3 times for 10 min in 1x TBST. Incubation with the secondary antibody (1:1000) was done at room temperature for 1 h. Finally, target proteins were detected using an X-ray film.

Statistical analysis

The gene expression matrix data were analyzed using R version 4.0.3 (Institute for Statistics and Mathematics, Vienna, Austria) (https://www.r-project.org). Comparisons between two groups were performed using the Wilcoxon test; between more than two, with the Kruskal-Wallis test. Hazard ratios, 95% confidence interval (95% CI), and *P* values were used as statistical metrics. Two-tailed $P < 0.05$ was inferred as statistically significant.

Results

Four differently expressed genes found in obese PCA patients

First, we tried to find the key genes which can lead to PCA occurrence in people with obesity. We discovered a data set: GSE79021 that entails data of PCA patients with normal weight and those with higher weight. We assigned these patients into a group with normal weight or the other one defined as obesity with a BMI > 25. A total of 28 genes exhibited differential expression between the two groups, and this difference is reflected in the volcano plot (**Figure 1A**). We also identified another data set GSE46602, with differently expressed genes

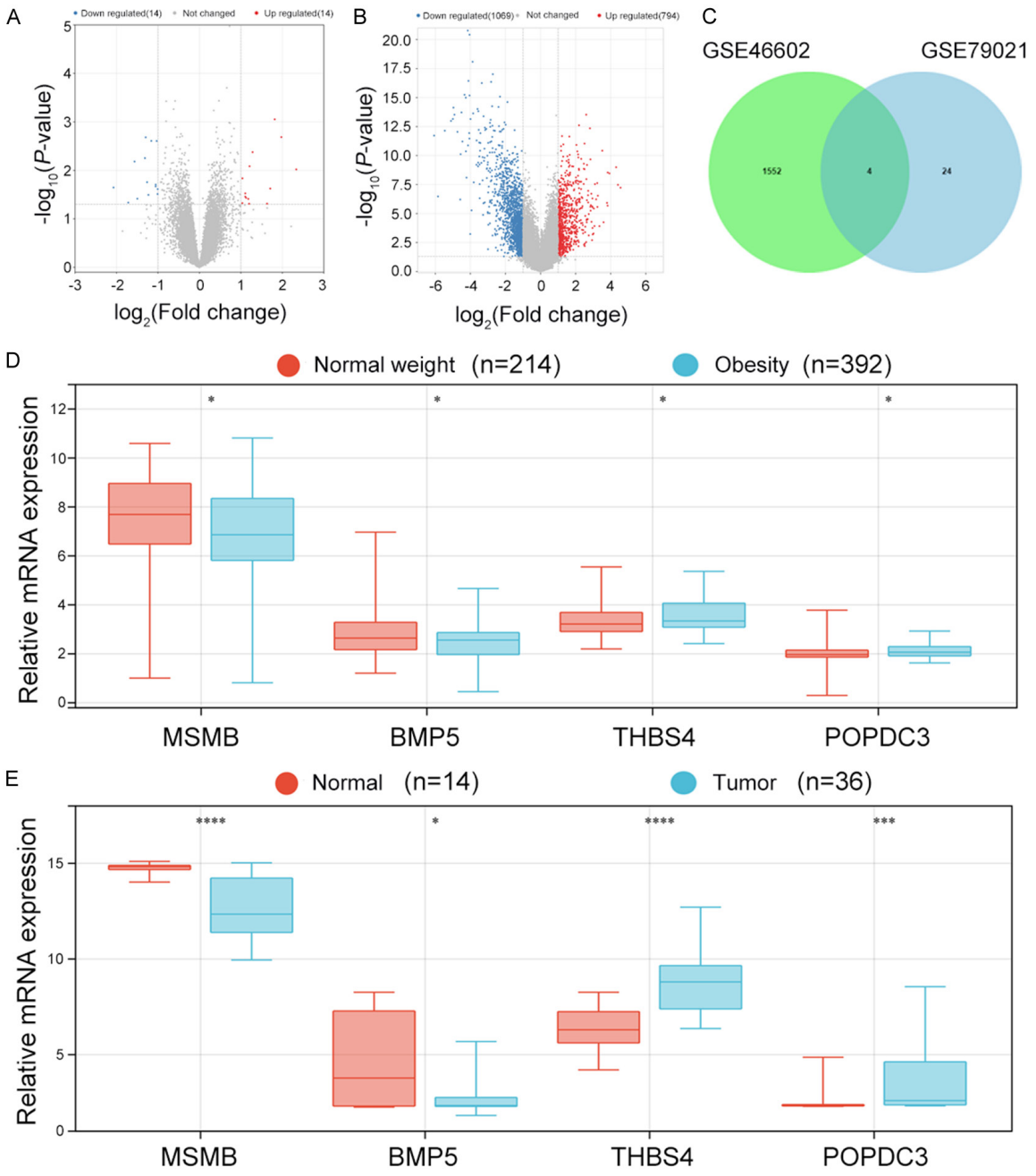


Figure 1. Four obesity-related genes associated with the occurrence of PCA. A. The differently expressed genes from GSE79021 shown in volcan map. B. The differently expressed genes from GSE46602 reflected in volcan map. C. Venn map reflected the common genes have different expression between GSE79021 and GSE46602. D. The expression of four obesity-related genes in PCA patients with normal and obesity weight from GSE79021 dataset. E. The expression of four obesity-related genes in PCA and normal samples from GSE46602 dataset. *represents $P < 0.05$, ***represents $P < 0.001$, ****represents $P < 0.0001$.

between PCA and normal prostate tissues (**Figure 1B**). By comparing the same differently expressed genes between GSE79021 and GSE46602 datasets, we uncovered four genes (*MSMB*, *BMP5*, *THBS4*, and *POPDC3*) maybe important for PCA occurrence in obese people (**Figure 1C**). Among them, *MSMB* and *BMP5* were downregulated, whereas *THBS4* and *POPDC3* were upregulated in obese PCA patients from GSE79021 (**Figure 1D**). Moreover, this expression pattern was also consistent in GSE46602, confirming the unique response of

the genes to PCA occurrence (**Figure 1E**). We also did a pathway enrichment analysis to infer the biological functions of the four genes and found the naba core matrisome pathway was among the most enriched for these genes (Supplementary Figure 1).

Obesity-related genes' expression in PCA form different databases

After finding obesity-related genes leading to PCA, we verified their expression in other publicly available datasets containing information from patients with PCA. We confirmed that *MSMB* and *BMP5* are downregulated, but *THBS4* and *POPDC3* are upregulated in tissues derived from patients with PCA in the TCGA database (**Figure 2A**). Since the TCGA database mainly consists of data from Western patient populations, we also assessed the expression of these genes in Chinese patient populations from CPGEA database. Among these obesity-related genes, *POPDC3* was unaffected by PCA in Chinese patients, while *THBS4* was upregulated as in the Western. Conversely, *MSMB* and *BMP5* were consistently downregulated in both patient populations (**Figure 2B**). We looked at another data set named GSE35988, with microarray data of PCA and normal prostate tissues. Intriguingly, *POPDC3* was downregulated in PCA tissues (**Figure 2C**). Because methylation levels of genes affect their function and are implicated in cancer initiation, we further examined whether the methylation levels of the obesity-related genes would change when PCA happened using data from the TCGA. Remarkably, methylation levels of these genes changed upon PCA occurrence (Supplementary Figure 2A-D). Methylation levels of *MSMB* and *BMP5* were upregulated (Supplementary Figure 2A, 2B), while those of *THBS4* and *POPDC3* were downregulated (Supplementary Figure 2C, 2D). These results indicate that the identified obesity-related genes are essential for PCA occurrence, likely instituted by changes in the methylation status of genes.

Risk prediction model of obesity-related genes in causing PCA

Since the expression of the discovered obesity-related genes changes with PCA occurrence, we guessed that these genes are risk factors for PCA. We performed a single-factor logistic regression using TCGA data and identified *THBS4* as a factor associated with a higher risk of PCA (Supplementary Figure 3A). However, the following multiple-factor logistic regression yielded none of the obesity-related genes were risk factors for PCA (Supplementary Figure 3B). Next, we constructed a nomogram to visualize our results. Indeed, it showed that *THBS4* has the highest risk for PCA (Supplementary Figure 3C). These results were supported with a calibration curve, which was calculated to validate the predictive ability of the nomogram (Supplementary Figure 3D).

Obesity-related genes affecting PCA progression

Next, we tried to find whether these obesity-related genes can affect PCA progression. As TNM (tumor, node, metastasis) staging classification is widely used to assess PCA severity [14], we analyzed the relationship between the expression of these genes and different categories of TNM staging. Because the M category (presence of distant metastasis) is rare among patients with PCA in TCGA, we did not include it in the analysis. Among the four obesity-related genes, *MSMB* influenced the primary tumor (T category) and regional lymph node involvement (N category) (Supplementary Figure 4A and 4E). By contrast, other genes showed no associations with T category except *THBS4* (Supplementary Figure 4B-D). Further, we found that *THBS4* can also affect the node-metastasis (N) tumor stage (Supplementary Figure 4G). However, other genes cannot affect the N tumor stage (Supplementary Figure 4F and 4H). We performed a similar analysis using the TNM staging information from Chinese patients with PCA using CPGEA data. In these patients, *MSMB* can influence T category, too (Supplementary Figure 5A), but other genes did not (Supplementary Figure 5B-D). In addition, none of the four genes can affect the N category (Supplementary Figure 5E-H). Finally, using data from GSE35988 data set, we found that the expression of all these four obesity-related genes changed when CRPC happened (Supplementary Figure 5I-L). These results indicated that some of the obesity-related genes can affect PCA progression.

