Multiple Omics Analysis Reveals E2F5 Predict Prognostic Biomarker in Diffuse Large B-Cell Lymphoma

Jinshui Tan
The First Affiliated Hospital of Xiamen University and Institute of Hematology
Mengya Zhong
The First Affiliated Hospital of Xiamen University and Institute of Hematology
Qinwei Chen
The First Affiliated Hospital of Xiamen University and Institute of Hematology
Zhen Lu
The First Affiliated Hospital of Xiamen University and Institute of Hematology
Jie Zha
The First Affiliated Hospital of Xiamen University and Institute of Hematology
Yong Zhou
The First Affiliated Hospital of Xiamen University and Institute of Hematology
Guowei Li
Affiliated Dongguan People's Hospital, Southern Medical University
Yirong Jiang
Huizhou Municipal Central Hospital
Bing Xu (xubing@xmu.edu.cn)
The First Affiliated Hospital of Xiamen University and Institute of Hematology

Research Article

Keywords: E2Fs, Diffuse large B-cell lymphoma (DLBCL), multi-omics analysis, DNA methylation, biomarkers

Posted Date: October 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-996006/v1

Abstract

Diffuse large B-cell lymphoma (DLBCL) is highly aggressive and fatal hematological malignancy. There are few biomarkers that can be used to predict the survival of DLBCL patients. Therefore, there is an urgent need to find new biological targets to improve the predictive value and sensitive diagnosis of DLBCL.E2F family play an essential role in tumorigenesis, however, remains obscure in DLBCL.E2F transcription factor family(E2Fs) mRNA expression between DLBCL and nonmalignant samples were screened by GEPIA,CCLE and EMBL-EBI. The associated regulation pathway in DLBLC was established using the GeneMANIA,Metascape ,SMATAPP database. Transcription analysis indicated E2F1/4/5/8 mRNA expression was significantly higher in patients and the cell lines.What's more, the high E2F5/8 expression had significantly lower survival rate. Further functional analysis showed that E2F1/3/5 were hypomethylated in DLBCL,which may associated with patient chemo-resistance. Subsequently,these genes with their co-expression genes mainly formed transcription factor complex, regulated G1/S transition of mitotic cell cycle and through TGF-beta signaling pathway to participate DLBCL tumorigenesis. This results demonstrate that E2F5 were potential prognostic biomarkers for better survival of DLBCL patients.

Introduction

DLBCL is a heterogeneous malignancy molecularly can be classified into at least three distinct subtypes (ABC, GCB, PMBL)[1]. According to adverse outcomes,biomarkers including MYC/BCL2 rearrangement, and TP53 mutations implicated worse prognosis[2]. R-CHOP regimen indeed improved overall survival, but more than 30% of patients will ultimately relapse [3].The cell origin phenotype and gene mutation are suggested as predictors in DLBCL, however, due to the heterogeneity and complexity of lymphoma, there are still many unknown oncogenes and tumor suppressor genes that need to be discovered. Thus, finding reliable biomarkers that can accurately detection of DLBCL is necessary.

The E2F family has been considered as an important regulator of the cell cycle,which plays vital roles in regulating transcription and tumor suppressor proteins. E2F family members are divided into transcriptional activators or repressors [4]. E2F1, E2F2 and E2F3A, as activator proteins, lead to adequate transition from the G1 to S phase and enhanced cellular proliferation [5 6]. E2F3b and E2F4/5/6,which are 'repressive' E2F branches, repress the transcription in quiescent and early G1 cells[7]. As atypical repressors, E2F7/8 suppress E2F1-induced cell cycle[8]. The E2Fs mRNA expression is aberrant in several human malignancies, such as breast cancer [9], gastric cancer[10], hepatocellular carcinoma[11], lung adenocarcinoma[12] and colon cancer[13]. To data, however, the overall biological role and clinical significance of the E2F factors in DLBCL remains not fully elucidated.

With the recent evolution of sequencing technologies, a vast amount sequencing data has been uploaded to the online repositories[14]. Here, we analyzed the unique expression patterns and significance for survival prognosis of eight E2Fs in DLBCL patients by bioinformatics.

Materials & Methods

Ethics statement

This research has been approved by the Ethics Committee of the First Affiliated Hospital of Xiamen University (Fujian, China). The TCGA and GTEx data was retrieved from published literature, and all procedures are implemented in accordance with relevant guidelines and regulations.

GEPIA Dataset Analysis

GEPIA (http://gepia.cancer-pku.cn/),which is a gene expression interactive analysis server,providing clinical and sequencing data from the TCGA and the GTEx projects. Meanwhile,it visualize this data by customizable functions depending on your need [15].

CCLE and EMBL-EBIDataset

The CCLE (https://www.broadinstitute.org/ccle) and EMBL-EBI(https://www.ebi.ac.uk)conduct a detailed multi-omics analysis of thousands of cell lines,which provides reference for patient stratification[16 17].

SMART APP

The SMART (Shiny Methylation Analysis Resource Tool) App combines multi-omics and clinical data with comprehensive DNA methylation analysis.It visualizes the relationship between gene expression and methylation in carcinoma [18].

Metascape

The Metascape (http://metascape.org),as a reliable, effective and intuitive tools,enables visualization analyses of multi-omic data. For this, it was utilized to conduct KEGG and GO(containing MF,BP,CC) analyses of E2Fs [19].

GeneMANIA analysis

GeneMANIA (http://www.genemania.org) uses genomics and proteomics data to predict genes with similar functions. Here it was employed for predict the function of E2Fs and co-expression genes in DLBCL [20].

Statistical analysis

The GSE31312 was analyzed by using R studio software version 1.2.1335. And it was used to verify OS and PFS between E2Fs expression of lymphoma patients.*p < 0.05, **p < 0.01, and ***p <0.001 mean statistically significant difference.

