

Diagnostic significance and carcinogenic mechanism of pan-cancer gene POU5F1 in liver hepatocellular carcinoma

Dingdong He

Wuhan University Zhongnan Hospital

Xiaokang Zhang

Wuhan University Zhongnan Hospital

Jiancheng Tu (✉ jianchengtu@whu.edu.cn)

Wuhan University Zhongnan Hospital <https://orcid.org/0000-0003-4304-1593>

Research

Keywords: POU5F1, LIHC, biomarker, immune infiltrates, pathogenesis

Posted Date: August 4th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-51170/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Cancer Medicine on September 26th, 2020. See the published version at <https://doi.org/10.1002/cam4.3486>.

Abstract

Background

The prognostic and clinicopathological significance of POU Class 5 Homeobox 1 (POU5F1) among various cancers is disputable heretofore. The diagnostic value and function mechanism of POU5F1 in liver hepatocellular carcinoma (LIHC) have not been studied thoroughly.

Methods

An integrative strategy of meta-analysis, bioinformatics and wet-lab approach was used to explore the diagnostic and prognostic significance of POU5F1 in various types of tumors, especially in LIHC. Meta-analysis was utilized to investigate the impact of POU5F1 on prognosis and clinicopathological parameters in various cancers. The expression level and diagnostic value of POU5F1 were assessed by qPCR in plasma collected from LIHC patients and controls. The correlation between POU5F1 and tumor infiltrating immune cells (TIICs) in LIHC was evaluated by CIBERSORT. Gene set enrichment analysis (GSEA) was performed based on TCGA. Hub genes and related pathways were identified on the basis of co-expression genes of POU5F1.

Results

Elevated POU5F1 was associated with poor OS, DFS, RFS and DSS in various cancers. POU5F1 was confirmed as an independent risk factor for LIHC and correlated with tumor occurrence, stage and invasion depth. The combination of POU5F1 and AFP in plasma was with high diagnostic validity (AUC = 0.902, $P < 0.001$). Specifically, the level of POU5F1 was correlated with infiltrating levels of B cells, T cells, dendritic cells and monocytes in LIHC. GSEA indicated POU5F1 participated in multiple cancer related pathways and cell proliferation pathways. Moreover, CBX3, CCHCR1 and NFYC were filtered as the central hub genes of POU5F1.

Conclusions

Our study identified POU5F1 as a pan-cancer gene could not only be a prognostic and diagnostic biomarker in various cancers, especially in LIHC, but functionally carcinogenic in LIHC.

Background

Cancer has become a key influence factor of morbidity and mortality in both developed and developing countries [1]. There will be an escalating trend of death rates caused by cancers in the future due to deficient cognition in the pathological processes and regulatory mechanisms of cancers [2]. Although the prognoses of cancers have been ameliorated through various therapeutic methods, the prognostic outcomes are invariably unsatisfactory in multiple kinds of cancers. Among all kinds of cancers, liver hepatocellular carcinoma (LIHC) acts as one of the major roles. According to the cancer statistics data from 2020, LIHC ranks sixth in mortality among all cancers [3]. As the main diagnostic and prognostic biomarker for LIHC, the sensitivity and specificity of α -fetoprotein (AFP) were ungratified in the early diagnosis of LIHC. Consequently, urgent requirements are raised to find novel biomarkers as potential diagnostic indicators and therapeutic targets of LIHC.

Many studies have certified cancer stem cells (CSCs) were associated with aggression, metastasis, and recrudescence in various cancers. In addition, several CSCs markers have been proved to contribute to poor prognosis of cancers [4–7], indicating the significance of CSCs markers in prognosis of malignancies. But due to the complexity of the regulatory network in tumor pathologic processes, the prognostic significance of CSCs markers has not been fully understood. With more in-depth studies on CSCs markers, some of these markers may become important targets in cancer diagnosis, therapy and prognosis.

POU Class 5 Homeobox 1 (POU5F1), is a transcription factor of the POU family which binds an octameric sequence motif so as to activate the expression of downstream genes [8]. POU5F1 has been identified to be one of the most important CSCs markers and participates in stemness maintenance of various tumors [9, 10]. Published literatures have certified that increased POU5F1 was correlated with clinicopathological features and prognosis not only in LIHC, but also in bladder carcinoma, non-small-cell lung carcinoma and oral squamous cell cancer [11–14]. POU5F1 may serve as an essential predict factor for multiple cancers in the near future.

Though plentiful researches have been performed, the prognostic significance of POU5F1 in cancers remains controversial and the functions of POU5F1 in the regulatory network of tumor are not fully recognized. Some studies have come to different or even totally opposite conclusions towards the prognostic value of POU5F1 and the role of POU5F1 played in tumor development. For instance, He et al. showed elevated POU5F1 in esophageal squamous cell carcinoma symbolized poor survival outcomes [15]. However, Ge et al. found that high expression of POU5F1 was connected with longer survival in esophageal squamous cell carcinoma [16]. The prognostic value of POU5F1 in LIHC was not statistically significant according to Qian et al [17], but was prominent in studies performed by Huang et al [18]. These disputes have not been settled in a reasonable way and the value of POU5F1 in tumor prognosis is still ambiguous. Meanwhile, current studies on the role of POU5F1 in LIHC mainly used tissue samples, hindering the clinical application of POU5F1 as a diagnostic biomarker due to the invasiveness. Researches focus on the POU5F1 status in plasma could make it easier in promotion.

In this study, we adopted an integrative strategy of meta-analysis, bioinformatics and wet-lab approach to explore the diagnostic and prognostic significance of POU5F1 in various types of tumors, especially in LIHC. First, we performed meta-analysis and trial sequential analysis (TSA) with a large sample size to evaluate the significance of POU5F1 for survival prognosis in various cancers. And LIHC was selected as the major target when combined the meta-analysis results with the survival analysis results from TCGA datasets. Then, we validated POU5F1 expression level in plasma and evaluated the diagnostic value of POU5F1 in LIHC. Further, a protein-protein interaction (PPI) network was constructed based on the co-expression genes of POU5F1 and central hub genes were recognized. Finally, cell signal transduction diagram was drawn to clarify the potential function pathways of POU5F1 in LIHC.

Methods

Literature search strategy

We comprehensively retrieved PubMed, Embase, Web of Science and Cochrane Library to search studies published during 1 January 2000 to 1 June 2019 with language limitation of English and screened studies reported prognosis and clinicopathological features in cancer patients with aberrant expression of POU5F1. To increase search sensitivity, we used a strategy involving both Medical Subject Heading terms and free-text words. The search strategy was segmented into 3 parts: “POU5F1 transcription factor or octamer transcription factor 4 or octamer transcription factor 3” and “neoplasms or malignant neoplasms or carcinoma” and “prognosis or prognostic factors or survival”. We also manually browsed the references of retrieved articles to recognize more eligible studies that might have been missed by the search strategy.

Literature inclusion and exclusion criteria

Published articles that met the 7 criteria would be enrolled: (1) evaluated the association between POU5F1 expression and clinical prognosis or clinicopathological parameters of cancers; (2) provided hazard ratios (HRs) and 95% CI or survival curves of POU5F1 relevant outcomes; (3) cohort studies (follow-up duration longer than 24 months); (4) whole paper was written in English; (5) available full-text articles; (6) research on human; (7) sample size of cancer patients was no less than 20. The exclusion criteria including: (1) absence of essential data, such as detection methods of POU5F1 expression, survival analyses data, accurate prognosis indicators; (2) reviews, case reports, letters, conference abstracts, animal trials, or duplicate publications.

Literature data extraction and quality evaluation

Each process of our research was strictly in conformity to preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines [19]. Important features of the eligible cohorts were recorded, including first author; published year; nation; sample size; tumor category; age and gender of the patients; detection method and cut off value for POU5F1; follow-up period; study design; clinicopathological parameters; outcome of interest, including overall survival (OS), disease-free survival (DFS), disease-specific survival (DSS) and recurrence-free survival (RFS). Newcastle-Ottawa Scale (NOS) was utilized for appraising quality of included cohorts. According to the NOS criteria, a cohort was considered of high quality when the total score was no less than 7 [20].

Trial sequential analysis

With new studies constantly enrolled into the cumulative meta-analysis, type I and type II errors might increase due to repetitive test of significance, fragmentary data, and ambiguous publication bias [21]. Trial sequential analysis (TSA) can overcome these obstacles and estimate a priori information size (APIS) which is considered as the minimal sample size required to draw a reliable conclusion. When the cumulative Z-curve fails to cross conventional boundary ($Z=1.96$), it indicates that the result is farfetched. If the Z-curve crosses the conventional boundary but doesn't reach the TSA boundary, meaning the trials show false positive results. If the Z-curve crosses both the conventional boundary and TSA boundary but not the APIS, it suggests that more researches are needed to support the conclusion. If the Z-curve crosses all of the three boundaries, a reliable conclusion has been certified. We implemented TSA by maintaining two-sided α of 5%, 15% relative risk reduction (RRR) and statistical test power of 80%. We performed TSA with fixed-effects model when I^2 was less than 30%. Elsewise, random-effects model would be executed.

Expression and survival analysis based on TCGA

Tumor Immune Estimation Resource (TIMER) is an online database that incorporates expression profiles of 10,009 samples across 23 cancer types from TCGA (<https://cistrome.shinyapps.io/timer/>) [22]. We utilized TIMER to confirm the expression level of POU5F1 in various cancers. Gene Expression Profiling Interactive Analysis (GEPIA), another online analysis tool, contains survival and clinicopathological data extracted from various cancers based on TCGA (<http://gepia.cancer-pku.cn/>) [23]. Survival analysis of OS and DFS were executed by GEPIA to find the correlation between POU5F1 expression and prognosis of various cancers.

Specimens

Plasma specimens of 30 LIHC patients from Zhongnan Hospital of Wuhan University (Wuhan, China) were collected during July 2017 and October 2019 and stored at -80°C until use. LIHC patients were identified on the basis of their pathology reports. Meanwhile, 30 healthy people without hepatic diseases nor abnormal liver biochemical outcomes

were enrolled as controls. Our study was authorized by the Medical Ethics Committee of Zhongnan Hospital of Wuhan University.

RNA extraction and quantitative PCR analysis

RNA was extracted from plasma by Total RNA Separate Extraction Kit (Biotek, China) according to the manufacturer's instruction. NanoDrop 2000C was applied to assess the concentration and purity of RNA. ReverTra Ace qPCR RT Kit (Toyobo, Japan) was used to reversely transcribe mRNA into complementary DNA (cDNA). The quantitative PCR (qPCR) was implemented using SYBR Green I UltraSYBR Mixture (CWBI, China) on Bio-Rad CFX96 (Bio-Rad Laboratories, USA). *Glyceraldehyde 3-phosphate dehydrogenase (GAPDH)* was taken as the endogenous reference gene. The detailed sequences of each pair primers were listed in Additional file 1. All experiments were repeated twice. The expression status of target gene was assessed by $2^{-\Delta Cq}$ method, in which ΔCq represents the value of the mean quantification cycle (Cq) of target gene subtracts the mean Cq of endogenous reference gene.

Tumor infiltrating immune cells reckoning

CIBERSORT provides a deconvolution algorithm that is able to distinguish 22 kinds of tumor infiltrating immune cells (TIICs) from other cell types in tissues [24]. Expression profiles of 50 normal liver tissues and 374 LIHC tumor tissues were downloaded from TCGA database and TIICs proportions of each sample were evaluated by R (Version 3.6.2) on the basis of CIBERSORT algorithm. Then TIICs proportions of normal liver tissues and LIHC tumor tissues were divided into 2 subgroups respectively based on the median of POU5F1 expression level and visualized through violin plots.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) is a bioinformatics method that inspects the statistical significance of a priori defined set of genes and verifies the differences between two biological states [25]. We divided TCGA LIHC samples into 2 phenotype subgroups on the grounds of the median expression level of POU5F1. Genes from the TCGA expression profiles were ranked in a list according to the degree of divergence between high POU5F1 subgroup and low POU5F1 subgroup through GSEA software 4.0. Then, Gene ontology (GO) gene sets and Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets were analyzed to identify functional terms and pathways enriched in each phenotype subgroup. Gene set permutations were executed 1000 times for each analysis. The criteria of significantly enriched pathways were normalized P value < 0.05 and the absolute value of normalized enrichment score (NES) > 1.5 .