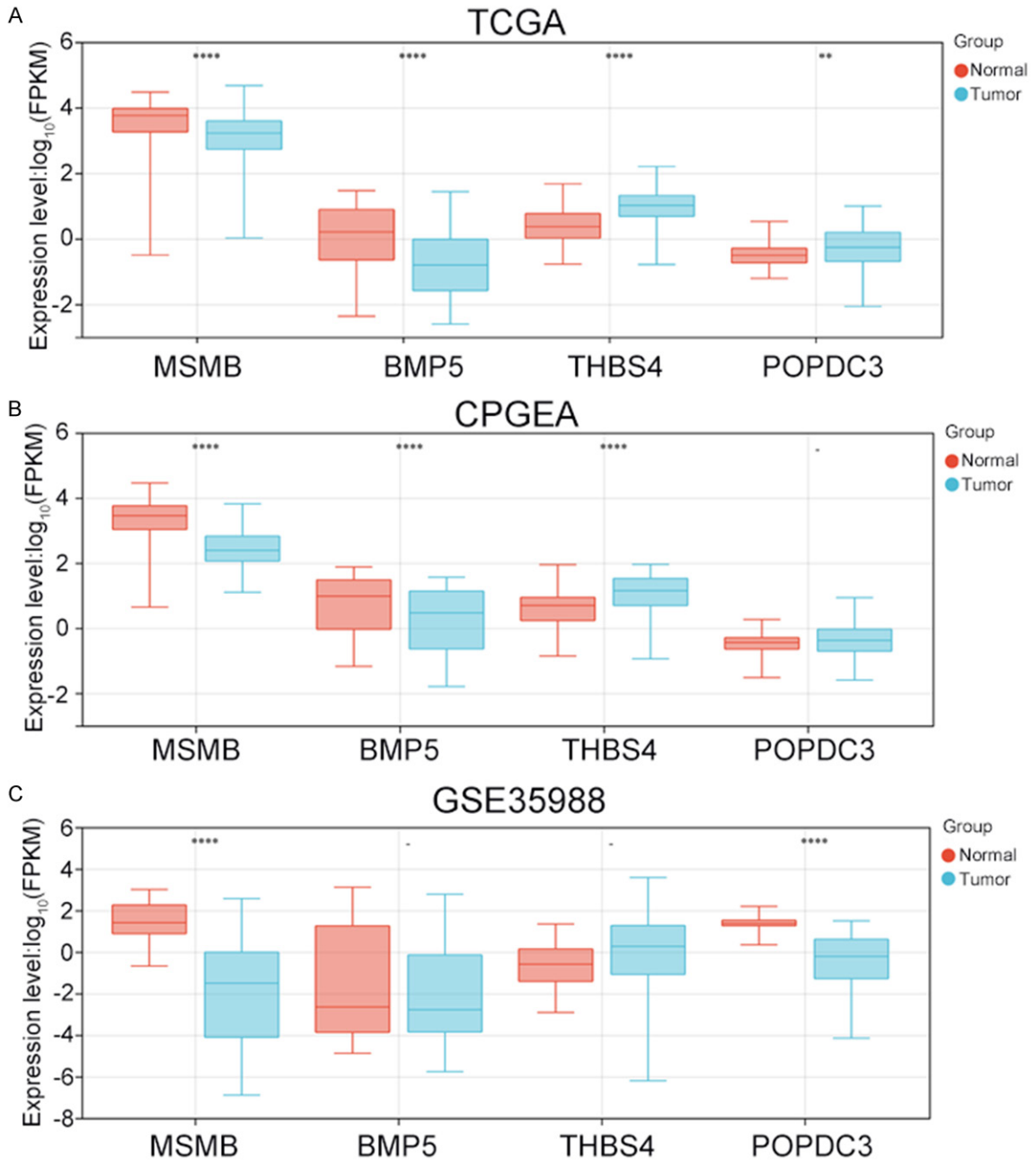


Figure 2. The expression of obesity-related genes in PCA form different databases. A. The expression of four obesity-related genes in PCA from TCGA database. B. The expression of obesity-related genes in PCA from CPGEA database. C. The expression of obesity-related genes in PCA from GSE35988 dataset. - represents no statistics difference. **represents $P < 0.01$, ****represents $P < 0.0001$.

MSMB influences the prognosis of PCA patients

Further, we suspected whether the identified obesity-related genes can influence PCA prognosis. Thus, we used data from the TCGA database and performed cox regression analysis to find the function of these genes in affecting PCA prognosis. With single- and multi-factor Cox regression, we found that the *MSMB* and T category of the TNM staging play a role in DFS of patients (**Figure 3A**, **3B**). Furthermore, the nomogram also predicted that the expression level of *MSMB* can impact the DFS status of PCA (**Figure 3C**). This finding was validated with a calibration curve (**Figure 3D**). Next, we built a

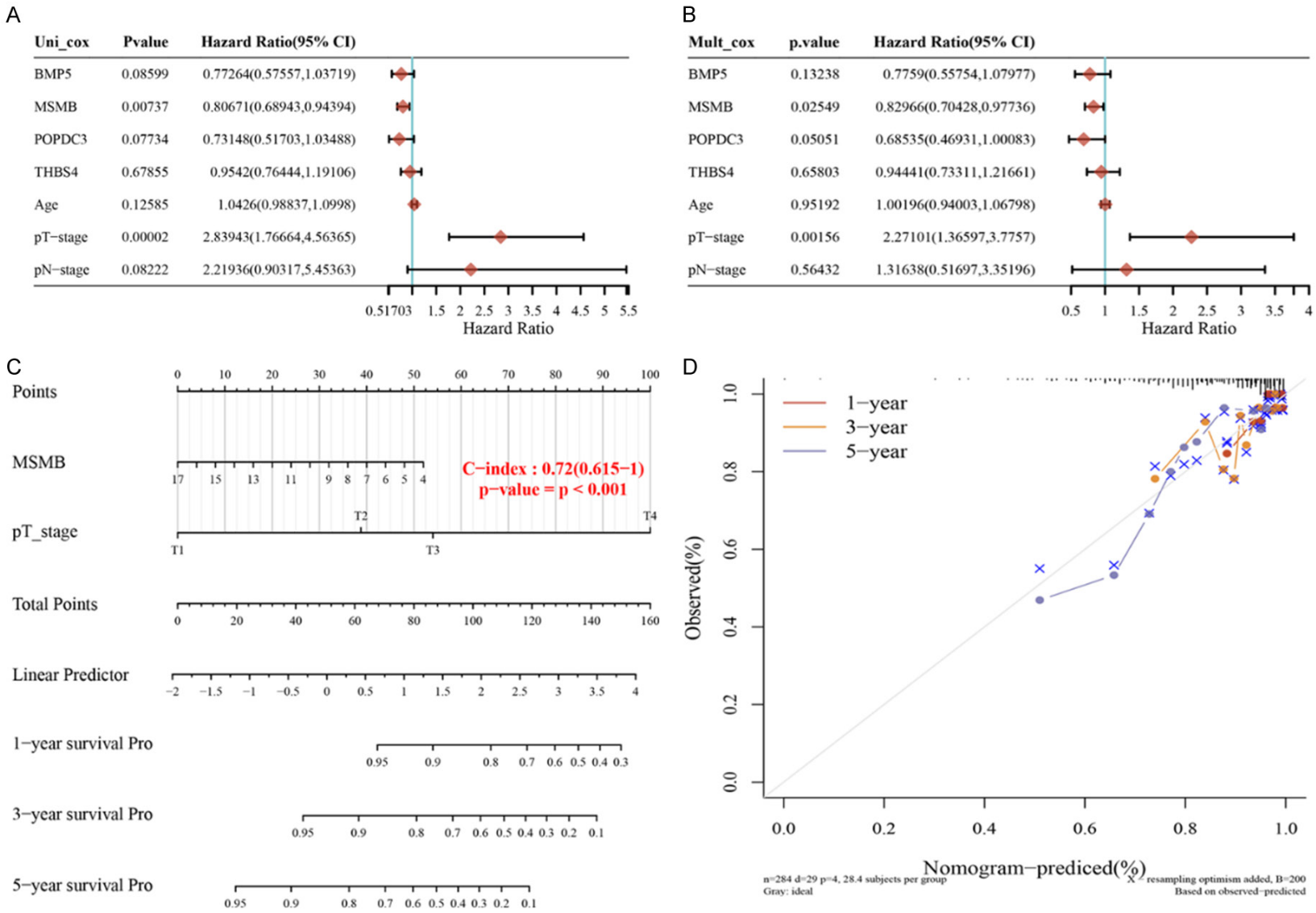


Figure 3. Cox regression reflects the risk of obesity-related genes in impacting patients' DFS. (A) Single-factor cox regression and (B) muti-factor regression reflect the role of obesity-related genes in influencing patients' DFS. (C) Nomogram reflects the function of obesity-related genes in affecting patients' DFS. (D) Calibration curve of nomogram.

lasso regression model and found the risk level of each gene in influencing DFS: risk score = (-0.2145* *MSMB* expression) + (-0.2167* *BMP5* expression) + (-0.0142* *THBS4* expression) + (-0.311* *POPDC3* expression) (Supplementary Figure 6A). We determined the model is effective by calculating Log (λ_1) (Supplementary Figure 6B). We also generated a Kaplan-Meier curve to estimate whether the four genes together can influence the DFS of patients with PCA. Indeed, the expression of these four genes together impact the DFS (P = 0.0013) (Supplementary Figure 6C). Finally, we demonstrated that our model has high accuracy by constructing a time-dependent receiver operating characteristic (ROC) curve (Supplementary Figure 6D). Together, our results indicate that obesity-related genes especially *MSMB* can affect PCA prognosis.

MSMB affects survival status of PCA patients

Because obesity-related genes were associated with PCA occurrence and progression, we investigated whether they are also related to the survival status of patients with PCA. Thus, we analyzed the association of obesity-related genes with DFS and OS of patients with PCA using GEPIA, an online web tool. Among the genes, only *MSMB* affects patients' DFS (Supplementary Figure 7A), but not others (Supplementary Figure 7B-D). However, none of genes can influence patients' OS (Supplementary Figure 8).

The correlation of obesity-related genes and immune cell infiltration in PCA

As obesity can cause immune cell infiltration and it may be one of the reasons for PCA occurrence [10], we sought to find the association between the obesity-related genes and immune cells in PCA by TIMER web tool. We analyzed the mutations in the genes and found that a mutation of *BMP5* and *POPDC3* can impact the infiltration level of immune cell in PCA (Supplementary Figure 9). Next, we investigated the specific correlation of each obesity-related genes and immune cells in PCA. We found purity and macrophages correlated with *MSMB* (Supplementary Figure 10A), all types of cells did with *BMP5* and *THTB4* (Supplementary Figure 10B, 10C). Macrophages, $CD4^+$ T cells, and dendritic cells were associated with *POPDC3* (Supplementary Figure 10D). These findings suggest that obesity-related genes may lead to PCA by affecting immune cell infiltration.