Results

E2Fs were highly expressed in DLBCL in the mRNA level

To call DEGs of E2Fs family in DLBCL, we used the GEPIA dataset to compare the mRNA expression of DLBCL patient samples from TCGA tumors (n=47) and GTEx healthy tissues (n=337). We used log2 (TPM + 1) to quantity mRNA expression data for E2Fs and found that E2F1/4/5/6/8 were highly-expressed in DLBCL, while E2F2/3/7 were not significantly different. (Fig.1)

Expression of E2Fs translation factors in lymphoma cell lines

After determining the expression of E2Fs in clinical lymphoma patients, we analyzed E2Fs' mRNA expression in cell lines through the CCLE and EMBL-EBI database. By assembling CCLE, we found that E2F1/2/4/5/8 were relatively highly expressed in DLBCL, comparing the average expression base line of all tumors (FigS1). Moreover, EMBL-EBI was employed to test mRNA level of E2F translation factors in DLBCL cell lines(n=21). E2F1/2/3/4/5/8 increased in most DLBCL cell lines (Fig.2), especially in SU-DHL-5, SU-DHL-10, WSU-NHL. Collectively, with the aim to further elucidate the major factors of E2Fs in DLCBL, three databases were searched for intersection of the predicted results. We found that E2F1/4/5/8 overexpress in the GEPIA, CCLE and EMBL-EBI. (Fig. 3)

Prognostic Values of E2F1/4/5/8 in Patients With DLBCL

Next, we initially assessed the E2F1/4/5/8 prognosis values in 471 DLBCL patients (GSE31312). It was determined that the higher expression of E2F1 mRNA predicts extended OS and PFS((p=0.0015,p=0.02, respectively). Elevated E2F5/8 were significantly correlated with shortened OS (p=0.0051,0.015, respectively) and PFS (p=0.019,0.15, respectively). (Fig.4). Next, multivariate Cox's hazard regression analysis results indicated E2F1 (HR = 1.45, 95%CI = 1.141–1.841, P = 0.002), E2F5 (HR = 0.74, 95% CI = 0.584-0.938, P = 0.013), E2F8 (HR = 1.607, 95%CI = 1.103-2.343, P = 0.014) expression level are independent prognostic factors for overall survival in DLBCL patients (Table1).

Functional enrichment analysis of E2F1/4/5/8 and co-expression genes in patients with DLBCL

To reveal the function and potential mechanism of E2F1/4/5/8, we constructed a network of E2F1/4/5/8 and their neighboring genes (TFDP1/2, RBL1/2, E2F2) by the GeneMANIA database(Fig5A). By analyzing Metascape, we found the E2F1/4/5/8 and their neighboring genes were mainly enriched in G1/S cell cycle. These genes may form protein-DNA complex.Then they enhanced the interaction between RNA polymerase and specific promoters to promote targeted gene expression by changing the structure of DNA.(Fig.5B and Table2).

The top KEGG pathways showed that the cell cycle signaling pathway and TGF-beta signaling pathway were significantly found to be involved in the development of multiple tumor and participated in the tumorigenesis and pathogenesis of lymphoma.(Fig.5C and Table 3.) In addition, to further understand the role of E2Fs in lymphoma, we conducted a relevant protein–protein interaction (PPI) analysis(Fig.5D).

The results showed that E2F2/3/4/5, RBL1/2, TFDP1/2, CDK2 mainly formed transcription factor complex, regulated the maligant proliferation of tumor through G1/S transition of mitotic cell cycle and TGF-beta signaling pathway.

E2Fs expression is negatively regulated by DNA methylation

In diffuse large B-cell lymphoma, DNA methylation is the most widely studied epigenetic process leading to chemotherapy resistance[21]. In order to explore the transcriptional regulation of E2Fs family in DLBCL, we analyzed the 450 K methylation array data using SMARTAPP Databases to verify whether E2Fs expression may be regulated by their DNA methylation status. By comparing the Correlation between gene expression and methylation in 47 DLBCL patients (TCGA), The method for Correlation Coefficient is Spearman, we found that E2F1/3/5 expression is negative with methylation sites (p=0.00013,0.026,0.0011, respectively) (Fig.6)

Gene–gene interaction network of the E2F1/3/5 with the DNA methylation associated genes

Above, we found that E2F1/3/5 were hypomethylated in the TCGA databases. Previous studies indicated that human DNA methylation is catalyzed by DNA transmethylase including DNMT1/3A/3B [22]. Therefore, we used the GeneMANIA database to predict the interaction relationship of the E2F1/3/5 and DNMT1/3A/3B (Fig.7). The 6 central nodes representing E2F1/3/5, DNMT1/3A/3B were surrounded by 20 nodes representing genes,which have similar functions(Fig.7A). E2F2/4, TFDP1/2 was correlated with central nodes, participating in the same reaction within a pathway. (Fig.7B). E2F2/4/6, TFDP1/2, TRDMT1, DNMT3L, DMAP1, ZBTB18 was correlated with central nodes, predicting functional relationships between protein interactions(Fig.7C). As for co-localization, E2F4 was correlated with E2F3/5, DNMT1(Fig.7D). In addition, HDGF, TFDP1/2 was correlated with E2F1/3, DNMT1 in terms of co-expression(Fig.7E).

In GEPIA Dataset, E2F1, E2F4, E2F5, E2F6, E2F8, TFDP1, HDGF and DNMT1 expression is higher in DLBCL patients (Fig.1, FigS2). Next, we found that the methylation of HDGF, PWWP2A, TFDP1, TRDMT1, E2F1, E2F3, E2F5 and DNMT1 were significantly negative with gene expression in SMART APP Databases. In conclusion, elevated E2F1, E2F5, TFDP1, DNMT1 and HDGF were negative with DNA methylation in DLBCL patients (Fig6,FigS3). In GSE31312, correlation analysis showed that E2F5 is positive correlation with TFDP1(R=0.18, P<0.001) and HDGF (R=0.13, P<0.01), DNMT1 is negative correlation with TFDP1(R=-0.16, P<0.001) and HDGF (R=-0.16, P<0.001) (Fig.8) in DLBCL patients.