Enrichment analysis of POU5F1 co-expression genes

Co-expression genes of POU5F1 were screened out by R based on expression profiles of TCGA. The cutoff line was set at $P < 0.05$ and the absolute value of spearman correlation coefficient > 0.45 . GO enrichment analysis and KEGG pathway analysis were performed through R package "clusterProfiler". $P < 0.05$ was taken as statistically significant enriched pathways.

PPI network establishment and hub genes identification

We utilized Search Tool for the Retrieval of Interacting Genes (STRING) database to establish protein-protein interaction (PPI) network, in order to discover the relationship among co-expression genes of POU5F1. The interaction score was set at 0.4 in STRING database. Cytoscape was used to enhance the legibility of PPI network on the basis of interaction data obtained from STRING database. We considered genes that interacted directly with POU5F1 as central hub genes and those that directly interacted with the central hub genes as subordinate hub genes.

Statistical analysis

All statistical analyses in this study were performed through Stata SE15 (Stata Corporation, USA), SPSS 25.0 (SPSS Inc., USA), GraphPad Prism 8.0 (GraphPad Inc., USA) and R (Version 3.6.2). Amalgamative HRs and relating 95% CIs in meta-analysis were computed by Stata SE15. If the studies didn't provide HRs or corresponding 95% CIs, these values were calculated by the equation: $HR = (P0/(1 - P0))/(P1/(1 - P1))$, in which P0 and P1 stood for survival rate of decreased and elevated POU5F1 subgroup, respectively. 95% CI was calculated through $\exp(\ln HR \pm 1.96 \times \text{stderr})$, exp represented exponential, lnHR was natural logarithm of HR, stderr meant standard error of HR. Several studies didn't report the relevant data about survival rate in subgroups, we utilized Engauge Digitizer Version 10.8 to collect representative data on Kaplan–Meier survival curves. Then the extracted data was imported into a computation sheet obtained from Tierney et al. for estimation of HRs and 95% CIs [26]. Heterogeneity of enrolled cohorts was evaluated through Chi square-based Q and I^2 analyses. We ran a meta-analysis with fixed-effects model when the heterogeneity was acceptable ($I^2 < 50\%$ or $P > 0.05$). Otherwise, the random-effects model was performed. Sensitivity analysis was conducted through sequentially expurgated each cohort to evaluate the stability of the amalgamative result. Potential publication biases were detected through Begg's and Egger's analysis.

Continuous variables with normal distribution were described by mean \pm standard deviation (SD). Median and inter-quartile ranges were used to describe abnormally distributed continuous variables. Student's *t* test and Mann-Whitney *U* tests were utilized for comparison between two groups. Correlation analyses were conducted by Pearson and Spearman correlation test. Chi-square test and Fisher's exact test were applied for evaluation of categorical variables. Cox proportional hazard regression was used for univariate and multivariate analysis. Odds ratios (ORs) were calculated by logistic regression. Receiver operation curve (ROC) was performed to assess the diagnostic values. Statistically significant threshold of two-sided *P* value was set at 0.05.

Results

Literature search results and quality evaluation

The retrieving procedure was illustrated in Additional file 2. The databases search obtained 1542 references. 1205 articles were left after exclusion of duplicates. 909 records were excluded by scanning the titles and abstracts. 296 studies were evaluated through browsing the full-text. Eventually, 57 studies containing 7401 patients were enrolled into our study [10-18, 27-74]. 16 types of cancers were contained, including acute myeloid leukemia, bladder cancer, breast cancer, cervical cancer, colorectal cancer, esophageal cancer, gallbladder adenocarcinoma, gastric cancer, head and neck cancer, LIHC, lung cancer, neuroblastomas, ovarian cancer, pancreatic cancer, papillary renal cell carcinoma and prostate cancer. The expression levels of POU5F1 were detected by immunohistochemistry (IHC) in 44 studies, qPCR in 10 studies, and immunofluorescence (IF) in the remaining 3 studies. Essential features of enrolled studies were exhibited in Additional file 3. The quality assessments of enrolled studies were implemented through NOS and 45 studies were rated as high-quality studies with comprehensive scores greater than 7 points (Additional file 4).

Overall analysis of POU5F1 expression and cancer prognosis

Among the included 57 studies, a total of 5,485 subjects in 48 studies described the relationship between POU5F1 expression and OS, 1649 subjects in 14 studies for DFS, 5 studies with 1249 subjects for DSS and 6 studies involved 636 subjects for RFS. According to the meta-analyses, the heterogeneities were not distinct in these 4 kinds of prognosis analyses (Fig. 1a, c). Therefore, we capitalized fixed-effects model to calculate the amalgamative HRs and relating 95% CIs. The results showed that increased POU5F1 was correlated with inferior outcomes for OS (HR = 2.45, 95% CI = 2.22–2.71, $P < 0.001$), DFS (HR = 2.66, 95% CI = 2.22–3.19, $P < 0.001$), DSS (HR = 4.03, 95% CI = 2.70–6.01, $P < 0.001$) and RFS (HR = 2.59, 95% CI = 1.85–3.63, $P < 0.001$).

Subgroup analysis for OS and DFS

Subgroup analyses were implemented for OS and DFS to clarify the connection between POU5F1 expression and cancer type, analysis type, sample size, detection method. Studies were defined as “other cancers” in the cancer type subgroup when there was only one enrolled study for each kind of cancer. As demonstrated in Fig. 1b and Table 1, elevated expression of POU5F1 predicted poor prognosis of OS in bladder cancer, breast cancer, colorectal cancer, esophageal cancer, gastric cancer, LIHC, head and neck cancer, lung cancer and other cancers, including ovarian cancer, cervical cancer, neuroblastomas and pancreatic cancer. But the prognostic value of POU5F1 was not obvious in overall survival of acute myeloid leukemia. Simultaneously, overexpression of POU5F1 was related to shorter DFS in head and neck cancer, breast cancer, LIHC, colorectal cancer and other cancers, including lung cancer, gastric cancer, cervical cancer and acute myeloid leukemia (Fig. 1d, Additional file 5). Furthermore, the subgroup category of analysis type, sample size and detection method also indicated the observable relationship between high level of POU5F1 and shorter OS and DFS.

Correlation of POU5F1 and clinicopathological characteristics

To explore why elevated POU5F1 could lead to worse prognosis in various cancers, the correlations between POU5F1 status and neoplastic clinicopathological parameters were evaluated (Table 2). The overexpression of POU5F1 was remarkably correlated with tumor size, TNM stage, tumor differentiation, tumor invasion depth, lymph node metastasis, distant metastasis, lymphovascular invasion, vascular invasion, tumor number and tumor recurrence. Non-statistically significant results were found about age, gender, tumor encapsulation, liver cirrhosis, HBsAg and smoke.

Reliability of pooled prognostic results

TSA was implemented to assess the reliability of our meta-analysis results (Additional file 6). The heterogeneity of OS ($I^2 = 4.40\%$), DFS ($I^2 = 20.49\%$) and DSS ($I^2 = 0.00\%$) was not obvious, so the fixed model was utilized to perform TSA. Whereas, heterogeneity appeared in RFS ($I^2 = 32.53\%$), thus random model was adopted. The accumulated Z-curve of OS crossed traditional boundary, TSA boundary and APIS, suggesting the conclusion was significantly reliable. The cumulative Z-curve of DFS, DSS and RFS crossed the conventional boundary and TSA boundary but didn't reach APIS, indicating the current trails have obtained positive results and more studies were required to support the results. Sensitivity analyses were executed to detect the stability of the conclusions about the prognostic value of POU5F1. No individual cohort could distinctly affect the pooled HRs of OS, DFS, DSS or RFS, meaning the conclusions were credible (Additional file 7). The underlying publication bias was appraised through Begg's and Egger's analysis. And there was no potential publication bias found in OS, DFS, DSS or RFS (Additional file 8).

Expression and prognostic role of POU5F1 in various cancers

To further verify the expression level of POU5F1 in various cancers, TIMER was adopted to analyze the expression profiles from TCGA. As displayed in Fig. 2a, POU5F1 was prominently upregulated in bladder urothelial carcinoma (BLCA); breast invasive carcinoma (BRCA); cholangiocarcinoma (CHOL); colon adenocarcinoma (COAD); head and neck squamous cell carcinoma (HNSC); kidney renal clear cell carcinoma (KIRC); kidney renal papillary cell carcinoma (KIRP); LIHC; rectum adenocarcinoma (READ); stomach adenocarcinoma (STAD); uterine corpus endometrial carcinoma (UCEC). Interestingly, downregulation of POU5F1 was only observed in kidney chromophobe (KICH). Further, survival analyses were carried out through GEPIA based on TCGA. Whereas, among the above-mentioned cancers, only LIHC showed statistically significant differences in both OS and DFS (Fig. 2b, Additional file 9). Hence, LIHC was selected as the main target to explore the underlying function role of POU5F1 played in.

Association between POU5F1 and clinicopathological variables of LIHC

The expression profiles and clinical characteristics of 374 LIHC patients were obtained from TCGA to probe into the relationship between POU5F1 expression status and clinicopathological characters. As shown in Fig. 3a-h, elevated POU5F1 was associated with tumor occurrence ($P < 0.001$), advanced histological grade ($P = 0.016$), stage ($P = 0.025$), tumor invasion depth ($P = 0.019$), and distant metastasis ($P = 0.018$). Logistic regression analysis indicated the expression of POU5F1 as a risk factor that was associated with poor prognostic clinicopathologic variables (Table 3). Increased POU5F1 was significantly correlated with tumor occurrence (OR = 65.63, $P < 0.001$), advanced stage (OR = 2.06, $P = 0.007$) and tumor invasion depth (OR = 2.00, $P = 0.001$). Besides, POU5F1 (HR = 1.64, $P = 0.038$) was identified as an independent risk factor for OS of LIHC through multivariate analysis, as were tumor stage (HR = 1.51, $P < 0.001$), tumor invasion depth (HR = 1.51, $P < 0.001$) and distant metastasis (HR = 3.73, $P = 0.026$) (Table 4).

Diagnostic value of POU5F1 in plasma

We detected the expression level of POU5F1 in plasma collected from 30 LIHC patients and 30 normal controls by qPCR to investigate the diagnostic value of POU5F1. The main clinical characters of enrolled subjects were listed in Table 5. Significantly higher alanine aminotransferase (ALT) ($P < 0.001$), aspartate aminotransferase (AST) ($P < 0.001$), γ -glutamyl transferase (GGT) ($P = 0.047$), AFP ($P < 0.001$) and glucose (GLU) ($P < 0.001$) were observed in LIHC patients. On the contrary, albumin (ALB) was much lower in LIHC patients compared with normal controls ($P < 0.001$). The results of qPCR revealed POU5F1 was upregulated in plasma of LIHC patients, which was consistent with the results from liver tissue samples based on TCGA (Fig. 3i). Moreover, elevated POU5F1 was associated with high level of ALT in plasma ($P < 0.001$) (Table 6). ROC analysis was utilized to assess the diagnostic value of POU5F1 in LIHC. As displayed in Fig. 3j, the predictive validity of POU5F1 (AUC = 0.790, Se = 73.3%, Sp = 80.0%, $P < 0.001$) was higher than AFP (AUC = 0.766, Se = 63.3%, Sp = 100.0%, $P < 0.001$). Encouragingly, the diagnostic validity was remarkably improved through the combination of POU5F1 and AFP (AUC = 0.902, Se = 83.3%, Sp = 80.0%, $P < 0.001$).

Relationship between POU5F1 and TIICs

In order to inquire into the mechanism of POU5F1 involved in the pathological progress of LIHC, we analyzed the correlation between POU5F1 expression and 22 types of TIICs through CIBERSORT algorithm on the basis of expression profiles from TCGA. As exhibited in Fig. 4a, T cells CD4 memory resting ($P = 0.019$), T cells regulatory ($P = 0.048$), macrophage M1 ($P = 0.012$) and dendritic cells resting ($P = 0.002$) increased in the high POU5F1 group of normal liver tissues, while T cells follicular helper ($P = 0.031$) decreased. In LIHC tumor tissues, B cells memory ($P = 0.001$) and T cells follicular helper ($P = 0.007$) were enriched in the high POU5F1 group, B cells naive ($P < 0.001$), monocytes ($P = 0.004$) and dendritic cells activated ($P = 0.006$) increased in the low POU5F1 group (Fig. 4b). Besides, B cells memory ($P < 0.001$), T cells follicular helper ($P = 0.039$) and dendritic cells activated ($P < 0.001$) were positively related with POU5F1 in LIHC tumor tissues (Fig. 4c-e). The anomalous correlation between POU5F1 and dendritic cells activated might be partially explained by the limited data from dendritic cells activated. Negative correlations were observed in B cells naive ($P < 0.001$) and monocytes ($P = 0.011$) with POU5F1 (Fig. 4f, g).