MSMB down expressed in PCA and CRPC samples and decreases PCA cells proliferation

Our previous results indicated that *MSMB* influences PCA development, prognosis and CRPC occurrence. In addition, it is consistently downregulated in obese PCA patients, suggesting *MSMB* plays a vital role in obese people suffered PCA. To verify this hypothesis, we collected 10 clinical samples of PCA and paired paracancerous normal tissues from patients with obesity. In addition, 6 paired tissues from obese CRPC patients were collected, too. The mRNA levels of MSMB were downregulated after PCA occurrence (**Figure 4A**), supporting our previous results. Moreover, the MSMB protein expression corresponded with the transcript (**Figure 4B**). The same protein trend was also found in CRPC patients, too (Supplementary Figure 11A). Then, we tried to confirm whether *MSMB* would decrease when PCA progressed into CRPC. We then compared the mRNA expression of MSMB between PCA and CRPC patients. We found that when patients suffered CRPC, the mRNA of MSMB had a lower expression (Supplementary Figure 11B). Next, we constructed plasmids carrying *MSMB* (oeMSMB) or not (oeControl) and transfected them into two types of PCA cells: PC-3 and DU145 which had been reported that express *MSMB* [15]. We showed that oeMSMB plasmids can elevate MSMB protein levels in PC-3 and DU145 cells (**Figure 4C**, **4D**). Subsequently, we did a CCK-8 assay to assess the effect of *MSMB* on cell proliferation. Remarkably, *MSMB* decreased the proliferation of these two types of PCA cell lines (**Figure 4E**, **4F**). This result indicates that *MSMB* can truly impact PCA occurrence and progression in obese people.

Discussion

The prevalence of people with higher weight and obesity has been increasing worldwide with the changing environment and food habits over the past decades [6, 16]. Obesity is driven by lots of factors, such as diet, environment, heredity and fat metabolism disorder and can lead to many complications by provoking metabolic dysfunction, such as hypertension and hyperlipemia [4]. For example, obesity is a risk

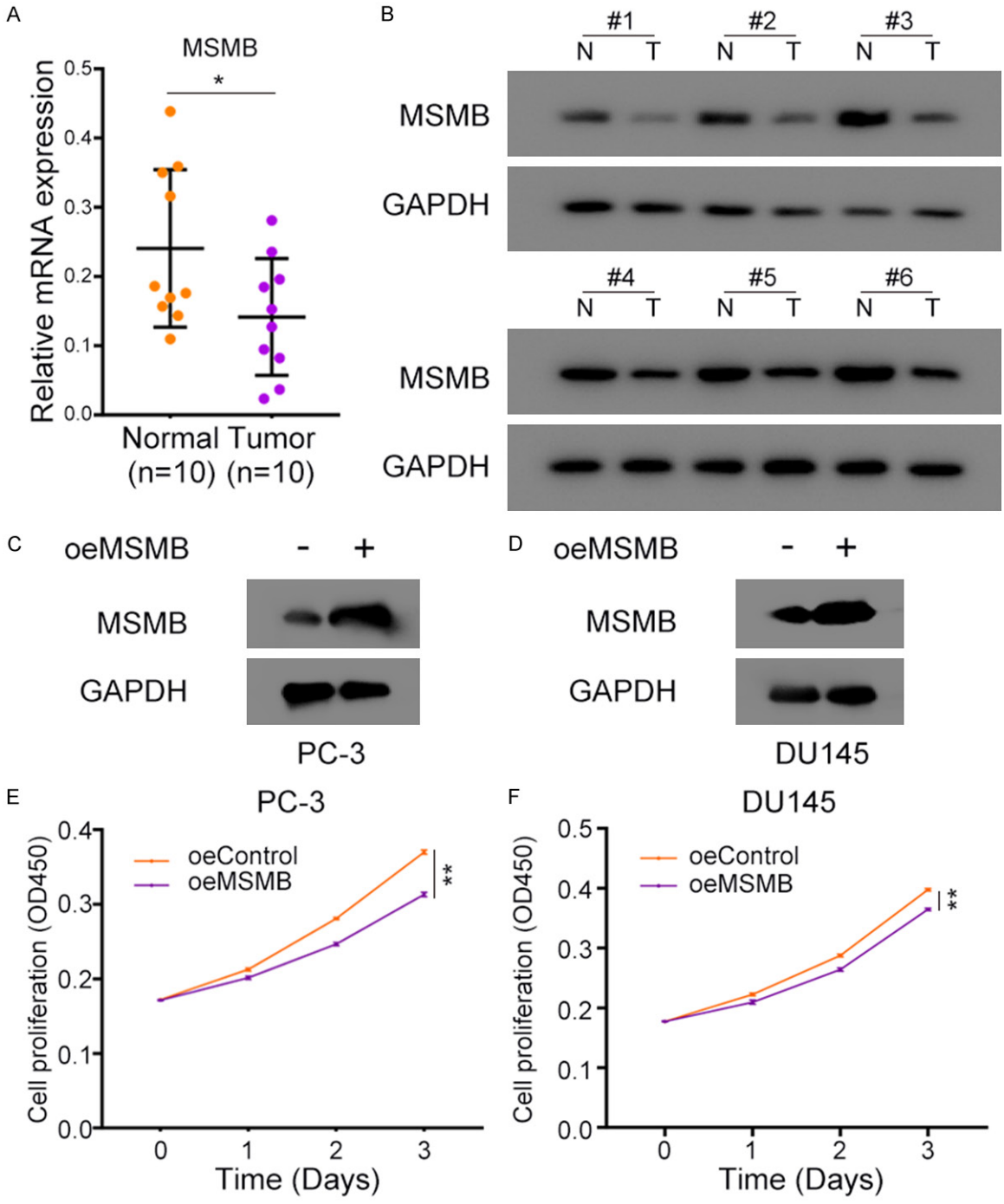


Figure 4. *MSMB* downregulated in obese PCA patients and influence PCA cells proliferation. The mRNA (A) and protein (B) level of MSMB between PCA tumor tissues and adjacent normal tissues in obese PCA patients. (C, D) The protein level of MSMB in two types of PCA cell lines: PC-3 (C) and DU145 (D) after PCA cells transfected by oeMSMB plasmids or not. Cell proliferation of PC-3 (E) and DU145 (F) was detected after cells transfected by oeControl and oeMSMB plasmids. *represents $P < 0.05$. **represents $P < 0.01$. The RT-qPCR and Western blot used GAPDH as endogenous control. N: normal tissues, T: tumor tissues.

factor for many chronic diseases, such as cardiovascular diseases, diabetes, chronic kidney disease, and musculoskeletal disorders [17-20]. Furthermore, obesity is also a risk factor for many types of cancer [21] and may cause breast and colorectal cancers [22, 23]. In addition, abundant evidence shows obesity is also a risk factor for PCA [8, 24].

Obesity can affect the periprostatic fat density which can increase the likelihood of PCA [25]. In addition, adipose tissue and adipokines may increase the chances of developing PCA [9]. Further, many obesity-induced inflammatory factors, such as TNF-α, IL-6, and IL-8, raise the risk of PCA [10]. Finally, deregulated insulin and the insulin-like growth factor (IGF) axes and a compromised adipokine pathway are other biological mechanisms in PCA initiation from obesity [26]. These all indicate that obesity promotes PCA and it may correlate with genetic alteration. However, the specific obesity-related genes leading to PCA are unknown. Therefore, in this study, we aimed to identify the genes that play a part in PCA from obesity.

Using publicly available electronic databases, we found that four obesity-related genes (*MSMB*, *BMP5*, *THBS4*, and *POPDC4*) are associated with PCA. Among these genes, we found *MSMB* can affect PCA occurrence, development, and prognosis. In addition, we also found these obesity-related genes were associated with various immune cells in PCA, suggesting these genes may promote PCA via enhancing immune cell infiltration. Finally, we verified the expression and function of *MSMB* in clinical obese PCA and CRPC patients and PCA cells. We found that *MSMB* was down expression in obese PCA and CRPC patients and it can inhibit PCA cells proliferation. All these results proved the important role of *MSMB* for PCA in obese people.

MSMB is a crucial prostate-specific protein secreted from luminal epithelial cells in the prostate [27]. It is a biomarker for PCA diagnosis since the levels of *MSMB* decrease upon PCA occurrence [28-30], which our results confirm. This gene may also influence PCA progression [31], consistent with our results, too. In addition, a single nucleotide polymorphism in the *MSMB* promoter (rs10993994) can cause PCA [32]. Furthermore, *MSMB* may promote PCA initiation via affecting enhancer of zeste 2 polycomb repressive complex 2 subunit (*EZH2*) expression [15], a gene that catalyzes DNA methylation which means that *MSMB* may impact PCA progression by affecting the methylation level of other genes.

We acknowledge our study has some limitations, although it offers valuable insights. First, the obesity-related genes analyzed in our study were from a single data set (GSE79021). Thus, it may not be comprehensive and lead to bias. Second, though we verified the expression of *MSMB* in different databases and clinical samples, the number of patients used in the study is small. Hence, future studies should include more patients to verify our results. Finally, we only demonstrated the change in *MSMB* expression after PCA and CRPC occurrence. However, how this gene promotes PCA requires additional research to elucidate the mechanism of PCA in patients with obesity. Nonetheless, our study is the first to focus on the specific genes which are important for PCA in obese people.

Conclusion

In conclusion, we found four genes associated with PCA in patients with obesity. One of these genes, *MSMB,* can affect PCA development. It can be served as a biomarker for predicting PCA occurrence and progression in obese people.

Acknowledgements

We thank Bullet Edits Limited for the linguistic editing and proofreading of the manuscript. This study was supported by Natural Science Foundation of Shanghai Municipal Science and Technology Committee (No. 22ZR1456800). National Natural Science Foundation of China (No. 82001610). Clinical Research Plan of SHDC (No. SHDC2020CR3074B). The Shanghai Science and Technology Innovation Action Plan (No. 20Y11904400). The datasets supporting the conclusions of this article are available in the public databases such as GEO, TCGA and CPGEA.