Discussion

DLBCL is the most common form of lymphoma. Current biological and clinical research shows that it is a highly heterogeneous and aggressive malignant tumor, which is manifested in clinical outcome, genetic characteristics, cell origin, etc.[23]. Although with the clinical application of the RCHOP treatment strategy, the cure rate of DLBCL has been significantly improved, 30-40% of patients still relapse[24]. Therefore, researches looking for biomarkers to predict clinical prognosis in DLBCL are necessary. Researches

suggested that the E2F transcription factors family plays a key role in cell cycle network regulation, which regulate cell proliferation, differentiation, apoptosis, and participate in physiological and pathological processes. [25,26]. However, the multifaceted roles of E2Fs in the development, metastasis, and prognostication of DLBCL remain to be clarified. Our study analyzed the transcription levels of E2F mRNA expression, potential pathways and prognostic (OS/PFS) values in DLBCL.

E2F1 preferentially binds to the retinoblastoma protein (pRB) in a cycle-dependent manner[27]. It can simultaneously mediate cell proliferation and p53-dependent or independent apoptosis [28]. Our results indicated that the higher expression of E2F1 mRNA means extended OS and PFS. Møller found that E2F1 as tumor suppressor gene in DLBCL and low E2F1 expression was associated with treatment failure of DLBCL, which may serving as prognostic markers for DLBCL patients [29]. The research results are consistent with ours. However, Samaka found that high E2F1 expression had shorter survival time of DLBCL cases and upregulation of E2F1 indicated that the tumor is more malignant [30]. Therefore, more detail researches need to do to prove the relation between the E2F1 and DLBCL patient survival.

As an oncogene, E2F5 can cooperate with other oncogenes promote cell malignant transformation. E2F5 is up-regulated and shows significant correlation with pathological variables and tumorigenesis[31]. E2F5 repressors as effectors of RB that control cell proliferation and apoptosis[8]. Previous studies reveal that E2F5 was highly expressed in several tumors[32], such as glioblastoma[33], breast cancer [34] and prostate cancer[35]. In this study, we found that Elevated expressions of E2F5 were significantly correlated with shortened OS/PFS in DLBCL patients.

As a tumor suppressor, E2F8 mediates transcriptional suppression by blocking cells from entering S phase[36]. Over-expression of E2F8 activates target genes that may promote cell cycle, mitosis, immune and other cancer related functions in Burkitt's lymphoma (BURK) and mantle cell lymphoma (MCL) [37]. Our research showed that elevated expressions of E2F8 were significantly correlated with shortened OS in DLBCL.

In our research, we revealed that E2F1/4/5/8 mRNA expression level were up-regulated in the patients and the cell lines of DLBCL. Further analysis showed that E2F2/3/4/5, RBL1/2, TFDP1/2, CDK2 mainly formed transcription factor complex participating cell cycle,and TGF-beta signaling pathway in DLBCL. Previous studies have shown that abnormally methylated genes affect its molecular expression regulation and patient's survival, which found that gene expression is associated with methylation in DLBCL [38]. We found that the E2F1/3/5 expression is negative with DNA methylation. Emerging studies have provided evidence that E2F and other transcription factors regulate DNMTs expression through E2F-Rb-HDAC-dependent and -independent pathways, which potentially contribute to tumor progression[39]. In our study, we found that elevated E2F1, E2F5, TFDP1, HDGF and DNMT1 were negative with DNA methylation in DLBCL patients. TFDP-1 and E2F factors jointly participate in the normal cell cycle procession. Somatic mutations in TFDP-1 uncouples the normal biological processes of the E2F pathway, which could lead to tumorigenesis[40]. HDGF upregulates in a variety of malignant tumors ,which is closely

related to the tumor stage, grade and proliferation [41]. Studies have found that cancers harboring mutant p53 genes are accompanied by high expression of HDGF[42]. In DLBCL patients, p53 mutations are accompanied by related gene expression dysregulation leading to poor prognosis[43]. In our research, we found that TFDP1and HDGF are positive correlation with E2F5, which is negative with DNMT1 expression in DLBCL patient databases. This result also consistent with observations in immortalized cell lines, indicating that only co-expressed with TFDP1,the transient overexpression of E2F5 can stimulate the cell proliferation[31]. We indicated that DNMT1-mediated DNA methylation of E2F5/TFDP1/HDGF ternary complex regulates prognosis in DLBCL patients. These deregulation makes cell cycle change and has prognostic significance in DLBCL.

Conclusions

We comprehensively analyzed the role of E2F transcription factors in the occurrence and development of DLBCL and its prognosis through bioinformatics. we found that E2F1/4/5/8 were over-expression in DLBCL cell lines and patients. And E2F2/3/4/5, RBL1/2, TFDP1/2, CDK2 mainly formed transcription factor complex, regulating cell cycle and TGF-beta signaling pathway in DLBCL. E2Fs expression and DNA methylation status showed that E2F1/3/5 expression might be negatively related to DNA methylation in DLBCL. These genes with DNA methylation genes were correlated with TFDP1 and HDGF, they may anticipate the pathogenesis of DLBCL. Furthermore, our study indicated that increased expression of E2F5 indicated worse prognosis in DLBCL, it may be potential targets for individualized treatment of DLBCL patients. In conclusion, this research demonstrates for the first time that E2F5 were potential prognostic biomarkers for better survival of DLBCL patients.