Identification of POU5F1 related pathways

POU5F1 related signaling pathways were analyzed through GSEA to identify pathways that were differentially activated in LIHC between low and high POU5F1 expression phenotypes. GO terms enriched in high POU5F1 phenotype mainly contained DNA replication, regulation of cell cycle G2 M phase transition, signal transduction by p53 class mediator and so on. GO terms including acute phase response and complement activation alternative pathway were enriched in low POU5F1 phenotype (Fig. 5a). Multiple cancers related KEGG pathways were enriched in the high POU5F1 phenotype, such as bladder cancer, colorectal cancer, non-small cell lung cancer and renal cell carcinoma. Several well-known cancers related signaling pathways were also enriched in the high POU5F1 phenotype, including

MTOR signaling pathway, p53 signaling pathway and WNT signaling pathway. While, PPAR signaling pathway were enriched in the low POU5F1 phenotype (Fig. 5b).

Enrichment analysis of POU5F1 co-expression genes

To explore genes that might potentially associated with POU5F1, co-expression analysis was performed and the expression status of the top 50 genes were displayed in Fig. 5c. GO functional enrichment analysis indicated these genes were enriched in cell proliferation-related terms, including chromosome segregation, nuclear division, organelle fission, DNA replication and so on (Fig. 5d). Cell cycle, spliceosome, p53 signaling pathway, pancreatic cancer and DNA replication were the main signaling pathways in which these POU5F1 co-expression genes enriched through KEGG pathway analysis (Fig. 5e).

PPI network construction and hub genes recognition

PPI network was constructed to reveal the intrinsic correlations among the POU5F1 co-expression genes. As exhibited in Fig. 6a, deeper color of each gene circle indicated increased correlation coefficient with POU5F1. Analogously, larger circle size indicated smaller *P* value. 3 genes (CBX3, CCHCR1, NFYC) were found to be directly associated with POU5F1 and were defined as central hub genes. BARD1, ZNF692, IQCC, FBXL19, GPD2 and KAT2A had direct connections with the central hub genes and were regarded as subordinate hub genes for POU5F1. The expression status of the central hub genes and subordinate hub genes were all positively correlated with the expression level of POU5F1 in LIHC on the basis of TCGA (Fig. 6b). In addition to KAT2A, shorter OS of LIHC was found to be correlated with the overexpression of all the hub genes. Elevated expression of all the hub genes except NFYC indicated poor DFS of LIHC (Fig. 7). Besides, the 9 hub genes were all prominently upregulated in LIHC patients based on TCGA (Fig. 8a).

Discussion

POU5F1 has been studied for a long period of time as a well-known CSCs marker, which participates in tumor invasion, differentiation and recurrence [67]. A growing number of researches have suggested the prognostic value of POU5F1 in various malignancies. But due to the limitation of sample size and methodology, the conclusions drawn by individual studies may be unauthentic to demonstrate the prognostic validity of POU5F1. We performed a meta-analysis that incorporated 16 types of cancers with 7401 subjects from 57 studies to come to more reliable conclusions. The amalgamative results indicated elevated POU5F1 was associated with poor OS, DFS, DSS and RFS in various cancers. Especially, TSA confirmed the sample size of current studies has far exceeded the APIS, suggesting it was quite credible to draw a conclusion that elevated POU5F1 was apparently connected with shorter OS in various cancers. Besides, the pooled estimates of clinicopathological parameters suggested POU5F1 played pivotal roles in tumorigenesis, tumor growth, invasion, metastasis and therapy resistance in multiple cancers. These results indicated POU5F1 might serve as a prognostic pan-cancer biomarker and potential therapy target.

POU5F1 was upregulated in BLCA, BRCA, CHOL, COAD, HNSC, KIRP, LIHC, READ and STAD based on TCGA, which was consistent with the meta-analysis results. Interestingly, differences in both OS and DFS between the high POU5F1 group and the low POU5F1 group were observed only in LIHC on the basis of TCGA, indicating POU5F1 played a unique and important role in the prognosis of LIHC. Further, DNA replication, regulation of cell cycle G2 M phase transition, bladder cancer, colorectal cancer, non-small cell lung cancer, renal cell carcinoma, MTOR signaling pathway, p53 signaling pathway and WNT signaling pathway were the main GO and KEGG terms enriched in the high POU5F1 phenotype according to the GSEA. The GSEA results suggested POU5F1 might participate in the pathological progress of LIHC and other cancers through promoting cell proliferation. Similar GO terms and KEGG pathways were found in the co-expression genes of POU5F1 and further validated the GSEA results.

Previous studies reported TILCs could independently predict the OS among cancer patients and reflect the status of lymph node [75]. Our study found there was a prominent decrease in B cells naive and increase in B cells memory in the high POU5F1 group of LIHC tumor tissues, hinting the elevated POU5F1 might promote the transformation of B cells naive into B cells memory in LIHC. T cells follicular helper decreased in the high POU5F1 group in normal liver tissues, but increased in the high POU5F1 group in LIHC tumor tissues. And the exact opposite results were found in dendritic cells activated. The typing and quantity conversion of TILCs in normal tissues and tumor tissues indicated the significant meanings of POU5F1 in regulating tumor immune microenvironment of LIHC. The mechanism of POU5F1 participates in the regulation of tumor immune microenvironment still needs further study.

We identified upregulated POU5F1 as an independent prognostic factor for poor prognosis of LIHC through cox regression, along with tumor stage, invasion depth and distant metastasis. Overexpression of POU5F1 was related with high level of ALT in plasma. In consideration of the high concentration of ALT was the indicative of liver cells destruction, we speculated the overexpression of POU5F1 might be associated with hepatocellular necrosis or apoptosis [76]. In addition, although the diagnostic value of POU5F1 in LIHC was quite gratifying, the necessity of applying POU5F1 and AFP together in the diagnosis of LIHC to improve the diagnostic specificity needed to be emphasized, in view of POU5F1 was upregulated in a variety of cancers and might reduce the diagnostic specificity.

The molecular regulation mechanisms and pathways by which POU5F1 participates in LIHC have not been thoroughly studied. To further explore the role of POU5F1 played in LIHC, we constructed PPI network using co-expression genes of POU5F1 and identified hub genes that interacted with POU5F1, including CBX3, CCHCR1, NFYC, BARD1, ZNF692, IQCC, FBXL19, GPD2 and KAT2A. Based on the related studies of hub genes, we visualized the pathways POU5F1 might play a role in LIHC (Fig. 8b). It has been reported that the transcription factor complex of POU5F1, SOX2 and KLF4 binds to Nanog promoter to induce cellular reprogramming and cancer stemness [77]. EpiCD translocates to the nucleus in a multiprotein complex and enhances the expression of POU5F1 by binding to the promoter of POU5F1 [78]. CBX3 has been confirmed to promote cell cycle transition by inducing CDK1 and PCNA [79]. The elevated POU5F1 in LIHC may influence the expression of CBX3 and then activate NF- κ B and PI3K/Akt pathway through BARD1 and ZNF692 [80, 81]. The promotion effect of CCND2 on cell proliferation is regulated by NFYC and may also be affected by POU5F1 [82]. The interaction between NFYC and KAT2A indicates POU5F1 participates tumor development through KAT2A mediated histone H3 succinylation [83]. GPD2 promotes HuH-7 cell mitochondrial energy metabolism which may be regulated by NFYC and POU5F1 [84]. As a central hub gene of POU5F1, CCHCR1 accelerates cell proliferation through EGFR [85]. FBXL19 induces Rac1 and Rac3 expression and inhibits apoptosis [86]. The relationship among FBXL19, CCHCR1 and POU5F1 needs further verification. Besides, loss-of-function of POU5F1 remarkably restrains propagation, metastasis and aggression of cancer stem cells through inhibition of the PI3K/Akt pathway, from which we could expect POU5F1 to be an underlying target for cancer therapy [87].

Conclusions

In summary, our study identified POU5F1 as a pan-cancer gene with significant prognostic value in various cancers, especially in LIHC. POU5F1 can serve as an independent prognostic factor for LIHC and the combination of AFP and POU5F1 in plasma has prominent diagnostic validity for LIHC. POU5F1 may influence the progression of LIHC by regulating the tumor immune microenvironment and participating in cell proliferation-related pathways. Further research should be performed to verify the function mechanism of POU5F1 in the pathogenesis of LIHC.

Abbreviations

POU5F1: POU Class 5 Homeobox 1; LIHC:liver hepatocellular carcinoma; TILs:tumor infiltrating immune cells; GSEA:Gene set enrichment analysis; AFP:α-fetoprotein; CSCs:cancer stem cells; TSA:trial sequential analysis; PPI:protein-protein interaction; HRs:hazard ratios; PRISMA:preferred reporting items for systematic reviews and meta-analyses; OS:overall survival; DFS:disease-free survival; DSS:disease-specific survival; RFS:recurrence-free survival; NOS:Newcastle-Ottawa Scale; APIS:a priori information size; RRR:relative risk reduction; TIMER:Tumor Immune Estimation Resource; GEPIA:Gene Expression Profiling Interactive Analysis; cDNA:complementary DNA; qPCR:quantitative PCR; GAPDH:Glyceraldehyde 3-phosphate dehydrogenase; Cq:quantification cycle; GO:Gene ontology; KEGG:Kyoto Encyclopedia of Genes and Genomes; NES:normalised enrichment score; STRING:Search Tool for the Retrieval of Interacting Genes; PPI:protein-protein interaction; SD:standard deviation; ORs:Odds ratios; ROC:Receiver operation curve; IHC:immunohistochemistry; IF:immunofluorescence; BLCA:bladder urothelial carcinoma; BRCA:breast invasive carcinoma; CHOL:cholangiocarcinoma; COAD:colon adenocarcinoma; HNSC:head and neck squamous cell carcinoma; KIRC:kidney renal clear cell carcinoma; KIRP:kidney renal papillary cell carcinoma; READ:rectum adenocarcinoma; STAD:stomach adenocarcinoma; UCEC:uterine corpus endometrial carcinoma; KICH:kidney chromophobe; ALT:alanine aminotransferase; AST:aspartate aminotransferase; GGT:γ-glutamyl transferase; GLU:glucose; ALB:albumin; NR:not reported; M:male; F:female; RT-PCR:reverse transcription polymerase chain reaction.

Declarations

Ethics approval and consent to participate

All experimental schemes were approved by the Ethics Committee of Zhongnan Hospital of Wuhan University.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The study was funded by National Basic Research Program of China (2012CB720605), Zhongnan Hospital of Wuhan University Science, Technology and Innovation Seed Fund (ZNPY2017054).