The study was approved by the ethic committee of Tongji Hospital, School of Medicine, Tongji University (SBKT-2021-220). Each participate volunteered to join and signed the informed consent form. The study conformed to the provisions of the Declaration of Helsinki.

Disclosure of conflict of interest

None.

Address correspondence to: Weidong Zhou, Gang Wu and Chengdang Xu, Department of Urology, Tongji Hospital, School of Medicine, Tongji University, Shanghai 200065, The People's Republic of China.

Tel: +86-21-66111530; E-mail: 1811219@tongji.edu.cn (WDZ); wg_urologist@163.com (GW); xuchengdang1990@163.com (CDX)

References

[1] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. CA Cancer J Clin 2021; 71: 7-33.

[2] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J. Cancer statistics in China, 2015. CA Cancer J Clin 2016; 66: 115-132.

[3] Hsing AW and Chokkalingam AP. Prostate cancer epidemiology. Front Biosci 2006; 11: 1388-1413.

[4] Lu X, Jin Y, Li D, Zhang J, Han J and Li Y. Multidisciplinary progress in obesity research. Genes (Basel) 2022; 13: 1772.

[5] GBD 2015 Obesity Collaborators; Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, Lee A, Marczak L, Mokdad AH, Moradi-Lakeh M, Naghavi M, Salama JS, Vos T, Abate KH, Abbafati C, Ahmed MB, Al-Aly Z, Alkerwi A, Al-Raddadi R, Amare AT, Amberbir A, Amegah AK, Amini E, Amrock SM, Anjana RM, Arnlov J, Asayesh H, Banerjee A, Barac A, Baye E, Bennett DA, Beyene AS, Biadgilign S, Biryukov S, Bjertness E, Boneya DJ, Campos-Nonato I, Carrero JJ, Cecilio P, Cercy K, Ciobanu LG, Cornaby L, Damtew SA, Dandona L, Dandona R, Dharmaratne SD, Duncan BB, Eshrati B, Esteghamati A, Feigin VL, Fernandes JC, Furst T, Gebrehiwot TT, Gold A, Gona PN, Goto A, Habtewold TD, Hadush KT, Hafezi-Nejad N, Hay SI, Horino M, Islami F, Kamal R, Kasaeian A, Katikireddi SV, Kengne AP, Kesavachandran CN, Khader YS, Khang YH, Khubchandani J, Kim D, Kim YJ, Kinfu Y, Kosen S, Ku T, Defo BK, Kumar GA, Larson HJ, Leinsalu M, Liang X, Lim SS, Liu P, Lopez AD, Lozano R, Majeed A, Malekzadeh R, Malta DC, Mazidi M, McAlinden C, McGarvey ST, Mengistu DT, Mensah GA, Mensink GBM, Mezgebe HB, Mirrakhimov EM, Mueller UO, Noubiap JJ, Obermeyer CM, Ogbo FA, Owolabi MO, Patton GC, Pourmalek F, Qorbani M, Rafay A, Rai RK, Ranabhat CL, Reinig N, Safiri S, Salomon JA, Sanabria JR, Santos IS, Sartorius B, Sawhney M, Schmidhuber J, Schutte AE, Schmidt MI, Sepanlou SG, Shamsizadeh M, Sheikhbahaei S, Shin MJ, Shiri R, Shiue I, Roba HS, Silva DAS, Silverberg JI, Singh JA, Stranges S, Swaminathan S, Tabares-Seisdedos R, Tadese F, Tedla BA, Tegegne BS, Terkawi AS, Thakur JS, Tonelli M, Topor-Madry R, Tyrovolas S, Ukwaja KN, Uthman OA, Vaezghasemi M, Vasankari T, Vlassov VV, Vollset SE, Weiderpass E, Werdecker A, Wesana J, Westerman R, Yano Y, Yonemoto N, Yonga G, Zaidi Z, Zenebe ZM, Zipkin B and Murray CJL. Health effects of overweight and obesity in 195 countries over 25 years. N Engl J Med 2017; 377: 13-27.

[6] Hawkes C, Smith TG, Jewell J, Wardle J, Hammond RA, Friel S, Thow AM and Kain J. Smart food policies for obesity prevention. Lancet 2015; 385: 2410-2421.

[7] Karra P, Winn M, Pauleck S, Bulsiewicz-Jacobsen A, Peterson L, Coletta A, Doherty J, Ulrich CM, Summers SA, Gunter M, Hardikar S and Playdon MC. Metabolic dysfunction and obesity-related cancer: beyond obesity and metabolic syndrome. Obesity (Silver Spring) 2022; 30: 1323-1334.

[8] Bandini M, Gandaglia G and Briganti A. Obesity and prostate cancer. Curr Opin Urol 2017; 27: 415-421.

[9] Adesunloye BA. Mechanistic insights into the link between obesity and prostate cancer. Int J Mol Sci 2021; 22: 3935.

[10] Olivas A and Price RS. Obesity, inflammation, and advanced prostate cancer. Nutr Cancer 2021; 73: 2232-2248.

[11] Julian V, Ciba I, Olsson R, Dahlbom M, Furthner D, Gomahr J, Maruszczak K, Morwald K, Pixner T, Schneider A, Pereira B, Duclos M, Weghuber D, Thivel D, Bergsten P and Forslund A. Association between metabolic syndrome diagnosis and the physical activity-sedentary profile of adolescents with obesity: a complementary analysis of the Beta-JUDO Study. Nutrients 2021; 14: 60.

[12] Chen X, Ma J, Wang X, Zi T, Qian D, Li C and Xu C. CCNB1 and AURKA are critical genes for prostate cancer progression and castration-resistant prostate cancer resistant to vinblastine. Front Endocrinol (Lausanne) 2022; 13: 1106175.

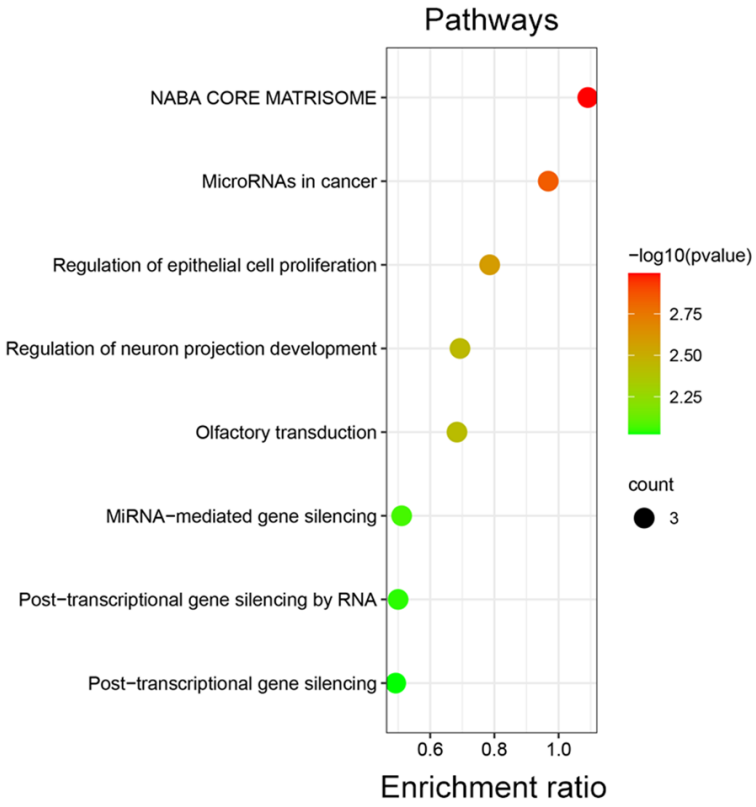
[13] Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC and Lempicki RA. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 2007; 35: W169-175.

[14] Paner GP, Stadler WM, Hansel DE, Montironi R, Lin DW and Amin MB. Updates in the eighth edition of the tumor-node-metastasis staging classification for urologic cancers. Eur Urol 2018; 73: 560-569.

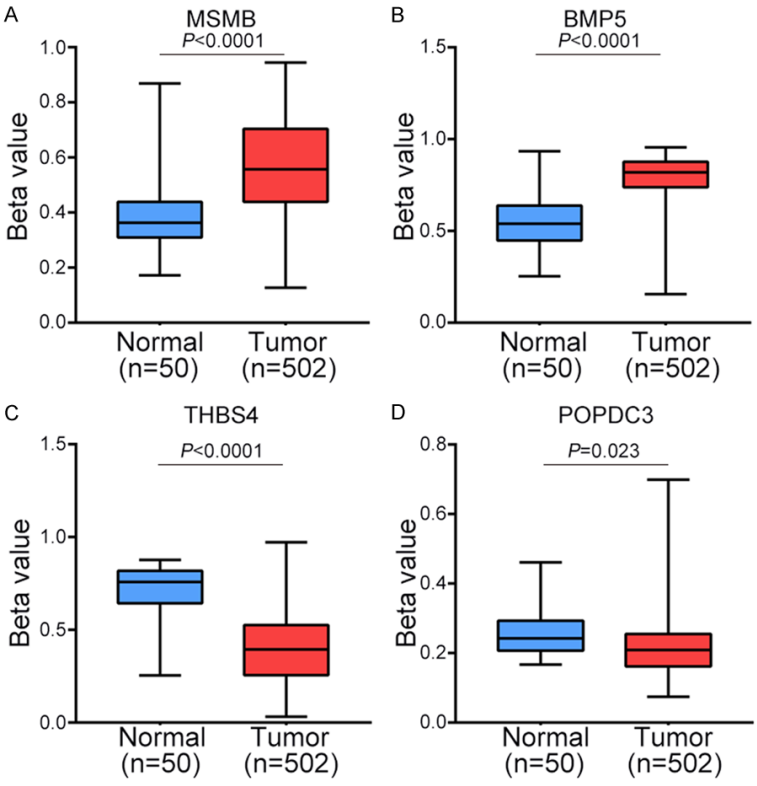
[15] Beke L, Nuytten M, Van Eynde A, Beullens M and Bollen M. The gene encoding the prostatic tumor suppressor PSP94 is a target for repression by the Polycomb group protein EZH2. Oncogene 2007; 26: 4590-4595.