References

1. Dunleavy, K., Grant, C. & Wilson, W. H. Using biologic predictive factors to direct therapy of diffuse large B-cell lymphoma. *Ther Adv Hematol*, **4**, 43–57 https://doi.org/10.1177/2040620712464508 (2013).
2. Wang, W. T. *et al.* The prognostic value of 25-hydroxy vitamin D deficiency and its interaction with c-Myc expression in diffuse large B cell lymphoma. *Ann Hematol*, https://doi.org/10.1007/s00277-020-04200-5 (2020).
3. Huang, W. *et al.* Clinical significance of PCDH10 promoter methylation in diffuse large B-cell lymphoma. *BMC Cancer*, **17**, 815 https://doi.org/10.1186/s12885-017-3810-7 (2017).
4. Emanuele, M. J. *et al.* Complex Cartography: Regulation of E2F Transcription Factors by Cyclin F and Ubiquitin. *Trends Cell Biol*, **30**, 640–652 https://doi.org/10.1016/j.tcb.2020.05.002 (2020).
5. DeGregori, J., Leone, G., Miron, A., Jakoi, L. & Nevins, J. R. Distinct roles for E2F proteins in cell growth control and apoptosis. *Proc Natl Acad Sci U S A*, **94**, 7245–7250 https://doi.org/10.1073/pnas.94.14.7245 (1997).

6. Wu, L. *et al.* The E2F1-3 transcription factors are essential for cellular proliferation. *Nature*, **414**, 457–462 https://doi.org/10.1038/35106593 (2001).
7. Kent, L. N. & Leone, G. The broken cycle: E2F dysfunction in cancer. *Nat Rev Cancer*, **19**, 326–338 https://doi.org/10.1038/s41568-019-0143-7 (2019).
8. Chen, H. Z., Tsai, S. Y. & Leone, G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer*, **9**, 785–797 https://doi.org/10.1038/nrc2696 (2009).
9. Sun, C. C. *et al.* Comprehensive Analysis of the Expression and Prognosis for E2Fs in Human Breast Cancer. *Mol Ther*, **27**, 1153–1165 https://doi.org/10.1016/j.ymthe.2019.03.019 (2019).
10. Manicum, T., Ni, F., Ye, Y., Fan, X. & Chen, B. C. Prognostic values of E2F mRNA expression in human gastric cancer. *Biosci Rep*, **38**, https://doi.org/10.1042/bsr20181264 (2018).
11. Huang, Y. L. *et al.* Promising diagnostic and prognostic value of E2Fs in human hepatocellular carcinoma. *Cancer Manag Res*, **11**, 1725–1740 https://doi.org/10.2147/cmar.S182001 (2019).
12. Sun, C. C. *et al.* Transcriptional E2F1/2/5/8 as potential targets and transcriptional E2F3/6/7 as new biomarkers for the prognosis of human lung carcinoma. *Aging (Albany NY)*, **10**, 973–987 https://doi.org/10.18632/aging.101441 (2018).
13. Yao, H., Lu, F. & Shao, Y. The E2F family as potential biomarkers and therapeutic targets in colon cancer. *PeerJ*, **8**, e8562 https://doi.org/10.7717/peerj.8562 (2020).
14. Sealfon, S. C. & Chu, T. T. RNA and DNA microarrays. *Methods Mol Biol*, **671**, 3–34 https://doi.org/10.1007/978-1-59745-551-0_1 (2011).
15. Tang, Z. *et al.* GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*, **45**, W98–w102 https://doi.org/10.1093/nar/gkx247 (2017).
16. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, **483**, 603–607 https://doi.org/10.1038/nature11003 (2012).
17. Li, W. *et al.* The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res*, **43**, W580–584 https://doi.org/10.1093/nar/gkv279 (2015).
18. Li, Y., Ge, D. & Lu, C. The SMART App: an interactive web application for comprehensive DNA methylation analysis and visualization. *Epigenetics Chromatin*, **12**, 71 https://doi.org/10.1186/s13072-019-0316-3 (2019).
19. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*, **10**, 1523 https://doi.org/10.1038/s41467-019-09234-6 (2019).
20. Warde-Farley, D. *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*, **38**, W214–220 https://doi.org/10.1093/nar/gkq537 (2010).
21. Clozel, T. *et al.* Mechanism-based epigenetic chemosensitization therapy of diffuse large B-cell lymphoma. *Cancer Discov*, **3**, 1002–1019 https://doi.org/10.1158/2159-8290.Cd-13-0117 (2013).
22. Lyko, F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat Rev Genet*, **19**, 81–92 https://doi.org/10.1038/nrg.2017.80 (2018).