Authors' contributions

JCT conceived and designed the workflow. DDH and XKZ performed the experiments and analyzed the data. DDH and XKZ wrote the manuscript. JCT revised the manuscript. All authors approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Wu S, Powers S, Zhu W, Hannun YA. Substantial contribution of extrinsic risk factors to cancer development. *Nature*. 2016;529:43–7.
2. Russnes HG, Lønning PE, Børresen-Dale AL, Lingjærde OC. The multitude of molecular analyses in cancer: the opening of Pandora's box. *Genome biology*. 2014;15:447.
3. Siegel RL, Miller KD, Jemal A. Cancer Statistics. 2020. *CA Cancer J Clin*. 2020;70:7–30.
4. Kim SI, Koo JS. Expression of cancer stem cell markers in breast phyllodes tumor. *Cancer Biomark*. 2020;10.
5. Gudbergsson JM, Christensen E, Kostrikov S, et al. Conventional Treatment of Glioblastoma Reveals Persistent CD44 + Subpopulations. *Mol Neurobiol*. 2020;10.1007/s12035-020-02004-2.
6. Gzil A, Zarębska I, Jaworski D, et al. The prognostic value of leucine-rich repeat-containing G-protein (Lgr5) and its impact on clinicopathological features of colorectal cancer. *J Cancer Res Clin Oncol*. 2020;10.1007/s00432-020-03314-7.
7. Zhang WJ, Zhou ZH, Guo M, Yang LQ, Xu YY, Pang TH, et al. High Infiltration of Polarized CD163(+) Tumor-Associated Macrophages Correlates with Aberrant Expressions of CSCs Markers, and Predicts Prognosis in Patients with Recurrent Gastric Cancer. *J Cancer*. 2017;8:363–70.
8. Wang Q, He W, Lu C, Wang Z, Wang J, Giercksky KE, et al. Oct3/4 and Sox2 are significantly associated with an unfavorable clinical outcome in human esophageal squamous cell carcinoma. *Anticancer research*. 2009;29:1233–41.
9. Cheng L, Sung MT, Cossu-Rocca P, Jones TD, MacLennan GT, De Jong J, et al. OCT4: biological functions and clinical applications as a marker of germ cell neoplasia. *J Pathol*. 2007;211:1–9.
10. Li C, Yan Y, Ji W, Bao L, Qian H, Chen L, et al. OCT4 positively regulates Survivin expression to promote cancer cell proliferation and leads to poor prognosis in esophageal squamous cell carcinoma. *PloS one*. 2012;7:e49693.
11. Cao L, Li C, Shen S, Yan Y, Ji W, Wang J, et al. OCT4 increases BIRC5 and CCND1 expression and promotes cancer progression in hepatocellular carcinoma. *BMC Cancer*. 2013;13:82.
12. Chang CC, Shieh GS, Wu P, Lin CC, Shiao AL, Wu CL. Oct-3/4 expression reflects tumor progression and regulates motility of bladder cancer cells. *Cancer research*. 2008;68:6281–91.
13. Chen YC, Hsu HS, Chen YW, Tsai TH, How CK, Wang CY, et al. Oct-4 expression maintained cancer stem-like properties in lung cancer-derived CD133-positive cells. *PloS one*. 2008;3:e2637.
14. Chiou SH, Yu CC, Huang CY, Lin SC, Liu CJ, Tsai TH, et al. Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2008;14:4085–95.
15. He W, Li K, Wang F, Qin YR, Fan QX. Expression of OCT4 in human esophageal squamous cell carcinoma is significantly associated with poorer prognosis. *World journal of gastroenterology*. 2012;18:712–9.
16. Ge N, Lin HX, Xiao XS, Guo L, Xu HM, Wang X, et al. Prognostic significance of Oct4 and Sox2 expression in hypopharyngeal squamous cell carcinoma. *Journal of translational medicine*. 2010;8:94.
17. Qian YW, Chen Y, Yang W, Fu J, Cao J, Ren YB, et al. p28(GANK) prevents degradation of Oct4 and promotes expansion of tumor-initiating cells in hepatocarcinogenesis. *Gastroenterology*. 2012;142:1547–58.e14.

18. Huang P, Qiu J, Li B, Hong J, Lu C, Wang L, et al. Role of Sox2 and Oct4 in predicting survival of hepatocellular carcinoma patients after hepatectomy. *Clinical biochemistry*. 2011;44:582–9.
19. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151:264–9, w64.
20. Zeng X, Zhang Y, Kwong JS, Zhang C, Li S, Sun F, et al. The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. *Journal of evidence-based medicine*. 2015;8:2–10.
21. Holst LB, Petersen MW, Haase N, Perner A, Wetterslev J. Restrictive versus liberal transfusion strategy for red blood cell transfusion: systematic review of randomised trials with meta-analysis and trial sequential analysis. *BMJ*. 2015;350:h1354.
22. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer research*. 2017;77:e108-e10.
23. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic acids research*. 2017;45:W98-w102.
24. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nature medicine*. 2015;21:938–45.
25. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102:15545–50.
26. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials*. 2007;8:16.
27. Cortes-Dericks L, Galetta D, Spaggiari L, Schmid RA, Karoubi G. High expression of octamer-binding transcription factor 4A, prominin-1 and aldehyde dehydrogenase strongly indicates involvement in the initiation of lung adenocarcinoma resulting in shorter disease-free intervals. *European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery*. 2012;41:e173-81.
28. Ravindran G, Sawant SS, Hague A, Kingsley K, Devaraj H. Association of differential β -catenin expression with Oct-4 and Nanog in oral squamous cell carcinoma and their correlation with clinicopathological factors and prognosis. *Head Neck*. 2015;37:982–93.
29. Chang TS, Wu YC, Chi CC, Su WC, Chang PJ, Lee KF, et al. Activation of IL6/IGFIR confers poor prognosis of HBV-related hepatocellular carcinoma through induction of OCT4/NANOG expression. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2015;21:201–10.
30. Chiou SH, Wang ML, Chou YT, Chen CJ, Hong CF, Hsieh WJ, et al. Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell-like properties and epithelial-mesenchymal transdifferentiation. *Cancer research*. 2010;70:10433–44.
31. Comisso E, Scarola M, Rosso M, Piazza S, Marzinotto S, Ciani Y, et al. OCT4 controls mitotic stability and inactivates the RB tumor suppressor pathway to enhance ovarian cancer aggressiveness. *Oncogene*. 2017;36:4253–66.
32. Dong Z, Zeng Q, Luo H, Zou J, Cao C, Liang J, et al. Increased expression of OCT4 is associated with low differentiation and tumor recurrence in human hepatocellular carcinoma. *Pathol Res Pract*. 2012;208:527–33.
33. Gwak JM, Kim M, Kim HJ, Jang MH, Park SY. Expression of embryonal stem cell transcription factors in breast cancer: Oct4 as an indicator for poor clinical outcome and tamoxifen resistance. *Oncotarget*. 2017;8:36305–18.

34. Hu J, Li J, Yue X, Wang J, Liu J, Sun L, et al. Expression of the cancer stem cell markers ABCG2 and OCT-4 in right-sided colon cancer predicts recurrence and poor outcomes. *Oncotarget*. 2017;8:28463–70.
35. Huang P, Chen J, Wang L, Na Y, Kaku H, Ueki H, et al. Implications of transcriptional factor, OCT-4, in human bladder malignancy and tumor recurrence. 29. London: Medical oncology (Northwood; 2012. pp. 829–34.
36. Javanbakht M, Akhavanmoghadam J, Talaei AJ, Aghyani M, Mozafari M, Khedmat L, et al. Differential expression of two genes Oct-4 and MUC5AC associates with poor outcome in patients with gastric cancer. *Clin Exp Pharmacol Physiol*. 2017;44:1099–105.
37. Jen J, Tang YA, Lu YH, Lin CC, Lai WW, Wang YC. Oct4 transcriptionally regulates the expression of long non-coding RNAs NEAT1 and MALAT1 to promote lung cancer progression. *Mol Cancer*. 2017;16:104.
38. Jiang WL, Zhang PF, Li GF, Dong JH, Wang XS, Wang YY. Oct-4 is associated with gastric cancer progression and prognosis. *OncoTargets therapy*. 2016;9:517–22.
39. Kaneko Y, Suenaga Y, Islam SM, Matsumoto D, Nakamura Y, Ohira M, et al. Functional interplay between MYCN, NCYM, and OCT4 promotes aggressiveness of human neuroblastomas. *Cancer Sci*. 2015;106:840–7.
40. Kim BW, Cho H, Choi CH, Ylaya K, Chung JY, Kim JH, et al. Clinical significance of OCT4 and SOX2 protein expression in cervical cancer. *BMC Cancer*. 2015;15:1015.
41. Kim K, Ro JY, Kim S, Cho YM. Expression of stem-cell markers OCT-4 and CD133: important prognostic factors in papillary renal cell carcinoma. *Human pathology*. 2012;43:2109–16.
42. Kong D, Su G, Zha L, Zhang H, Xiang J, Xu W, et al. Coexpression of HMGA2 and Oct4 predicts an unfavorable prognosis in human gastric cancer. 31. London: Medical oncology (Northwood; 2014. p. 130.
43. Kosaka T, Mikami S, Yoshimine S, Miyazaki Y, Daimon T, Kikuchi E, et al. The prognostic significance of OCT4 expression in patients with prostate cancer. *Human pathology*. 2016;51:1–8.
44. Li C, Zhu M, Lou X, Liu C, Chen H, Lin X, et al. Transcriptional factor OCT4 promotes esophageal cancer metastasis by inducing epithelial-mesenchymal transition through VEGF-C/VEGFR-3 signaling pathway. *Oncotarget*. 2017;8:71933–45.
45. Li N, Deng W, Ma J, Wei B, Guo K, Shen W, et al. Prognostic evaluation of Nanog, Oct4, Sox2, PCNA, Ki67 and E-cadherin expression in gastric cancer. 32. London: Medical oncology (Northwood; 2015. p. 433.
46. Li X, Wang J, Xu Z, Ahmad A, Li E, Wang Y, et al. Expression of Sox2 and Oct4 and their clinical significance in human non-small-cell lung cancer. *Int J Mol Sci*. 2012;13:7663–75.
47. Li XL, Jia LL, Shi MM, Li X, Li ZH, Li HF, et al. Downregulation of KPNA2 in non-small-cell lung cancer is associated with Oct4 expression. *Journal of translational medicine*. 2013;11:232.
48. Liu C, Cao X, Zhang Y, Xu H, Zhang R, Wu Y, et al. Co-expression of Oct-4 and Nestin in human breast cancers. *Molecular biology reports*. 2012;39:5875–81.
49. Liu CG, Lu Y, Wang BB, Zhang YJ, Zhang RS, Lu Y, et al. Clinical implications of stem cell gene Oct-4 expression in breast cancer. *Annals of surgery*. 2011;253:1165–71.
50. Liu T, Sun B, Zhao X, Li Y, Gu Q, Dong X, et al. OCT4 expression and vasculogenic mimicry formation positively correlate with poor prognosis in human breast cancer. *Int J Mol Sci*. 2014;15:19634–49.
51. Lu Y, Zhu H, Shan H, Lu J, Chang X, Li X, et al. Knockdown of Oct4 and Nanog expression inhibits the stemness of pancreatic cancer cells. *Cancer letters*. 2013;340:113–23.
52. Luo W, Li S, Peng B, Ye Y, Deng X, Yao K. Embryonic stem cells markers SOX2, OCT4 and Nanog expression and their correlations with epithelial-mesenchymal transition in nasopharyngeal carcinoma. *PloS one*. 2013;8:e56324.
53. Matsuoka J, Yashiro M, Sakurai K, Kubo N, Tanaka H, Muguruma K, et al. Role of the stemness factors sox2, oct3/4, and nanog in gastric carcinoma. *J Surg Res*. 2012;174:130–5.