[16] Roberto CA, Swinburn B, Hawkes C, Huang TT, Costa SA, Ashe M, Zwicker L, Cawley JH and Brownell KD. Patchy progress on obesity prevention: emerging examples, entrenched barri-

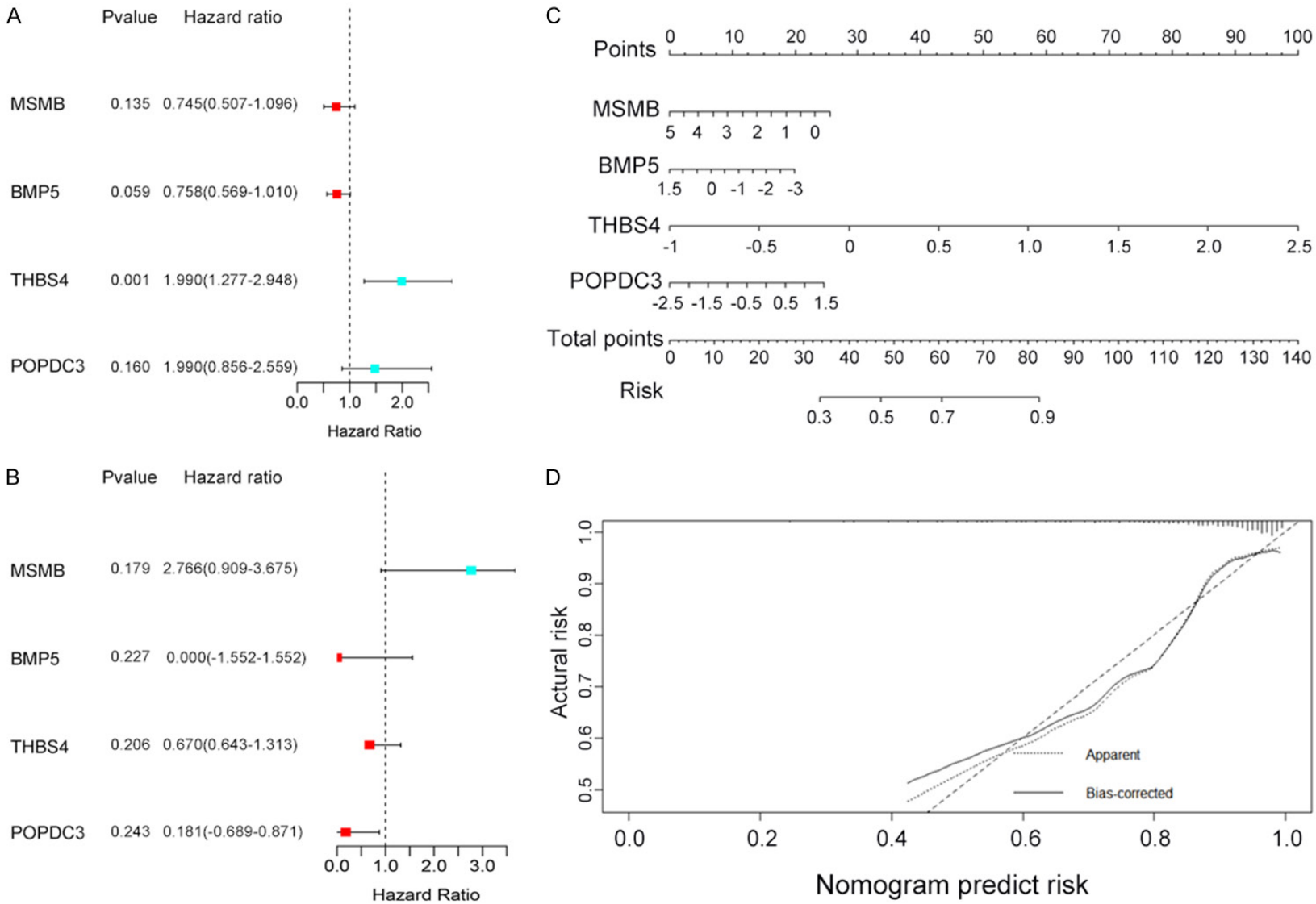
ers, and new thinking. Lancet 2015; 385: 2400-2409.
[17] Singh GM, Danaei G, Farzadfar F, Stevens GA, Woodward M, Wormser D, Kaptoge S, Whitlock G, Qiao Q, Lewington S, Di Angelantonio E, Vander Hoorn S, Lawes CM, Ali MK, Mozaffarian D and Ezzati M; Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group; Asia-Pacific Cohort Studies Collaboration (APCSC); Diabetes Epidemiology: Collaborative analysis of Diagnostic criteria in Europe (DECODE); Emerging Risk Factor Collaboration (ERFC); Prospective Studies Collaboration (PSC). The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. PLoS One 2013; 8: e65174.
[18] Emerging Risk Factors Collaboration, Wormser D, Kaptoge S, Di Angelantonio E, Wood AM, Pennells L, Thompson A, Sarwar N, Kizer JR, Lawlor DA, Nordestgaard BG, Ridker P, Salomaa V, Stevens J, Woodward M, Sattar N, Collins R, Thompson SG, Whitlock G and Danesh J. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. Lancet 2011; 377: 1085-1095.
[19] Jiang L, Rong J, Wang Y, Hu F, Bao C, Li X and Zhao Y. The relationship between body mass index and hip osteoarthritis: a systematic review and meta-analysis. Joint Bone Spine 2011; 78: 150-155.
[20] Jiang L, Tian W, Wang Y, Rong J, Bao C, Liu Y, Zhao Y and Wang C. Body mass index and susceptibility to knee osteoarthritis: a systematic review and meta-analysis. Joint Bone Spine 2012; 79: 291-297.
[21] Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F and Straif K; International Agency for Research on Cancer Handbook Working Group. Body Fatness and Cancer--Viewpoint of the IARC Working Group. N Engl J Med 2016; 375: 794-798.
[22] Bardou M, Barkun AN and Martel M. Obesity and colorectal cancer. Gut 2013; 62: 933-947.
[23] Lee K, Kruper L, Dieli-Conwright CM and Mortimer JE. The impact of obesity on breast cancer diagnosis and treatment. Curr Oncol Rep 2019; 21: 41.
[24] Allott EH, Masko EM and Freedland SJ. Obesity and prostate cancer: weighing the evidence. Eur Urol 2013; 63: 800-809.
[25] van Roermund JG, Hinnen KA, Tolman CJ, Bol GH, Witjes JA, Bosch JL, Kiemeney LA and van Vulpen M. Periprostatic fat correlates with tumour aggressiveness in prostate cancer patients. BJU Int 2011; 107: 1775-1779.
[26] Strong AL, Burow ME, Gimble JM and Bunnell BA. Concise review: the obesity cancer paradigm: exploration of the interactions and crosstalk with adipose stem cells. Stem Cells 2015; 33: 318-326.
[27] Whitaker HC, Warren AY, Eeles R, Kote-Jarai Z and Neal DE. The potential value of microseminoprotein-beta as a prostate cancer biomarker and therapeutic target. Prostate 2010; 70: 333-340.
[28] Bjartell AS, Al-Ahmadie H, Serio AM, Eastham JA, Eggener SE, Fine SW, Udby L, Gerald WL, Vickers AJ, Lilja H, Reuter VE and Scardino PT. Association of cysteine-rich secretory protein 3 and beta-microseminoprotein with outcome after radical prostatectomy. Clin Cancer Res 2007; 13: 4130-4138.
[29] Nam RK, Reeves JR, Toi A, Dulude H, Trachtenberg J, Emami M, Daigneault L, Panchal C, Sugar L, Jewett MA and Narod SA. A novel serum marker, total prostate secretory protein of 94 amino acids, improves prostate cancer detection and helps identify high grade cancers at diagnosis. J Urol 2006; 175: 1291-1297.
[30] Reeves JR, Dulude H, Panchal C, Daigneault L and Ramnani DM. Prognostic value of prostate secretory protein of 94 amino acids and its binding protein after radical prostatectomy. Clin Cancer Res 2006; 12: 6018-6022.
[31] Bergstrom SH, Jaremo H, Nilsson M, Adamo HH and Bergh A. Prostate tumors downregulate microseminoprotein-beta (MSMB) in the surrounding benign prostate epithelium and this response is associated with tumor aggressiveness. Prostate 2018; 78: 257-265.
[32] Whitaker HC, Kote-Jarai Z, Ross-Adams H, Warren AY, Burge J, George A, Bancroft E, Jhavar S, Leongamornlert D, Tymrakiewicz M, Saunders E, Page E, Mitra A, Mitchell G, Lindeman GJ, Evans DG, Blanco I, Mercer C, Rubinstein WS, Clowes V, Douglas F, Hodgson S, Walker L, Donaldson A, Izatt L, Dorkins H, Male A, Tucker K, Stapleton A, Lam J, Kirk J, Lilja H, Easton D; IMPACT Study Steering Committee; IMPACT Study Collaborators; UK GPCS Collaborators, Cooper C, Eeles R and Neal DE. The rs10993994 risk allele for prostate cancer results in clinically relevant changes in microseminoprotein-beta expression in tissue and urine. PLoS One 2010; 5: e13363.