23. Reddy, A. *et al.* Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma., **171**, 481–494 https://doi.org/.e415 (2017).
24. Younes, A. Prognostic Significance of Diffuse Large B-Cell Lymphoma Cell of Origin: Seeing the Forest and the Trees. *J Clin Oncol*, **33**, 2835–2836 https://doi.org/10.1200/jco.2015.61.9288 (2015).
25. Cam, H. & Dynlacht, B. D. Emerging roles for E2F: beyond the G1/S transition and DNA replication., **3**, 311–316 https://doi.org/10.1016/s1535-6108(03)00080-1 (2003).
26. Timmers, C. *et al.* E2f1, E2f2, and E2f3 control E2F target expression and cellular proliferation via a p53-dependent negative feedback loop. *Mol Cell Biol*, **27**, 65–78 https://doi.org/10.1128/mcb.02147-06 (2007).
27. Dyson, N. J. RB1: a prototype tumor suppressor and an enigma. *Genes Dev*, **30**, 1492–1502 https://doi.org/10.1101/gad.282145.116 (2016).
28. Denechaud, P. D., Fajas, L. & Giralt, A. E2F1, a Novel Regulator of Metabolism. *Front Endocrinol (Lausanne)*, **8**, 311 https://doi.org/10.3389/fendo.2017.00311 (2017).
29. Møller, M. B. *et al.* Frequent disruption of the RB1 pathway in diffuse large B cell lymphoma: prognostic significance of E2F-1 and p16INK4A., **14**, 898–904 https://doi.org/10.1038/sj.leu.2401761 (2000).
30. Samaka, R. M., Aiad, H. A., Kandil, M. A., Asaad, N. Y. & Holah, N. S. The Prognostic Role and Relationship between E2F1 and SV40 in Diffuse Large B-Cell Lymphoma of Egyptian Patients. Anal Cell Pathol (Amst) 2015, 919834, doi:10.1155/2015/919834 (2015).
31. Polanowska, J. *et al.* Human E2F5 gene is oncogenic in primary rodent cells and is amplified in human breast tumors. *Genes Chromosomes Cancer*, **28**, 126–130 (2000).
32. Tian, H. *et al.* miR-132 targeting E2F5 suppresses cell proliferation, invasion, migration in ovarian cancer cells. *Am J Transl Res*, **8**, 1492–1501 (2016).
33. Fang, D. Z. *et al.* MicroRNA-129-3p suppresses tumor growth by targeting E2F5 in glioblastoma. *Eur Rev Med Pharmacol Sci*, **22**, 1044–1050 https://doi.org/10.26355/eurrev_201802_14387 (2018).
34. Cai, C., Huo, Q., Wang, X., Chen, B. & Yang, Q. SNHG16 contributes to breast cancer cell migration by competitively binding miR-98 with E2F5. *Biochem Biophys Res Commun*, **485**, 272–278 https://doi.org/10.1016/j.bbrc.2017.02.094 (2017).
35. Majumder, S. *et al.* Deregulated E2F5/p38/SMAD3 Circuitry Reinforces the Pro-Tumorigenic Switch of TGFβ Signaling in Prostate Cancer. *J Cell Physiol*, **231**, 2482–2492 https://doi.org/10.1002/jcp.25361 (2016).
36. Yuan, R. *et al.* Cyclin F-dependent degradation of E2F7 is critical for DNA repair and G2-phase progression. *Embo j*, **38**, e101430 https://doi.org/10.15252/embj.2018101430 (2019).
37. Yang, A. P. *et al.* Integrated analysis of 10 lymphoma datasets identifies E2F8 as a key regulator in Burkitt's lymphoma and mantle cell lymphoma. *Am J Transl Res*, **11**, 4382–4396 (2019).
38. Chambwe, N. *et al.* Variability in DNA methylation defines novel epigenetic subgroups of DLBCL associated with different clinical outcomes., **123**, 1699–1708 https://doi.org/10.1182/blood-2013-

07-509885 (2014).

39. Kimura, H., Nakamura, T., Ogawa, T., Tanaka, S. & Shiota, K. Transcription of mouse DNA methyltransferase 1 (Dnmt1) is regulated by both E2F-Rb-HDAC-dependent and -independent pathways. *Nucleic Acids Res*, **31**, 3101–3113 https://doi.org/10.1093/nar/gkg406 (2003).
40. Ishida, H. *et al.* Identification and characterization of novel isoforms of human DP-1: DP-1{alpha} regulates the transcriptional activity of E2F1 as well as cell cycle progression in a dominant-negative manner. *J Biol Chem*, **280**, 24642–24648 https://doi.org/10.1074/jbc.M500189200 (2005).
41. Bao, C., Wang, J., Ma, W., Wang, X. & Cheng, Y. HDGF: a novel jack-of-all-trades in cancer. *Future Oncol*, **10**, 2675–2685 https://doi.org/10.2217/fon.14.194 (2014).
42. Sasaki, Y. *et al.* p53 negatively regulates the hepatoma growth factor HDGF. *Cancer Res*, **71**, 7038–7047 https://doi.org/10.1158/0008-5472.Can-11-1053 (2011).
43. Ichikawa, A. *et al.* Mutations of the p53 gene as a prognostic factor in aggressive B-cell lymphoma. *N Engl J Med*, **337**, 529–534 https://doi.org/10.1056/nejm199708213370804 (1997).

Tables

Table I
Cox regression of E2F1, E2F4, E2F5, E2F8 in Patients with DLBCL

	OS			PFS		
	P	HR	95.0% CI	P	HR	95.0% CI
Age (≥60 vs. <60), years	0.275	1.145	0.898-1.46	0.017*	1.35	1.054-1.729
Sex (males vs. females)	0.511	0.925	0.732-1.168	0.636	0.944	0.745-1.197
Stage (low vs. high)	0.25	0.855	0.655-1.117	0.276	0.856	0.648-1.132
ECOG score (low vs. high)	0.952	1.011	0.713-1.432	0.482	1.138	0.793-1.635
IPI score (low vs. high)	0.552	1.099	0.805-1.499	0.261	1.198	0.874-1.642
Treatment response (CR+PR vs. PD+SD)	0.011*	2.758	1.258-6.047	0.022*	2.433	1.135-5.214
Subtype (non-GCB vs. GCB)	0.798	1.031	0.813-1.398	0.437	1.1	0.865-1.4
E2F1 expression (low vs. high)	0.002*	1.45	1.141-1.841	0.016*	1.652	1.097-2.486

	OS			PFS		
	P	HR	95.0% CI	P	HR	95.0% CI
Age (≥60 vs. <60), years	0.22	1.164	0.913-1.484	0.03*	1.315	1.027-1.684
Sex (males vs. females)	0.654	0.948	0.751-1.197	0.494	0.92	0.724-1.168
Stage (low vs. high)	0.471	0.907	0.697-1.182	0.404	0.889	0.674-1.172
ECOG score (low vs. high)	0.567	1.108	0.779-1.575	0.706	1.07	0.751-1.525
IPI score (low vs. high)	0.869	1.026	0.753-1.399	0.387	1.15	0.838-1.578
Treatment response (CR+PR vs. PD+SD)	0.022*	2.485	1.138-5.427	0.058*	2.057	0.976-4.334
Subtype (non-GCB vs. GCB)	0.509	1.083	0.855-1.372	0.41	1.106	0.87-1.408
E2F4 expression (low vs. high)	0.205	1.271	0.877-1.841	0.031*	0.767	0.603-0.976