54. Miyoshi N, Fujino S, Ohue M, Yasui M, Takahashi Y, Sugimura K, et al. The POU5F1 gene expression in colorectal cancer: a novel prognostic marker. *Surg Today*. 2018;48:709–15.
55. Sawant S, Gokulan R, Dongre H, Vaidya M, Chaukar D, Prabhash K, et al. Prognostic role of Oct4, CD44 and c-Myc in radio-chemo-resistant oral cancer patients and their tumourigenic potential in immunodeficient mice. *Clinical oral investigations*. 2016;20:43–56.
56. Tang YA, Chen CH, Sun HS, Cheng CP, Tseng VS, Hsu HS, et al. Global Oct4 target gene analysis reveals novel downstream PTEN and TNC genes required for drug-resistance and metastasis in lung cancer. *Nucleic acids research*. 2015;43:1593–608.
57. Wang D, Lu P, Zhang H, Luo M, Zhang X, Wei X, et al. Oct-4 and Nanog promote the epithelial-mesenchymal transition of breast cancer stem cells and are associated with poor prognosis in breast cancer patients. *Oncotarget*. 2014;5:10803–15.
58. Wang G, Zhou H, Gu Z, Gao Q, Shen G. Oct4 promotes cancer cell proliferation and migration and leads to poor prognosis associated with the survivin/STAT3 pathway in hepatocellular carcinoma. *Oncol Rep*. 2018;40:979–87.
59. Wang QH, Zhang M, Shi CT, Xie JJ, Chen F, Shi QF, et al. High Oct4 predicted worse prognosis of right-sided colon cancer patients. *Future oncology (London England)*. 2018;14:2279–91.
60. Xiang Y, Zhou X. Octamer-binding transcription factor 4 correlates with complex karyotype, FLT3-ITD mutation and poorer risk stratification, and predicts unfavourable prognosis in patients with acute myeloid leukaemia. *Hematology (Amsterdam Netherlands)*. 2018;23:721–8.
61. Xin YH, Bian BS, Yang XJ, Cui W, Cui HJ, Cui YH, et al. POU5F1 enhances the invasiveness of cancer stem-like cells in lung adenocarcinoma by upregulation of MMP-2 expression. *PloS one*. 2013;8:e83373.
62. Xing CG, Lu XG, Zhang YS, Zhou F, Xu XP. Expression of embryonic stem cell marker Oct-4 and its prognostic significance in rectal adenocarcinoma. *Chin J Cancer Res*. 2010;22:106–11.
63. Yang Y, Wang Y, Yin C, Li X. Clinical significance of the stem cell gene Oct-4 in cervical cancer. *Tumour biology: the journal of the International Society for Oncodevelopmental Biology Medicine*. 2014;35:5339–45.
64. Yin JY, Tang Q, Zhai LL, Zhou LY, Qian J, Lin J, et al. High expression of OCT4 is frequent and may cause undesirable treatment outcomes in patients with acute myeloid leukemia. *Tumour biology: the journal of the International Society for Oncodevelopmental Biology Medicine*. 2015;36:9711–6.
65. Yin X, Li YW, Jin JJ, Zhou Y, Ren ZG, Qiu SJ, et al. The clinical and prognostic implications of pluripotent stem cell gene expression in hepatocellular carcinoma. *Oncology letters*. 2013;5:1155–62.
66. Yin X, Li YW, Zhang BH, Ren ZG, Qiu SJ, Yi Y, et al. Coexpression of stemness factors Oct4 and Nanog predict liver resection. *Ann Surg Oncol*. 2012;19:2877–87.
67. You L, Guo X, Huang Y. Correlation of Cancer Stem-Cell Markers OCT4, SOX2, and NANOG with Clinicopathological Features and Prognosis in Operative Patients with Rectal Cancer. *Yonsei Med J*. 2018;59:35–42.
68. Zhang JM, Wei K, Jiang M. OCT4 but not SOX2 expression correlates with worse prognosis in surgical patients with triple-negative breast cancer. *Breast cancer (Tokyo Japan)*. 2018;25:447–55.
69. Zhang X, Han B, Huang J, Zheng B, Geng Q, Aziz F, et al. Prognostic significance of OCT4 expression in adenocarcinoma of the lung. *Jpn J Clin Oncol*. 2010;40:961–6.
70. Zhao RC, Zhou J, Chen KF, Gong J, Liu J, He JY, et al. The prognostic value of combination of CD90 and OCT4 for hepatocellular carcinoma after curative resection. *Neoplasma*. 2016;63:288–98.
71. Zhao Y, Li C, Huang L, Niu S, Lu Q, Gong D, et al. Prognostic value of association of OCT4 with LEF1 expression in esophageal squamous cell carcinoma and their impact on epithelial-mesenchymal transition, invasion, and migration. *Cancer medicine*. 2018;7:3977–87.

72. Zhou H, Hu YU, Wang W, Mao Y, Zhu J, Zhou B, et al. Expression of Oct-4 is significantly associated with the development and prognosis of colorectal cancer. *Oncology letters*. 2015;10:691–6.
73. Zhou J, Dong D, Cheng R, Wang Y, Jiang S, Zhu Y, et al. Aberrant expression of KPNA2 is associated with a poor prognosis and contributes to OCT4 nuclear transportation in bladder cancer. *Oncotarget*. 2016;7:72767–76.
74. Zou Q, Yang L, Yang Z, Huang J, Fu X. PSCA and Oct-4 expression in the benign and malignant lesions of gallbladder: implication for carcinogenesis, progression, and prognosis of gallbladder adenocarcinoma. *BioMed research international*. 2013;2013:648420.
75. Azimi F, Scolyer RA, Rumcheva P, Moncrieff M, Murali R, McCarthy SW, et al. Tumor-infiltrating lymphocyte grade is an independent predictor of sentinel lymph node status and survival in patients with cutaneous melanoma. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*. 2012;30:2678–83.
76. Senior JR. Alanine aminotransferase: a clinical and regulatory tool for detecting liver injury-past, present, and future. *Clin Pharmacol Ther*. 2012;92:332–9.
77. Lee S, Wottrich S, Bonavida B. Crosstalks between Raf-kinase inhibitor protein and cancer stem cell transcription factors (Oct4, KLF4, Sox2, Nanog). *Tumour biology: the journal of the International Society for Oncodevelopmental Biology Medicine*. 2017;39:1010428317692253.
78. Oishi N, Yamashita T, Kaneko S. Molecular biology of liver cancer stem cells. *Liver cancer*. 2014;3:71–84.
79. Chen LY, Cheng CS, Qu C, Wang P, Chen H, Meng ZQ, et al. Overexpression of CBX3 in Pancreatic Adenocarcinoma Promotes Cell Cycle Transition-Associated Tumor Progression. *International journal of molecular sciences*. 2018;19(6).
80. Cimmino F, Formicola D, Capasso M. Dualistic Role of BARD1 in Cancer. *Genes*. 2017;8.
81. Xing Y, Ren S, Ai L, Sun W, Zhao Z, Jiang F, et al. ZNF692 promotes colon adenocarcinoma cell growth and metastasis by activating the PI3K/AKT pathway. *Int J Oncol*. 2019;54:1691–703.
82. Steiman-Shimony A, Shtrikman O, Margalit H. Assessing the functional association of intronic miRNAs with their host genes. *RNA (New York, NY)*. 2018;24:991–1004.
83. Wang Y, Guo YR, Liu K, Yin Z, Liu R, Xia Y, et al. KAT2A coupled with the α -KGDH complex acts as a histone H3 succinyltransferase. *Nature*. 2017;552:273–7.
84. Mikeli M, Fujikawa M, Nagahisa K, Yasuda S, Yamada N, Tanabe T. Contribution of GPD2/mGPDH to an alternative respiratory chain of the mitochondrial energy metabolism and the stemness in CD133-positive HuH-7 cells. *Genes to cells: devoted to molecular cellular mechanisms*. 2020;25:139–48.
85. Suomela S, Elomaa O, Skoog T, Ala-aho R, Jeskanen L, Pärssinen J, et al. CCHCR1 is up-regulated in skin cancer and associated with EGFR expression. *PloS one*. 2009;4:e6030.
86. Cai J, Culley MK, Zhao Y, Zhao J. The role of ubiquitination and deubiquitination in the regulation of cell junctions. *Protein cell*. 2018;9:754–69.
87. Lin H, Sun LH, Han W, He TY, Xu XJ, Cheng K, et al. Knockdown of OCT4 suppresses the growth and invasion of pancreatic cancer cells through inhibition of the AKT pathway. *Mol Med Rep*. 2014;10:1335–42.

Tables

Table 1 Subgroup analyses on pooled HRs of POU5F1 for OS

Categories	No. of studies	No. of patients	Pooled HR (95% CI)	Significant z	P value	Heterogeneity I ² (%)	P value	Model
[1] OS	48	5485	2.45 (2.22-2.71)	17.74	0.000	4.3	0.389	Fixed
[2] Cancer type								
1) Head and neck cancer	4	321	2.44 (1.48-4.02)	3.51	0.000	19.1	0.295	Fixed
2) Esophageal cancer	5	450	2.49 (1.81-3.45)	5.54	0.000	48.9	0.098	Fixed
3) Breast cancer	3	343	3.83 (2.57-5.70)	6.60	0.000	0.0	0.723	Fixed
4) Lung cancer	7	706	2.39 (1.86-3.07)	6.78	0.000	0.0	0.935	Fixed
5) Gastric cancer	5	969	2.02 (1.53-2.67)	4.93	0.000	13.8	0.326	Fixed
6) Hepatocellular cancer	9	1082	2.39 (1.94-2.95)	8.10	0.000	1.6	0.421	Fixed
7) Colorectal cancer	6	752	2.31 (1.82-2.94)	6.80	0.000	19.7	0.285	Fixed
8) Bladder cancer	3	360	2.97 (2.04-4.33)	5.67	0.000	0.0	0.715	Fixed
9) Acute myeloid leukemia	2	239	2.14 (0.69-6.66)	1.31	0.190	79.7	0.027	Random
10) Other cancers	4	263	3.08 (1.72-5.52)	3.79	0.000	0.0	0.654	Fixed
[3] Analysis type								
1) Multivariate	28	3405	2.61 (2.28-2.99)	13.91	0.000	1.6	0.440	Fixed
2) Univariate	20	2080	2.28 (1.97-2.64)	11.10	0.000	4.4	0.402	Fixed
[4] Sample size								
1) ≥110	22	3713	2.40 (2.12-2.72)	13.83	0.000	6.0	0.380	Fixed

2)	<110	26	1772	2.54 (2.15- 2.99)	11.12	0.000	5.6	0.382	Fixed
[5]	Detection method								
1)	IHC	36	4197	2.35 (2.10- 2.63)	14.91	0.000	0.0	0.620	Fixed
2)	RT-PCR	9	901	2.58 (1.99- 3.34)	7.21	0.000	19.2	0.272	Fixed
3)	IF	3	387	3.47 (2.41- 5.00)	6.68	0.000	35.4	0.213	Fixed

Table 2 Pooled ORs for the correlation between elevated POU5F1 and clinicopathological characteristics

Clinicopathological parameters	No. of studies	No. of patients	Risk of high POU5F1 OR (95% CI)	Significant z	P value	Heterogeneity I ² (%)	P value	Model
Age (≥60 vs <60)	16	1694	1.08 (0.88-1.32)	0.69	0.489	0.0	0.768	Fixed
Gender (Male vs Female)	35	3850	1.05 (0.90-1.23)	0.65	0.517	0.0	0.844	Fixed
Tumor size (≥5 cm vs <5 cm)	14	1967	1.38 (1.13-1.68)	3.21	0.001	42.2	0.048	Fixed
TNM stage (III-IV vs I-II)	22	2347	2.72 (2.23-3.31)	9.99	0.000	26.0	0.130	Fixed
Tumor differentiation (Well-Moderate vs Poor)	20	2632	3.08 (2.08-4.56)	5.62	0.000	67.0	0.000	Random
Tumor invasion depth (T3-T4 vs T1-T2)	15	1861	2.31 (1.82-2.93)	6.91	0.000	10.6	0.334	Fixed
Lymph node metastasis (Positive vs Negative)	25	3534	3.11 (2.66-3.63)	14.31	0.000	4.4	0.400	Fixed
Distant metastasis (Positive vs Negative)	10	1437	2.86 (1.96-4.19)	5.43	0.000	0.0	0.991	Fixed
Lymphovascular invasion (Positive vs Negative)	4	451	1.91 (1.25-2.94)	2.96	0.003	0.0	0.640	Fixed
Vascular invasion (Positive vs Negative)	6	727	2.34 (1.65-3.31)	4.80	0.000	11.1	0.345	Fixed
Tumor number (Multiple vs Single)	5	531	1.65 (1.06-2.55)	2.23	0.026	0.0	0.812	Fixed
Tumor Recurrence (Positive vs Negative)	5	546	5.05 (3.33-7.55)	7.62	0.000	31.5	0.212	Fixed
Tumor encapsulation (Incomplete vs Complete)	5	560	1.36 (0.95-1.94)	1.69	0.091	20.6	0.283	Fixed
Liver cirrhosis (Positive vs Negative)	6	712	1.01 (0.68-1.48)	0.03	0.979	0.0	0.817	Fixed

HBsAg (Positive vs Negative)	5	659	1.00 (0.62-1.61)	0.01	0.995	0.0	0.793	Fixed
Smoke (Yes vs No)	4	406	1.30 (0.80-2.10)	1.06	0.287	0.0	0.878	Fixed

Table 3 Correlations between elevated POU5F1 and clinicopathological characteristics in LIHC patients based on TCGA

Clinical characteristics	Total (N)	Risk of high POU5F1 OR (95% CI)	P value
Status (Tumor free vs With tumor)	421	65.63 (8.97 - 480.14)	<0.001
Age	370	1.00 (0.98 - 1.01)	0.820
Gender (Male vs Female)	371	1.07 (0.69 - 1.65)	0.767
Grade (III vs I)	177	1.89 (0.99 - 3.61)	0.054
Stage (III vs I)	256	2.06 (1.22 - 3.50)	0.007
T (III vs I)	261	2.00 (1.17 - 3.41)	0.011
N (Positive vs Negative)	256	1.10 (0.15 - 7.93)	0.925
M (Positive vs Negative)	270	0.26 (0.03 - 2.32)	0.225

Table 4 Univariate and multivariate analysis of OS in LIHC patients based on TCGA