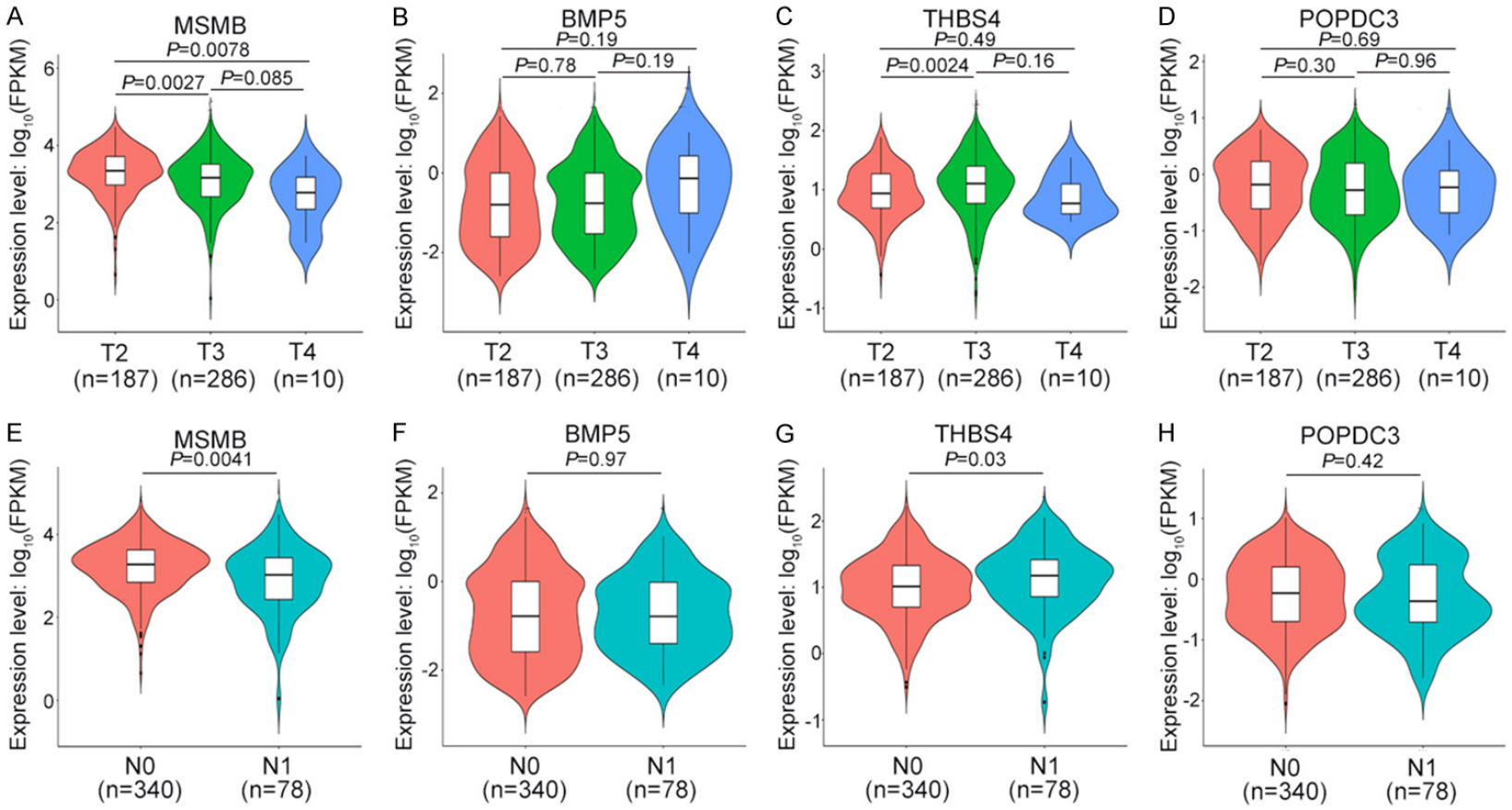
Supplementary Figure 1. The pathways of obesity-related genes enriched.



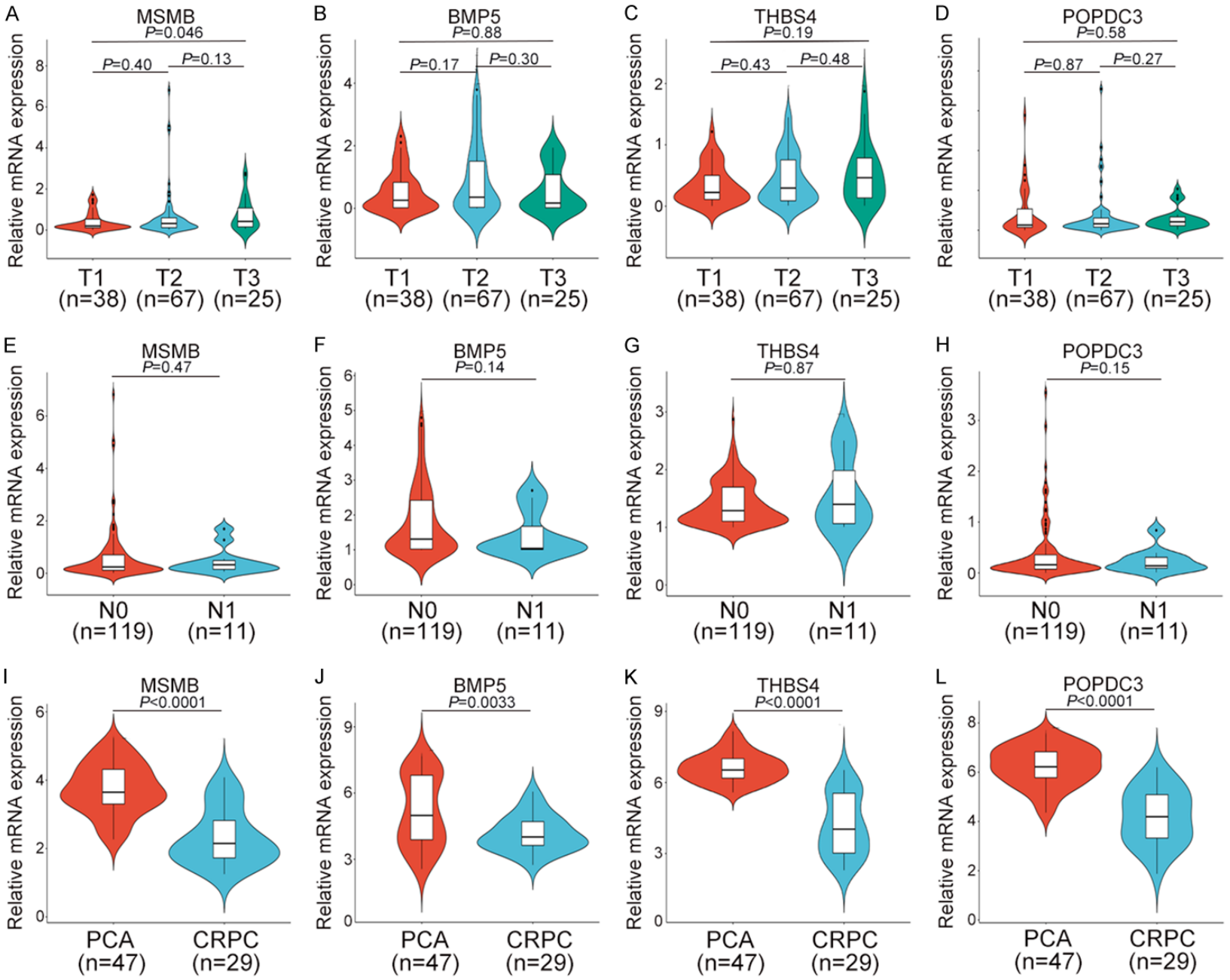
Supplementary Figure 2. The methylation level of obesity-related genes in PCA (data from TCGA database). (A) *MSMB* (B) *BMP5* (C) *THBS4* (D) *POPDC3*.



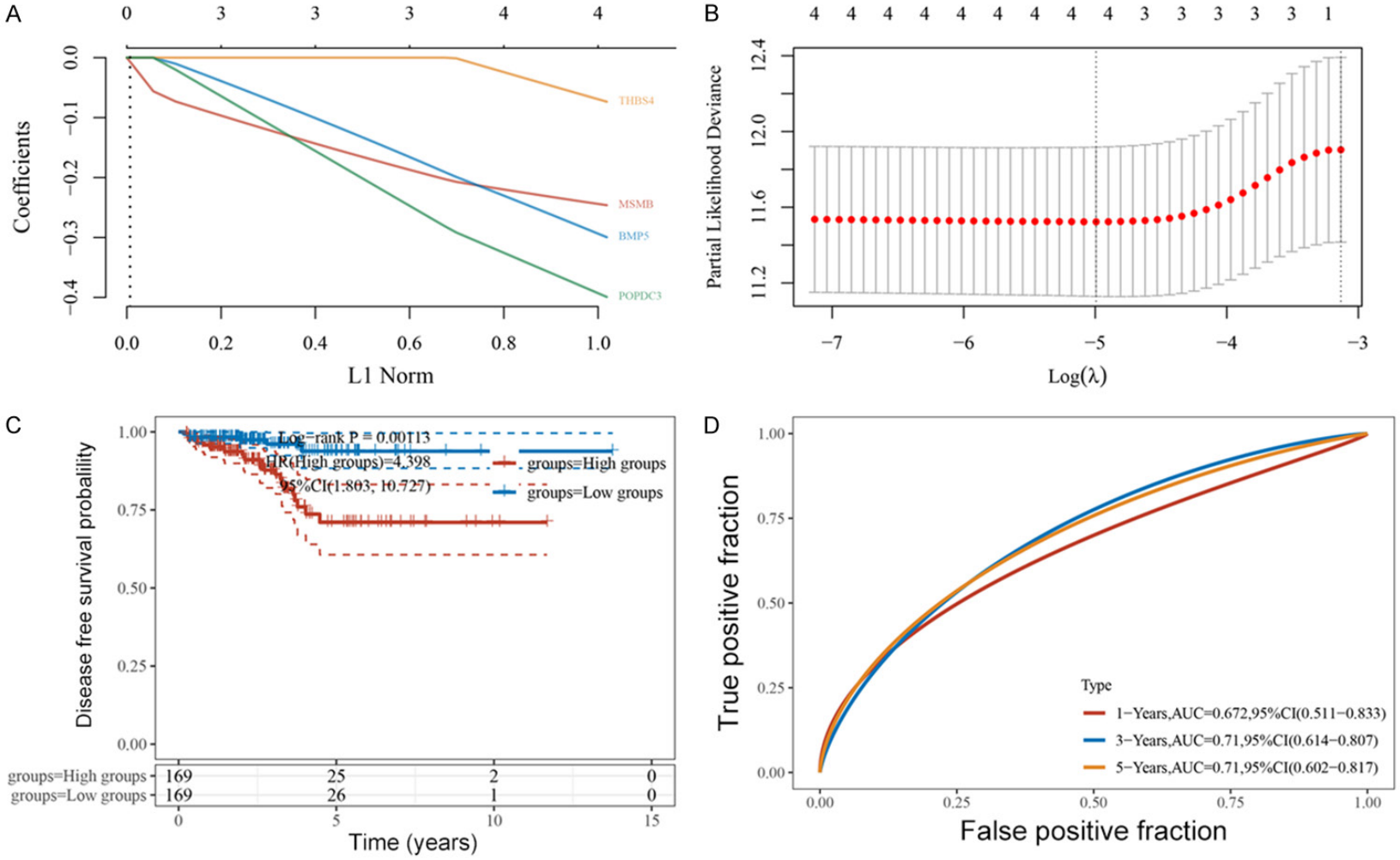
Supplementary Figure 3. Logistic regression reflects the risk of obesity-related genes in leading to PCA. (A) Single-factor logistic regression and (B) muti-factor regression reflect the value of obesity-related genes in causing PCA. (C) The nomogram reflects the risk of obesity-related genes in causing PCA. (D) Calibration curve of nomogram. Data from TCGA database.