	OS			PFS		
	P	HR	95.0% CI	P	HR	95.0% CI
Age (≥60 vs. <60), years	0.29	1.139	0.895-1.451	0.021	1.339	1.044-1.718
Sex (males vs. females)	0.531	0.928	0.734-1.173	0.668	0.949	0.748-1.204
Stage (low vs. high)	0.411	0.895	0.687-1.166	0.353	0.876	0.663-1.158
ECOG score (low vs. high)	0.469	1.138	0.802-1.616	0.872	1.03	0.723-1.465
IPI score (low vs. high)	0.925	1.015	0.745-1.383	0.287	1.188	0.865-1.632
Treatment response (CR+PR vs. PD+SD)	0.035*	2.325	1.063-5.089	0.081*	1.947	0.921-4.116
Subtype (non-GCB vs. GCB)	0.475	1.09	0.861-1.38	0.319	1.131	0.888-1.439
E2F5 expression (low vs. high)	0.013*	0.74	0.584-0.938	0.317	0.831	0.578-1.195

	OS			PFS		
	P	HR	95.0% CI	P	HR	95.0% CI
Age (≥60 vs. <60), years	0.158	1.192	0.934-1.52	0.025*	1.329	1.037-1.704
Sex (males vs. females)	0.634	0.945	0.748-1.193	0.537	0.928	0.731-1.177
Stage (low vs. high)	0.476	0.908	0.696-1.184	0.229	0.842	0.636-1.114
ECOG score (low vs. high)	0.61	1.094	0.774-1.548	0.807	0.957	0.672-1.363
IPI score (low vs. high)	0.729	1.056	0.775-1.44	0.198	1.231	0.897-1.69
Treatment response (CR+PR vs. PD+SD)	0.009*	2.885	1.31-6.352	0.036*	2.233	1.056-4.721
Subtype (non-GCB vs. GCB)	0.721	1.044	0.824-1.324	0.214	1.166	0.915-1.486
E2F8 expression (low vs. high)	0.014*	1.607	1.103-2.343	0.011*	0.731	0.575-0.93

OS: Overall Survival; PFS: progression Free Survival; ECOG: Eastern Cooperative Oncology Group; IPI: International Prognostic Index; CR: Complete Response; PR: Partial Response; PD: Progressive Disease; SD: Stable Disease; GCB: germinal center B-cell-likely lymphoma; *: $p < 0.05$.

Table II:

GO Analysis of E2F1/4/5/8 and co-expression genes in Patients with DLBCL using Metascape

GO	Category	Description	Count	%	Log10 (P)	Log10 (q)
GO:0000082	Biological Processes	G1/S transition of mitotic cell cycle	11	50.00	-15.59	-11.24
GO:0000083	Biological Processes	regulation of transcription involved in G1/S transition of mitotic cell cycle	5	22.73	-10.37	-7.40
GO:0001216	Molecular Functions	bacterial-type RNA polymerase transcriptional activator activity, sequence-specific DNA binding	3	13.64	-8.20	-5.49
GO:0001228	Molecular Functions	DNA-binding transcription activator activity, RNA polymerase II-specific	7	31.82	-7.12	-4.53
GO:0051301	Biological Processes	cell division	6	27.27	-4.90	-2.45
GO:1990841	Molecular Functions	promoter-specific chromatin binding	3	13.64	-4.74	-2.30
GO:0032993	Cellular Components	protein-DNA complex	3	13.64	-3.12	-0.75
GO:0006979	Biological Processes	response to oxidative stress	3	13.64	-2.11	0.00

Top 8 clusters with their representative enriched terms (one per cluster). "Count" is the number of genes in the user-provided lists with membership in the given ontology term. "%" is the percentage of all of the user-provided genes that are found in the given ontology term (only input genes with at least one ontology term annotation are included in the calculation). "Log10(P)" is the p-value in log base 10. "Log10(q)" is the multi-test adjusted p-value in log base 10.

Table III:

KEGG Analysis of E2F1/4/5/8 and co-expression genes in Patients with DLBCL using Metascape

GO	Category	Description	Count	%	Log10(P)	Log10(q)
hsa04110	KEGG Pathway	Cell cycle	10	45.45	-17.31	-14.61
hsa04350	KEGG Pathway	TGF-beta signaling pathway	5	22.73	-7.96	-5.57
hsa05203	KEGG Pathway	Viral carcinogenesis	5	22.73	-6.07	-3.95
hsa05222	KEGG Pathway	Small cell lung cancer	4	18.18	-6.04	-3.95
hsa05222	KEGG Pathway	Small cell lung cancer	4	18.18	-6.04	-3.95

Top 4 clusters with their representative enriched terms (one per cluster). "Count" is the number of genes in the user-provided lists with membership in the given ontology term. "%" is the percentage of all of the user-provided genes that are found in the given ontology term (only input genes with at least one ontology term annotation are included in the calculation). "Log10(P)" is the p-value in log base 10. "Log10(q)" is the multi-test adjusted p-value in log base 10.