Clinicopathologic variable	Univariate analysis	P value	Multivariate analysis	P value
	HR (95% CI)		HR (95% CI)	
Age	1.01 (1.00 - 1.03)	0.064		
Gender (Male vs Female)	1.18 (0.82 - 1.69)	0.380		
Grade (III vs I)	1.09 (0.82 - 1.44)	0.560		
Stage (III vs I)	1.63 (1.31 - 2.02)	0.000	1.51 (1.20 - 1.89)	< 0.001
T (III vs I)	1.61 (1.30 - 1.99)	0.000	1.51 (1.21 - 1.88)	< 0.001
N (Positive vs Negative)	1.94 (0.48 - 7.93)	0.355		
M (Positive vs Negative)	3.88 (1.22 - 12.35)	0.022	3.73 (1.17 - 11.87)	0.026
POU5F1 (High vs Low)	1.92 (1.34 - 2.76)	0.000	1.64 (1.03 - 2.62)	0.038

Table 5 The main clinical features of research subjects

Characteristics	Control (n = 30)	LIHC (n = 30)	P value
Gender			0.007
Male (%)	17 (56.67)	27 (90.00)	
Female (%)	13 (43.33)	3 (10.00)	
Age (y)			0.070
< 55 (%)	19 (63.33)	11 (36.67)	
≥ 55 (%)	11 (36.67)	19 (63.33)	
ALT (U/L)	18.00 (13.00 - 24.00)	43.00 (23.50 - 68.00)	<0.001
AST (U/L)	21.50 (18.00 - 27.00)	49.50 (31.25 - 83.00)	<0.001
ALP (U/L)	88.00 (73.25 - 157.00)	98.00 (78.50 - 220.00)	0.414
GGT (U/L)	24.50 (18.75 - 47.25)	34.50 (23.75 - 71.50)	0.047
TP (g/L)	69.20 (60.60 - 72.73)	63.05 (59.10 - 72.80)	0.232
ALB (g/L)	44.95 (42.83 - 46.75)	35.90 (33.08 - 38.50)	<0.001
CEA (ng/mL)	1.88 (1.23 - 2.55)	2.10 (1.50 - 3.11)	0.179
AFP (ng/mL)	2.77 (1.72 - 3.65)	34.83 (2.47 - 311.20)	<0.001
GLU (mmol/L)	5.08 (4.42 - 5.33)	5.86 (5.02 - 7.49)	<0.001

Table 6 Relationship between POU5F1 expression and clinical characteristics of LIHC patients

Characteristics	Patient number (n = 30)	Low expression (n = 15)	High expression (n = 15)	P value
Gender				0.999
Male	27	13 (48.15)	14 (51.85)	
Female	3	2 (66.67)	1 (33.33)	
Age				0.450
< 55	11	4 (36.36)	7 (63.64)	
≥ 55	19	11 (57.89)	8 (42.11)	
AFP (ng/mL)				0.450
< 200	19	11 (57.89)	8 (42.11)	
≥ 200	11	4 (36.36)	7 (63.64)	
CEA (µg/L)				0.999
< 5	27	13 (48.15)	14 (51.85)	
≥ 5	3	2 (66.67)	1 (33.33)	
ALT (u/L)				< 0.001
< 46	18	14 (77.78)	4 (22.22)	
≥ 46	12	1 (8.33)	11 (91.67)	
AST (u/L)				0.715
< 46	14	8 (57.14)	6 (42.86)	
≥ 46	16	7 (43.75)	9 (56.25)	
GGT (u/L)				0.700
< 55	20	9 (45.00)	11 (55.00)	
≥ 55	10	6 (60.00)	4 (40.00)	

Figures

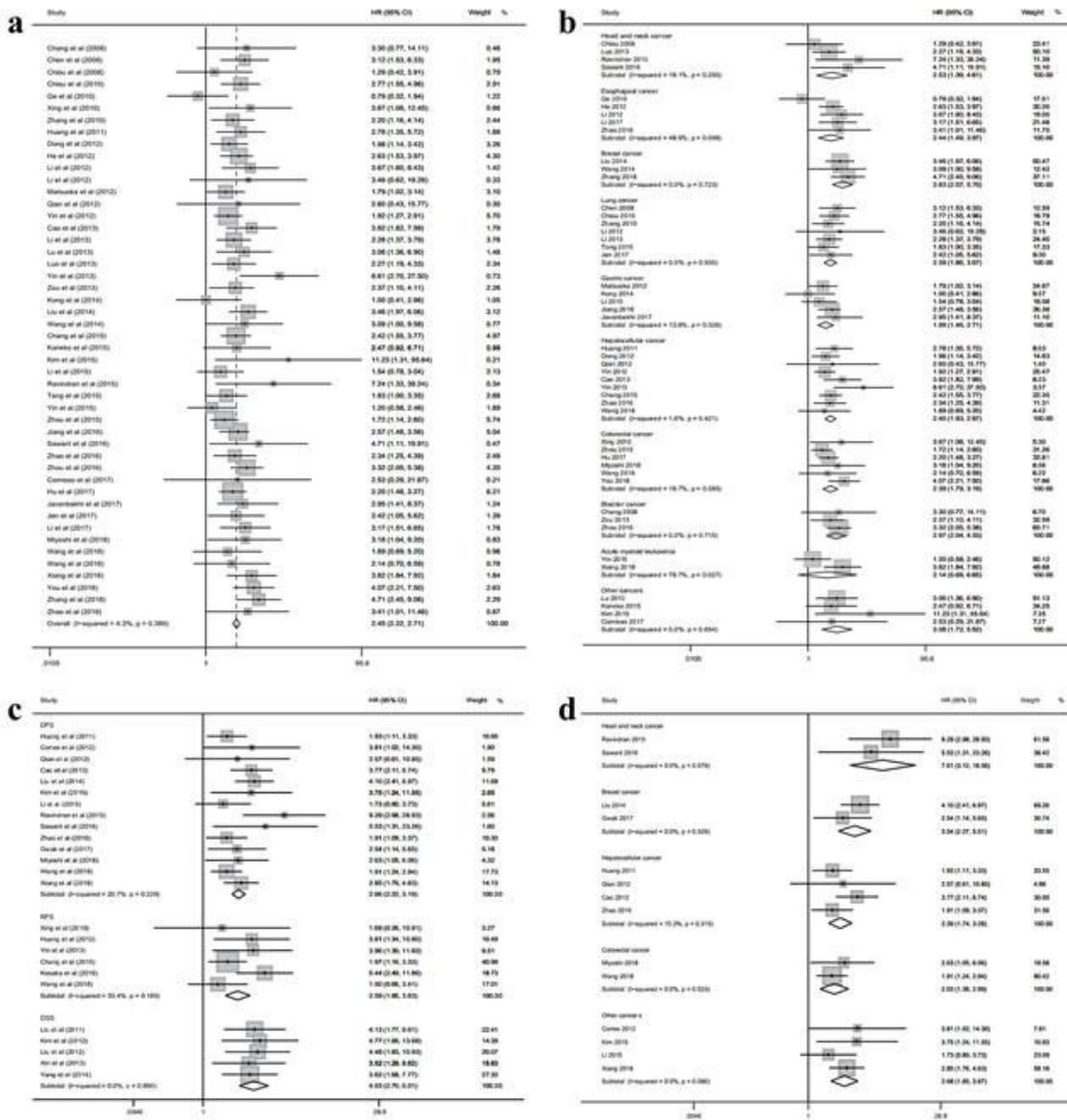


Figure 1

Forest plots of HRs for OS, DFS, RFS, DSS with elevated POU5F1 expression. a HRs for OS. b HRs for OS subgroup analysis of cancer type. c HRs for DFS, RFS and DSS. d HRs for DFS subgroup analysis of cancer type.

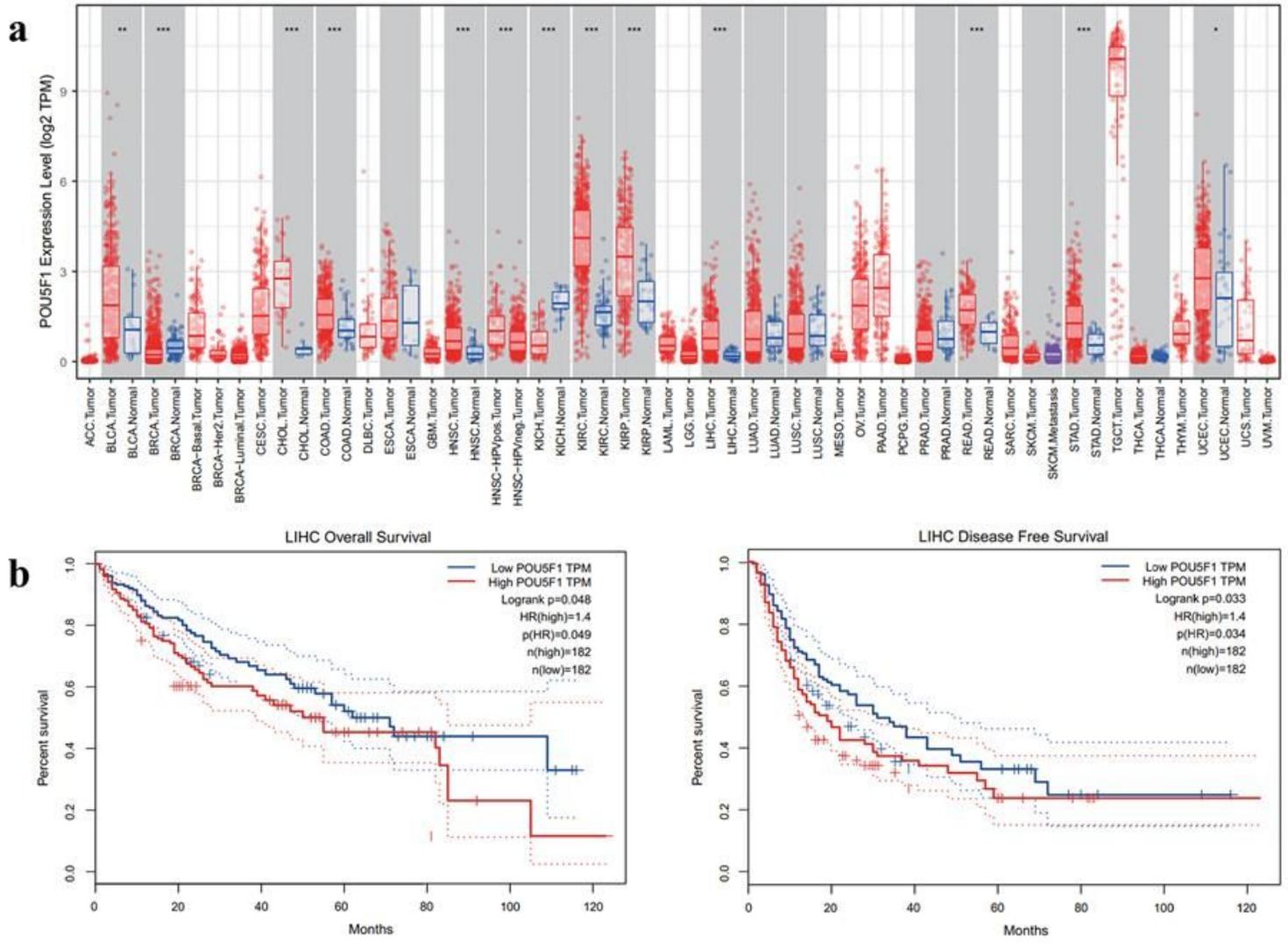


Figure 2

Expression status of POU5F1 in various cancers and survival analysis in LIHC. a Expression status of POU5F1 in various cancers based on TCGA. Statistical significance was assigned at $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***). b OS and DFS of LIHC patients with high ($n = 182$) or low ($n = 182$) POU5F1 levels in LIHC tissues.