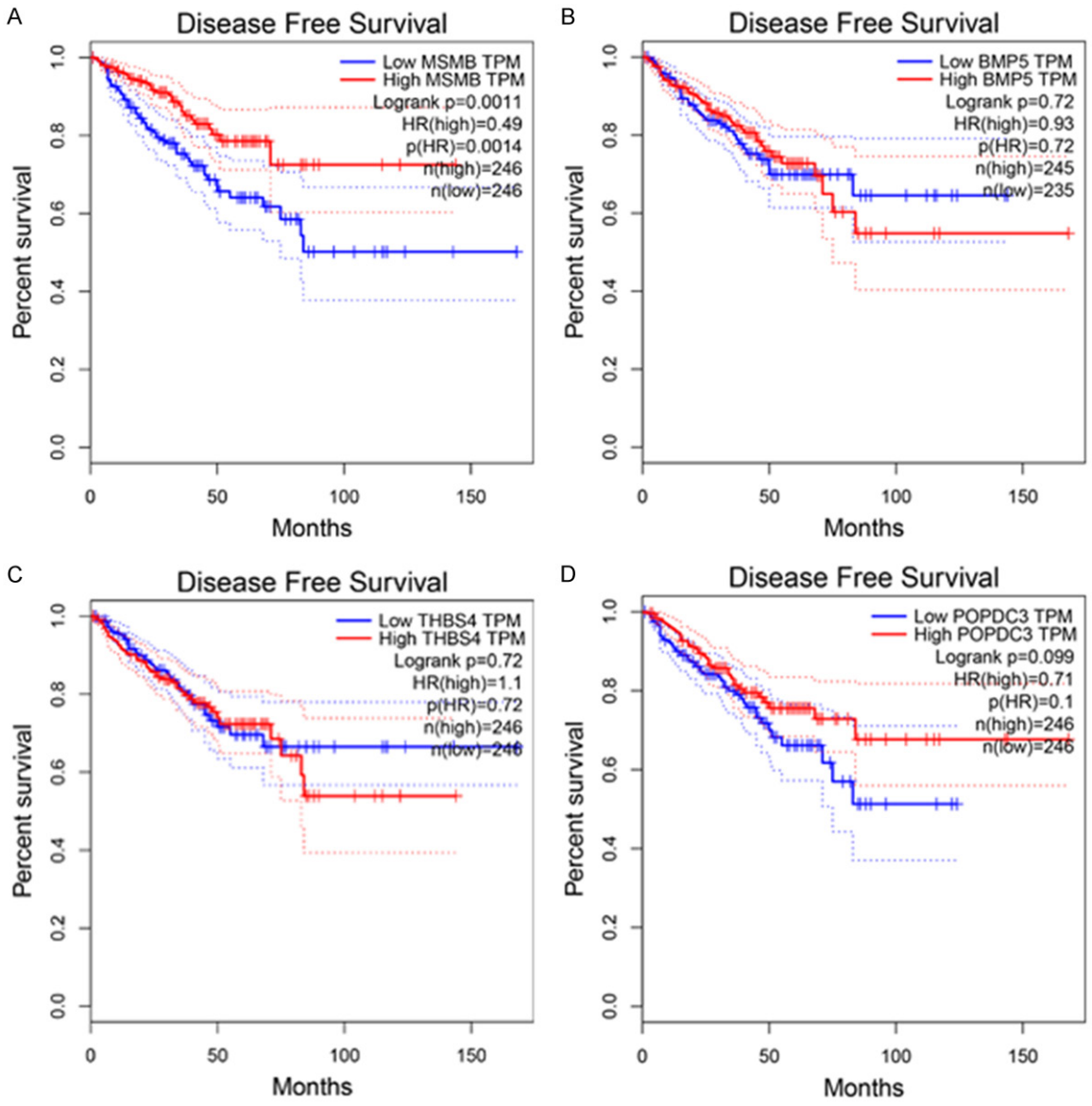
Supplementary Figure 4. The expression of obesity-related genes in different TNM tumor stage of PCA from TCGA. (A-D) The expression of obesity-related genes in different T category (A) *MSMB* (B) *BMP5* (C) *THBS4* (D) *POPDC3.* (E, F) The expression of obesity-related genes in different N category (E) *MSMB* (F) *BMP5* (G) *THBS4* (H) *POPDC3.*


A
MSMB
P=0.046
P=0.40
P=0.13
Relative mRNA expression
T1 (n=38)
T2 (n=67)
T3 (n=25)
B
BMP5
P=0.88
P=0.17
P=0.30
C
THBS4
P=0.19
P=0.43
P=0.48
D
POPDC3
P=0.58
P=0.87
P=0.27
E
MSMB
P=0.47
N0 (n=119)
N1 (n=11)
F
BMP5
P=0.14
G
THBS4
P=0.87
H
POPDC3
P=0.15
I
MSMB
P<0.0001
PCA (n=47)
CRPC (n=29)
J
BMP5
P=0.0033
K
THBS4
P<0.0001
L
POPDC3
P<0.0001

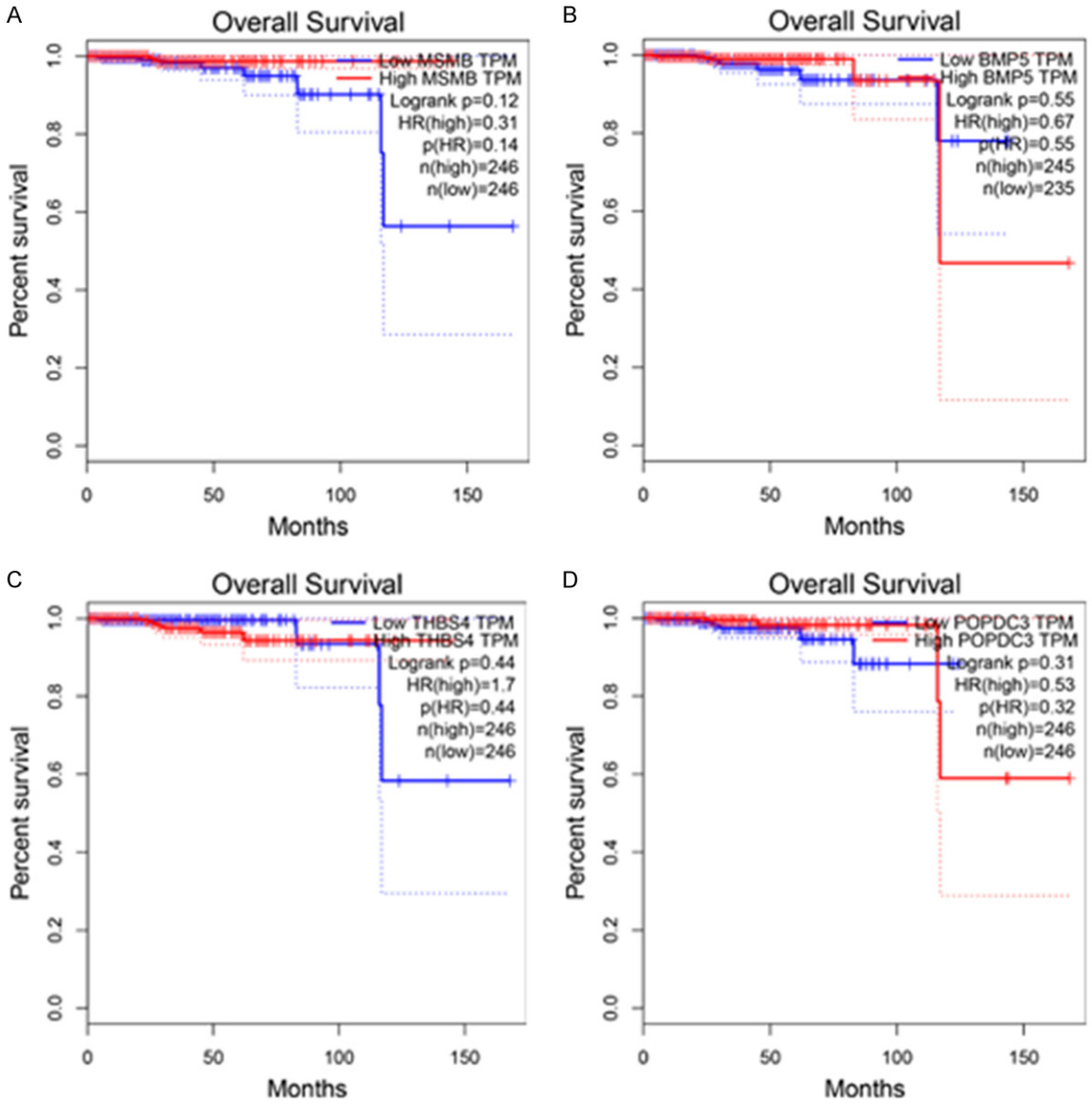
Supplementary Figure 5. The expression of obesity-related genes in different tumor stage of PCA. (A-D) The expression of obesity-related genes in different T stage of Chinese patients (A) *MSMB* (B) *BMP5* (C) *THBS4* (D) *POPDC3.* (E, F) The expression of obesity-related genes in different N stage of Chinese patients (E) *MSMB* (F) *BMP5* (G) *THBS4* (H) *POPDC3.* Data from CPGEA database. (I-L) The expression of (I) *MSMB* (J) *BMP5* (K) *THBS4* (L) *POPDC3* between PCA and CPRC samples from GSE35988 data set.