Figures

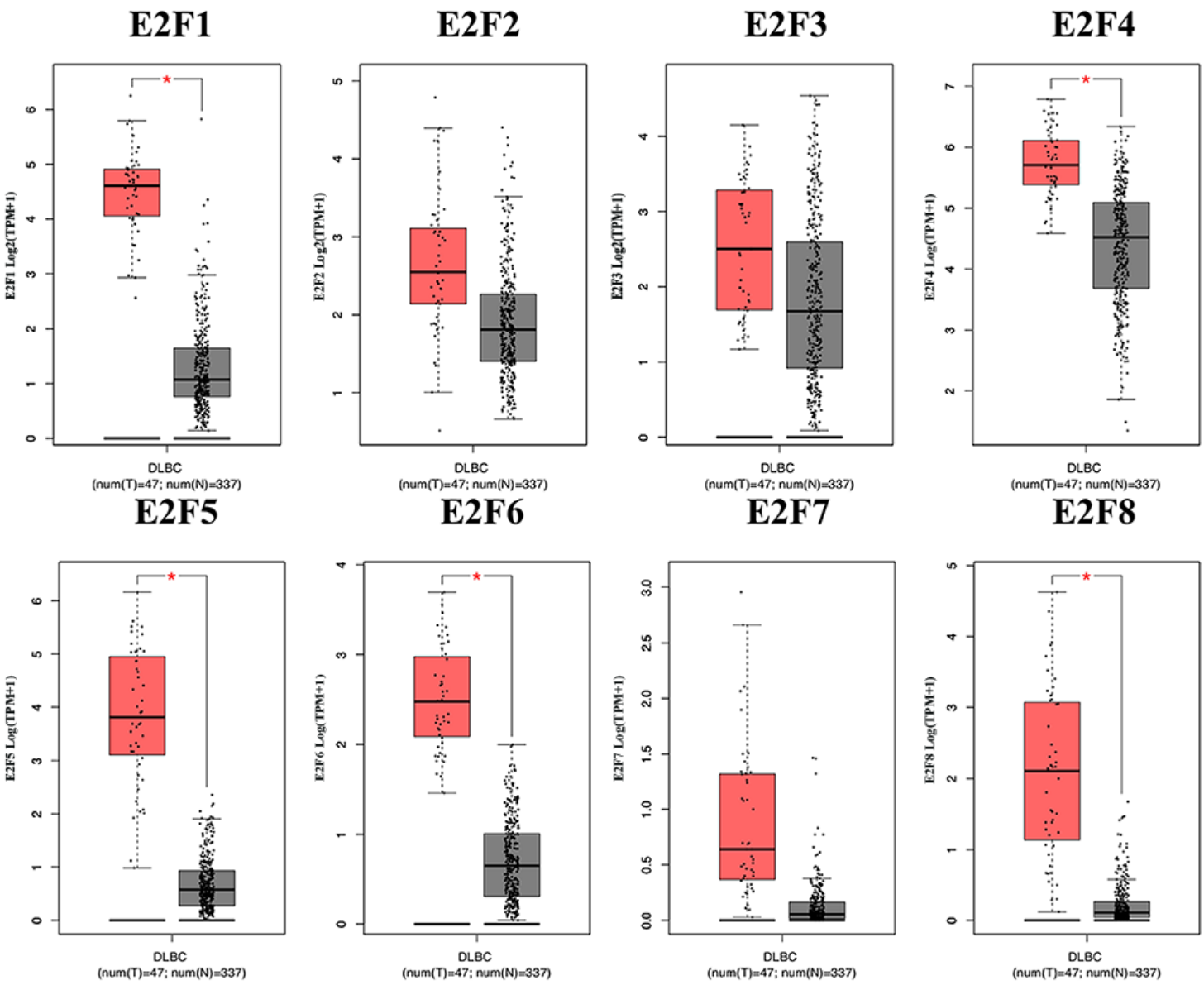

Figure 1

The E2F family members' mRNA level in DLBC patients (GEPIA). The box plots compare the differential expression of E2Fs in tumor tissues and normal tissues; *Indicate that the p-value <0.05.