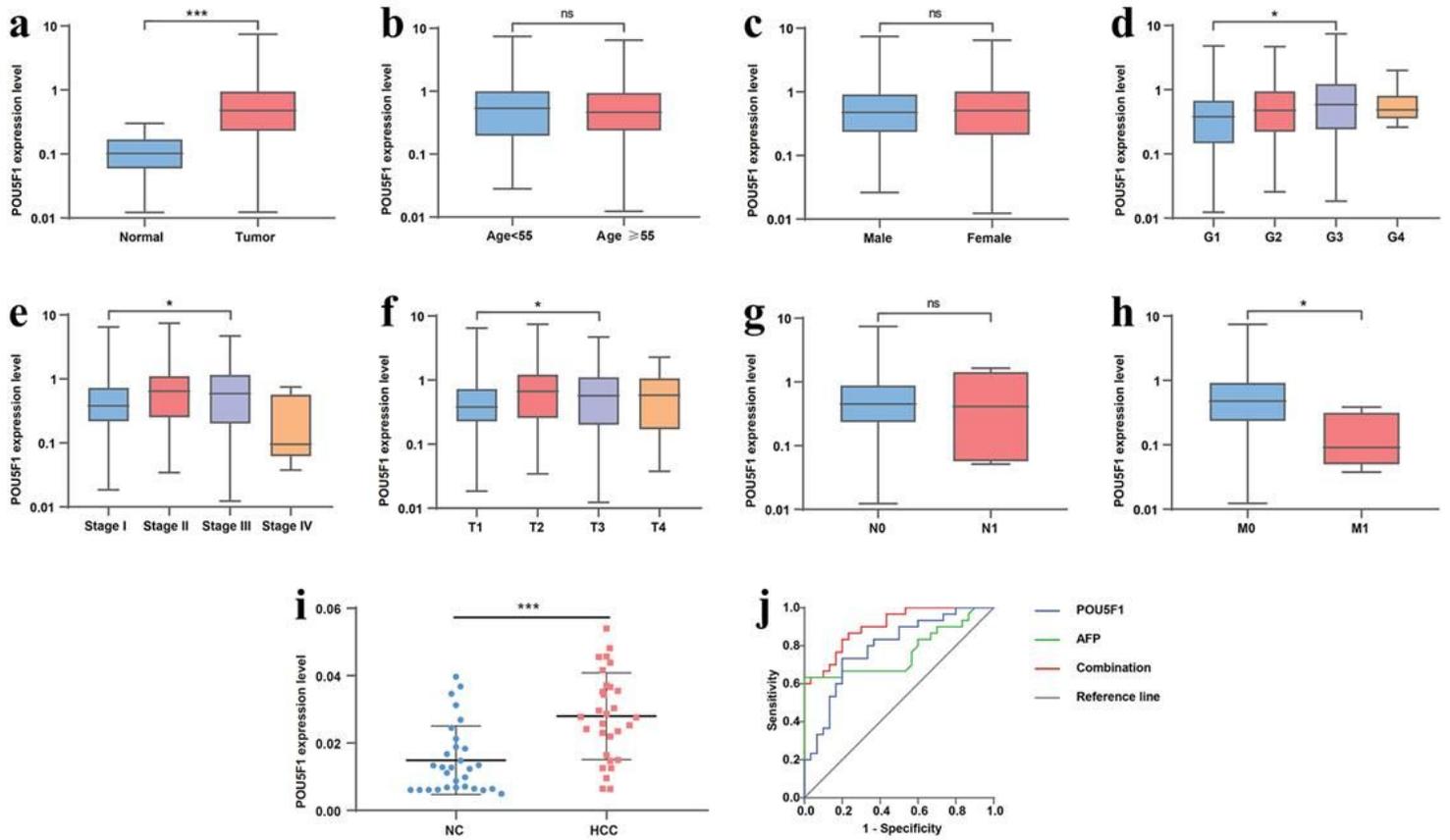


Figure 3

Association between POU5F1 expression and clinicopathologic characteristics and the diagnostic value of POU5F1 in LIHC. a Cancer status. b Age. c Gender. d Grade. e Clinical stage. f Tumor invasion depth. g Lymph node metastasis. h Distant metastasis. i Expression level of POU5F1 in plasma collected from 30 controls and 30 LIHC patients. j ROC based on POU5F1 and AFP levels in plasma separately or combinedly.

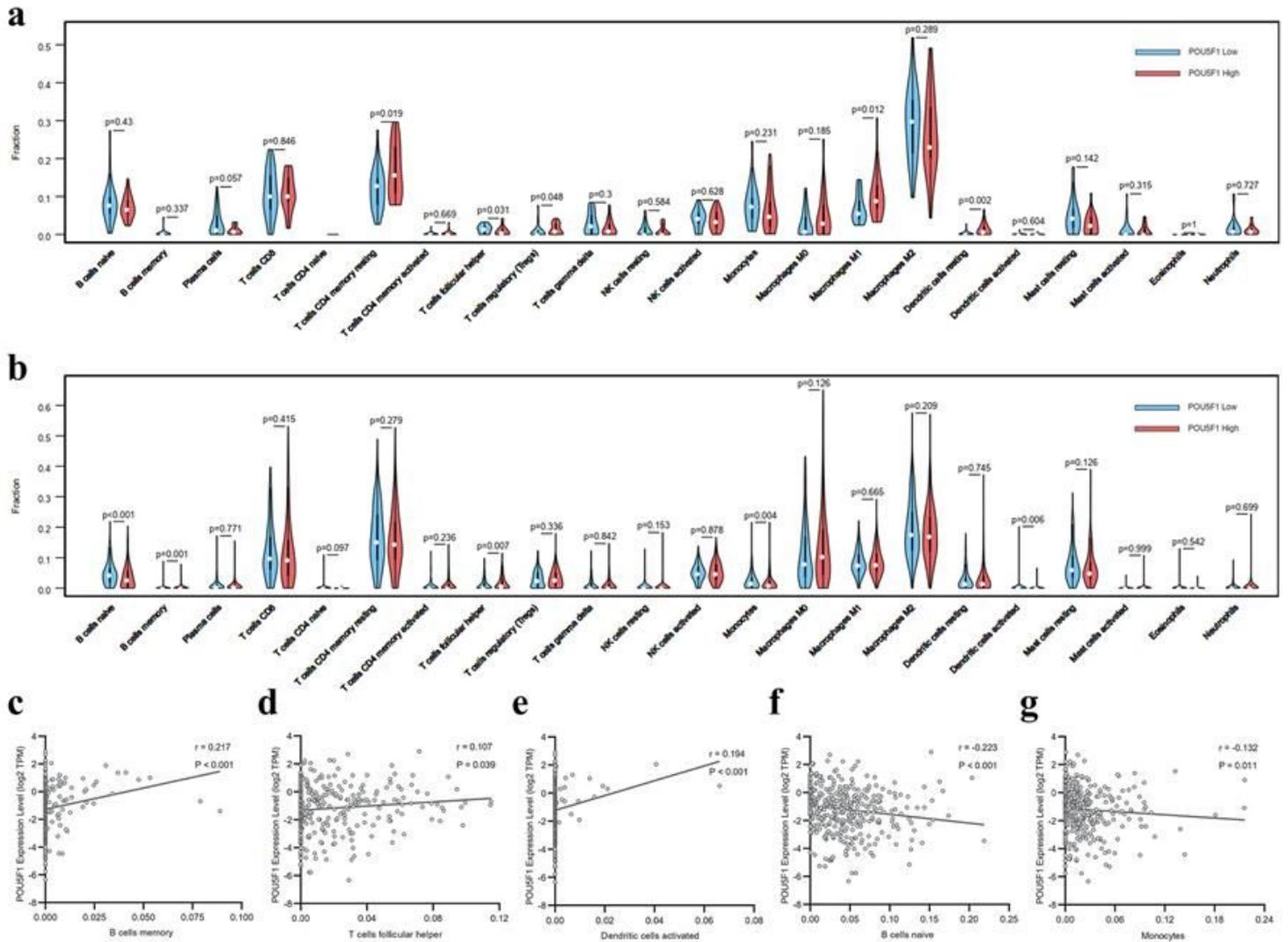


Figure 4

The proportion of 22 subpopulations of TIICs in normal liver tissues and LIHC tissues. a TIICs in normal liver tissues. b TIICs in LIHC tissues. Correlation between POU5F1 level and c B cells memory, d T cells follicular helper, e dendritic cells activated, f B cells naive, g monocytes.

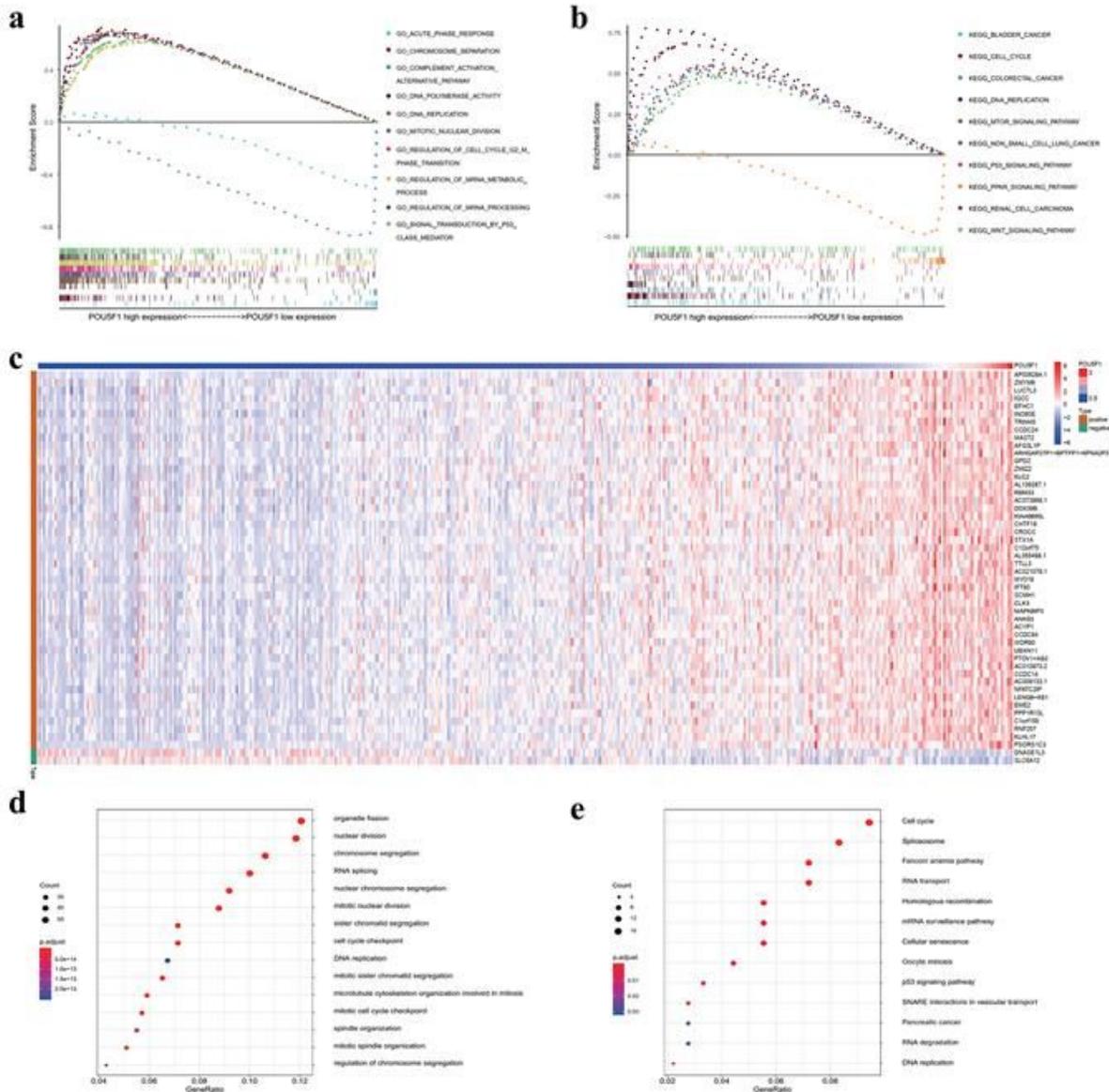


Figure 5

GSEA for POU5F1 and enrichment analysis of the co-expression genes of POU5F1 in LIHC. a GSEA of POU5F1 based on GO gene sets. b GSEA of POU5F1 based on KEGG gene sets. c Representative expression heat map of the top 50 co-expression genes of POU5F1. d GO enrichment analysis of the co-expression genes of POU5F1. e KEGG enrichment analysis of the co-expression genes of POU5F1.