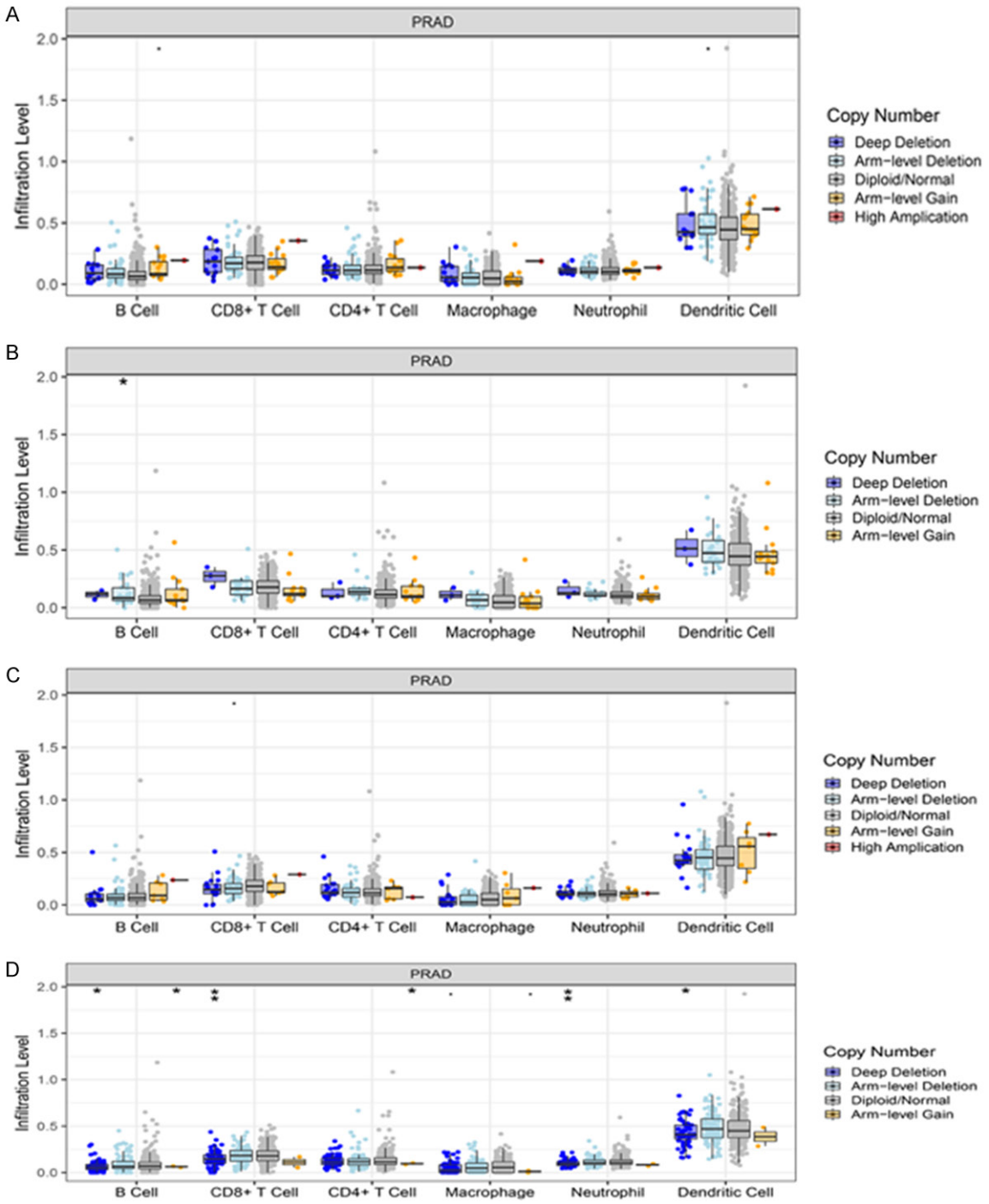
Supplementary Figure 6. Lasso regression reflects the value of obesity-related genes in affecting patients' prognosis. A. The lasso regression of obesity-related genes in affecting PCA patients' prognosis. B. Coefficient of regression model of Lasso. C. Kaplan-Meier curve reflects obesity-related genes effect on PCA patients' DFS. D. Time-depend ROC curve reflects the accuracy of the lasso regression model.



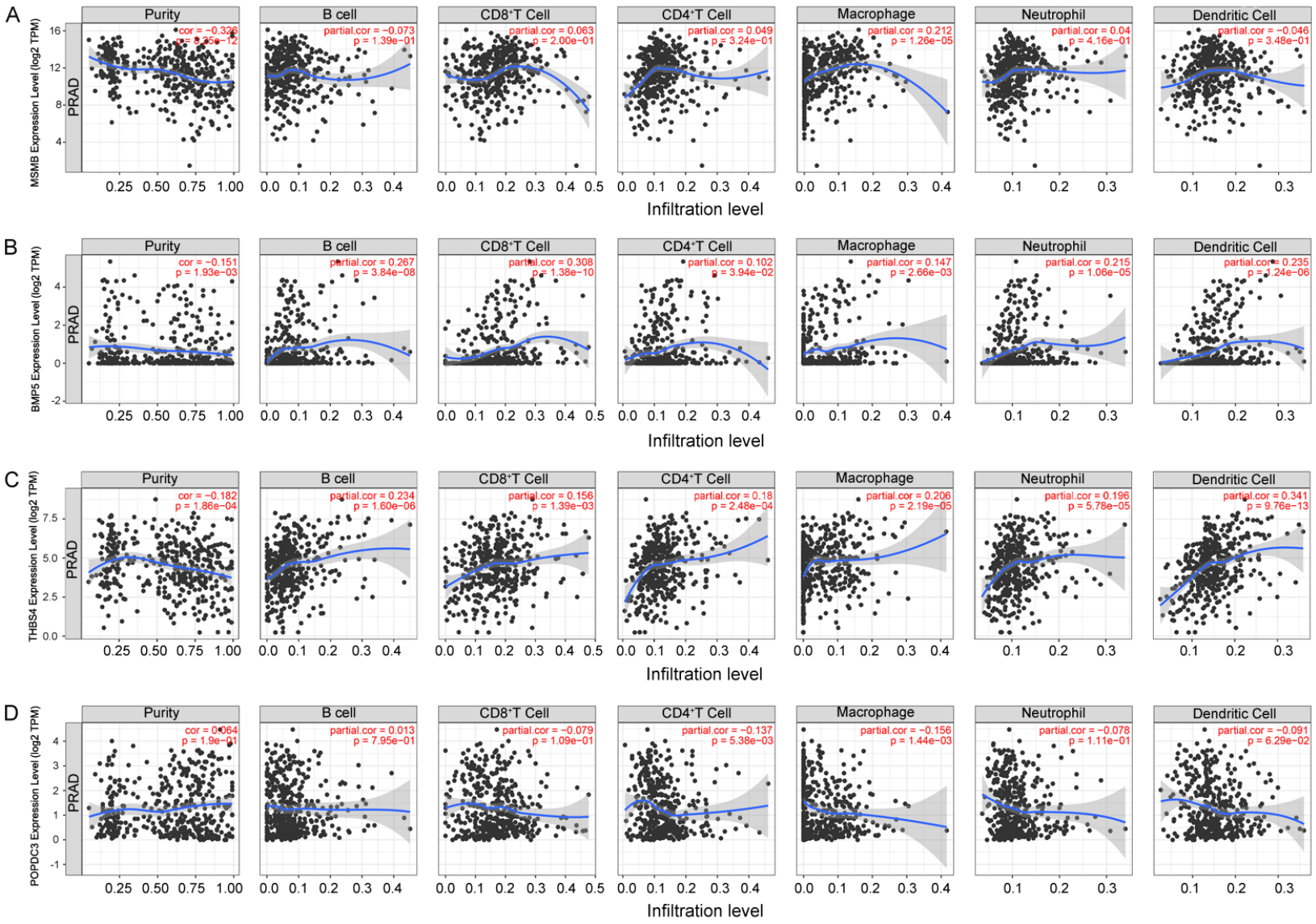
Supplementary Figure 7. The expression level of obesity-related genes in affecting PCA patients' DFS (data from GEPIA online webtool). (A) *MSMB* (B) *BMP5* (C) *THBS4* (D) *POPDC3*.



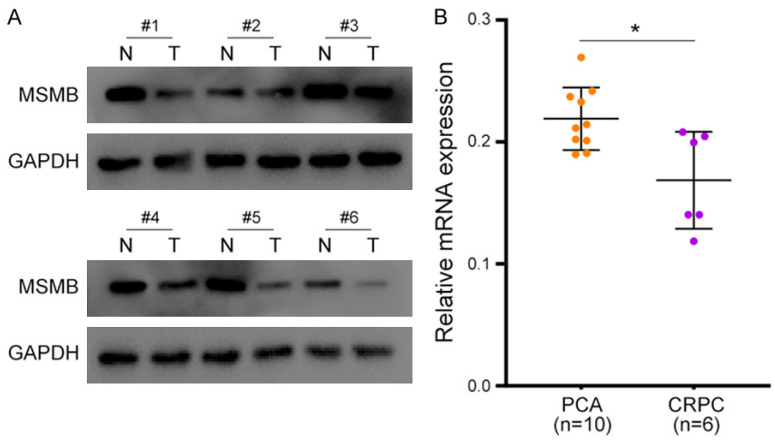
Supplementary Figure 8. The expression level of obesity-related genes in impacting PCA patients' OS (data from GEPIA online webtool). (A) *MSMB* (B) *BMP5* (C) *THBS4* (D) *POPDC3*.



Supplementary Figure 9. The mutation types of obesity-related genes in affecting immune cells infiltration in PCA (Data from TIMER webtool). (A) *MSMB* (B) *BMP5* (C) *THBS4* (D) *POPDC3*. *represents $P < 0.05$, **represents $P < 0.01$.



Supplementary Figure 10. The correlation of mRNA level of obesity-related genes and immune cells infiltration in PCA (Data from TIMER webtool). (A) *MSMB* (B) *BMP5* (C) *THBS4* (D) *POPDC3*.



Supplementary Figure 11. *MSMB* downregulated in obese CRPC tissue samples. A. Compared with normal adjacent prostate tissue, the protein level of MSMB down expression in CRPC tumor tissues. B. The mRNA level of MSMB lower in CRPC samples than PCA samples. *represents $P < 0.05$. N: normal tissues, T: tumor tissues.