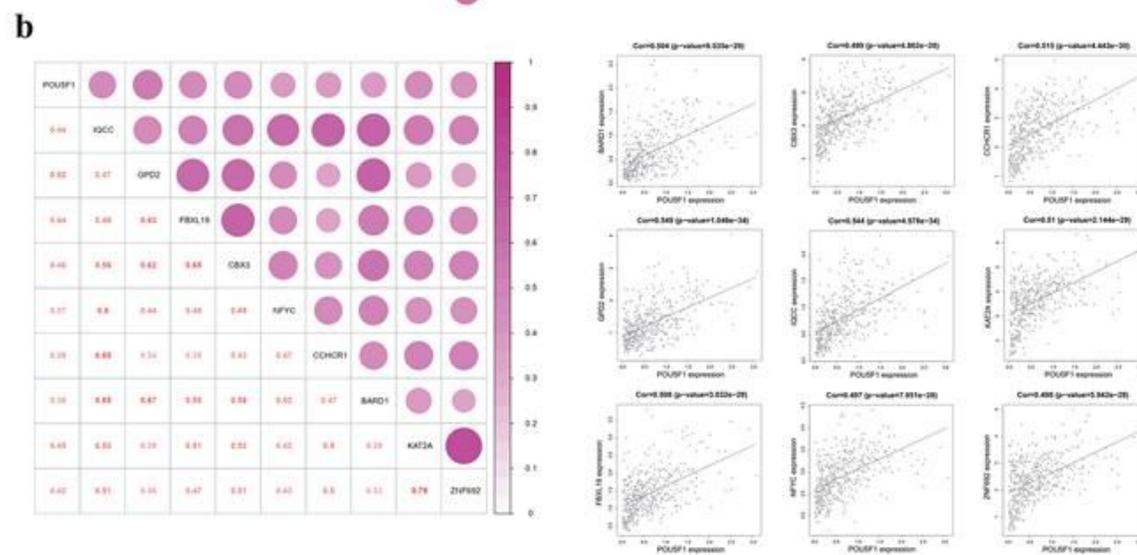
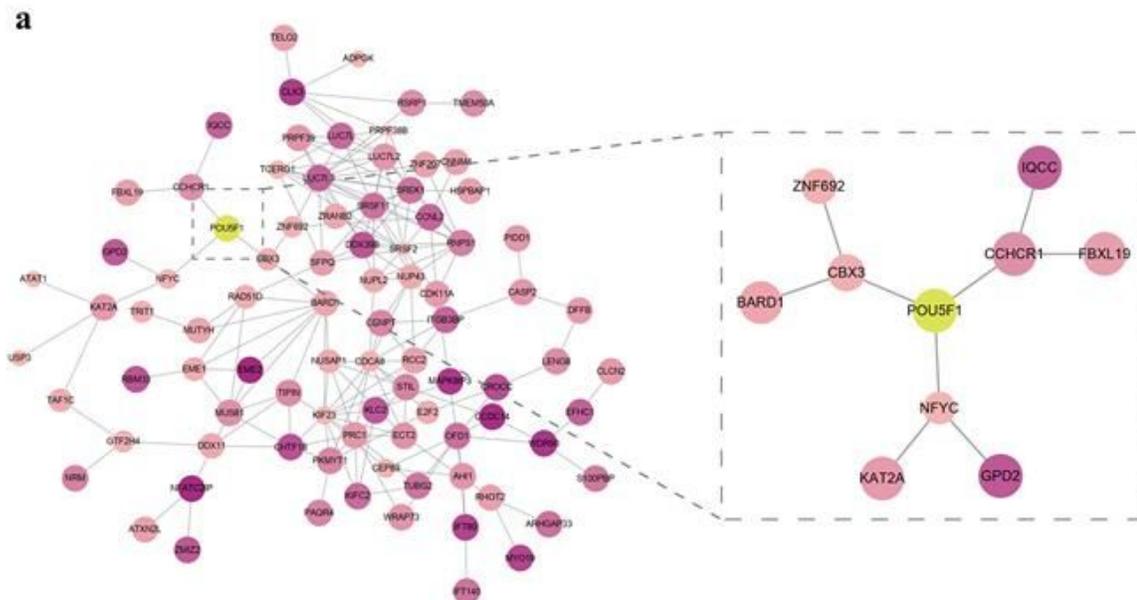


Figure 6

PPI network and the correlation between POU5F1 and the hub genes. a PPI network and the 9 hub genes interacted with POU5F1. b Correlation between expression of POU5F1 and the 9 hub genes.

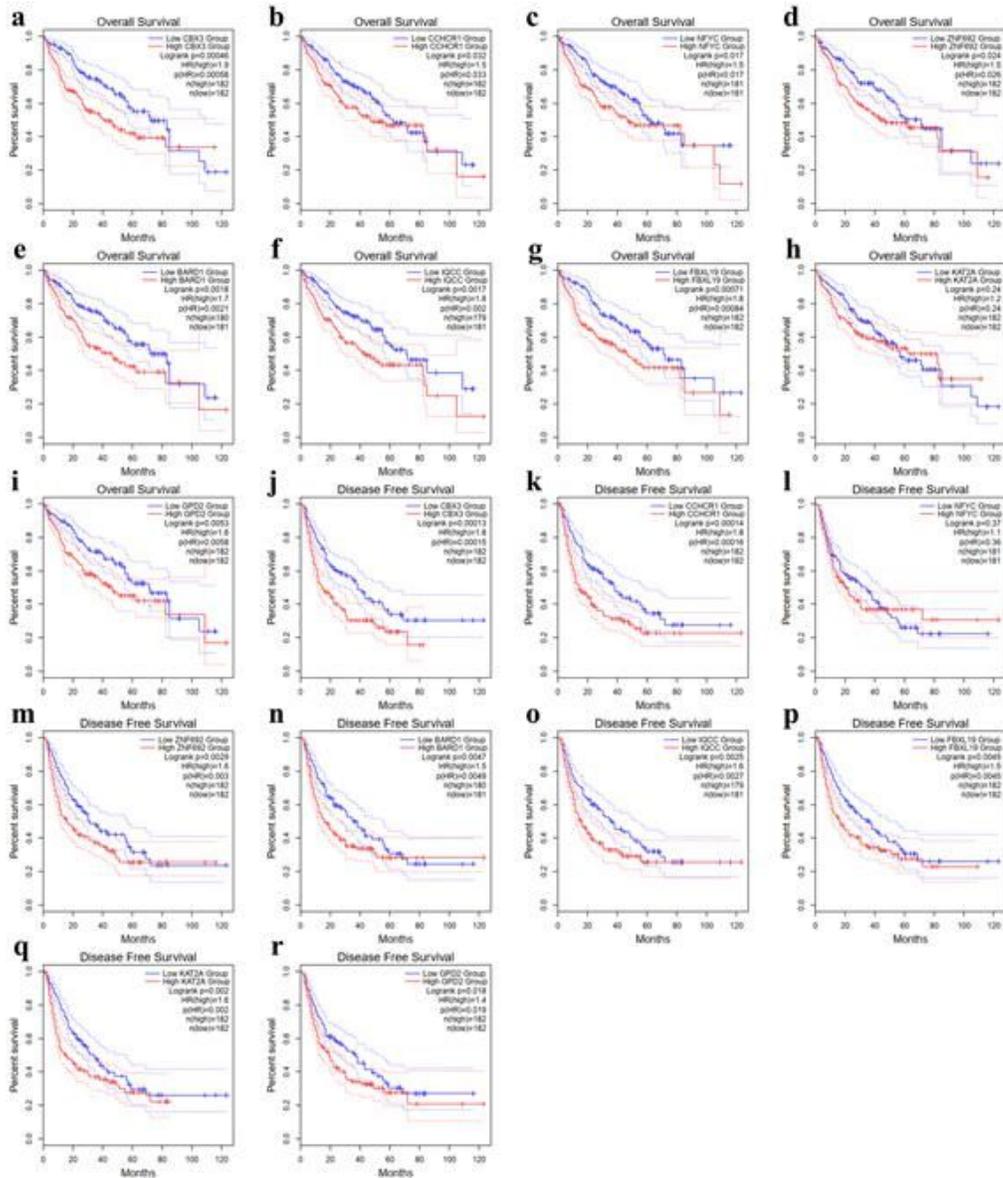


Figure 7

Kaplan–Meier survival analysis of 9 hub genes of POU5F1 in LIHC based on TCGA. a-i Overall survival of 9 hub genes. j-r Disease-free survival of 9 hub genes.

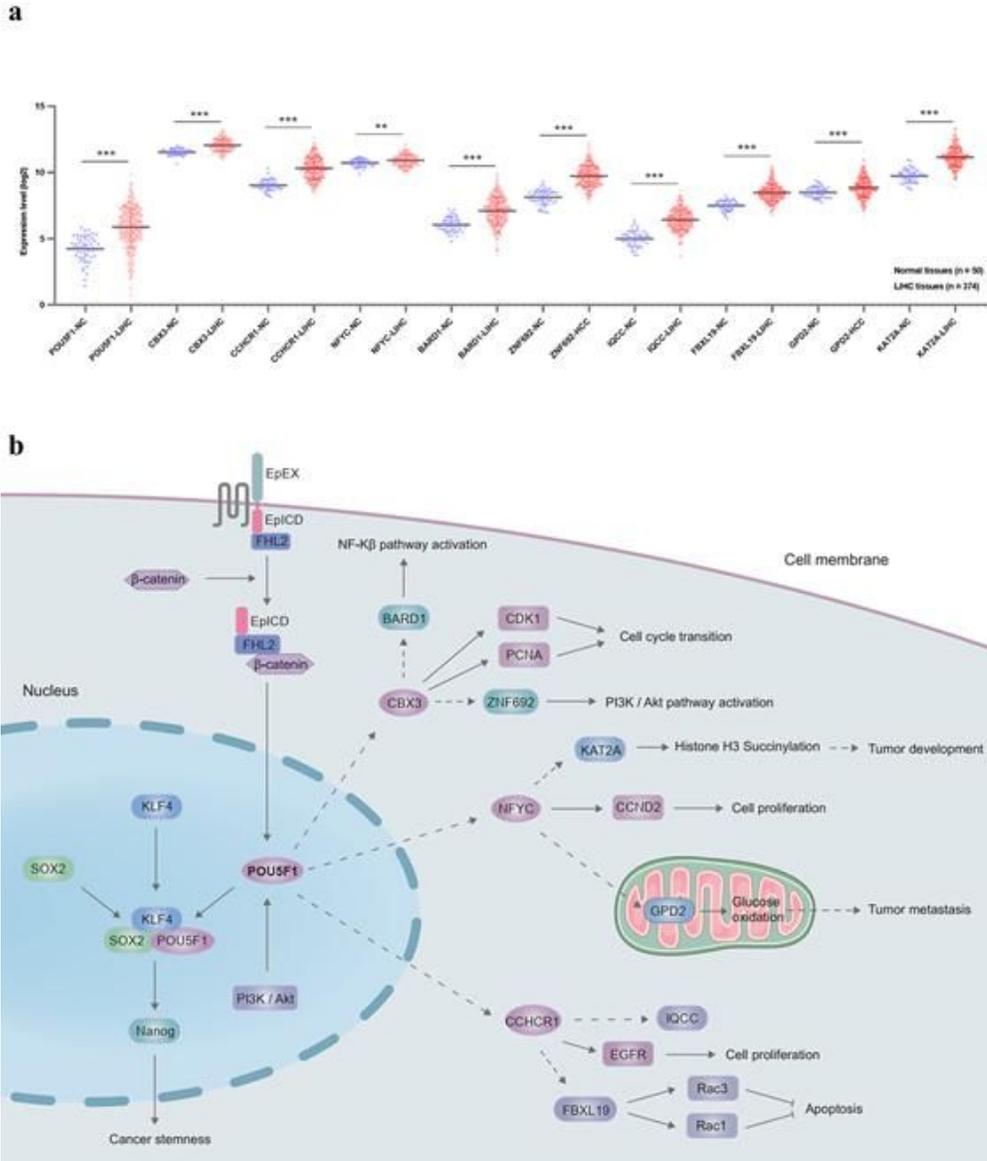


Figure 8

Expression levels and potential cell signal transduction pathways of POU5F1 and hub genes in LIHC. a Expression levels of POU5F1 and the 9 hub genes in LIHC. b Potential cell signal transduction pathways of POU5F1 and the 9 hub genes in LIHC.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile9.pdf](#)
- [Additionalfile9.pdf](#)
- [Additionalfile8.pdf](#)
- [Additionalfile8.pdf](#)
- [Additionalfile7.pdf](#)
- [Additionalfile7.pdf](#)

- [Additionalfile6.pdf](#)
- [Additionalfile6.pdf](#)
- [Additionalfile5.pdf](#)
- [Additionalfile5.pdf](#)
- [Additionalfile4.pdf](#)
- [Additionalfile4.pdf](#)
- [Additionalfile3.pdf](#)
- [Additionalfile3.pdf](#)
- [Additionalfile2.pdf](#)
- [Additionalfile2.pdf](#)
- [Additionalfile1.pdf](#)
- [Additionalfile1.pdf](#)