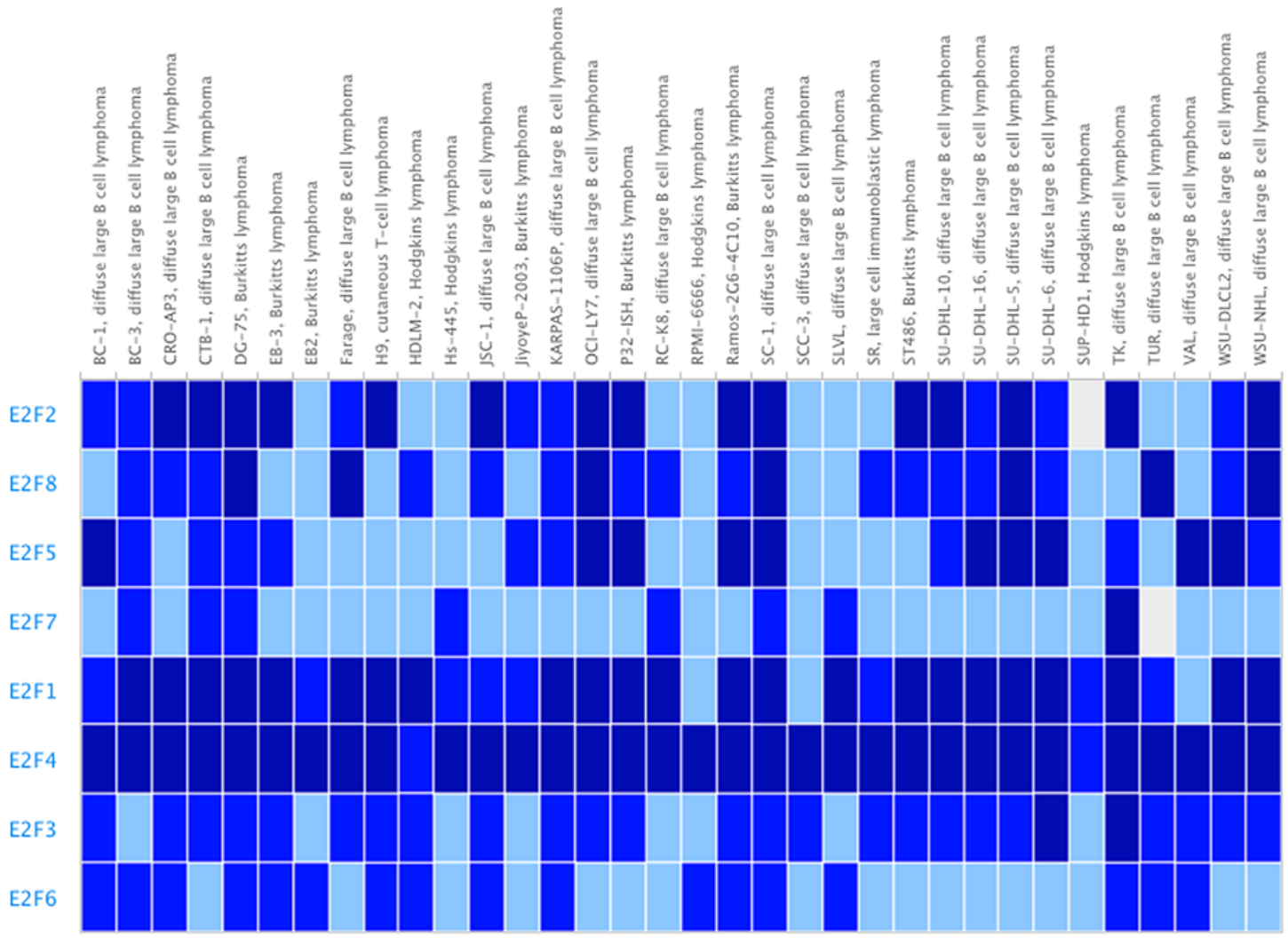

Figure 2

The mRNA level of E2Fs in Lymphoma Cell Lines of EMBL-EBI

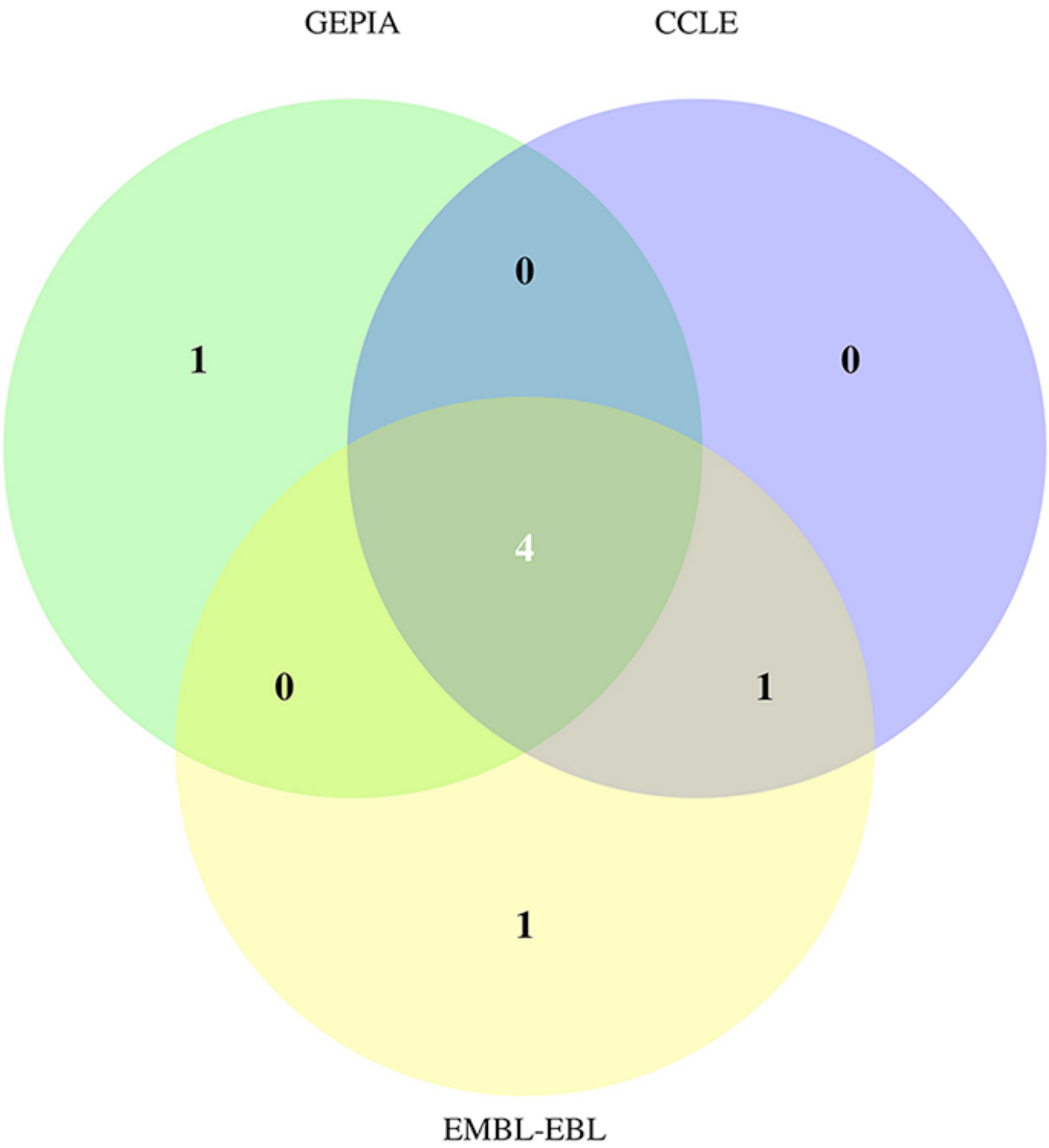

Figure 3

Intersection of the predicted results in the GEPIA, CCLE and EMBL-EBI.

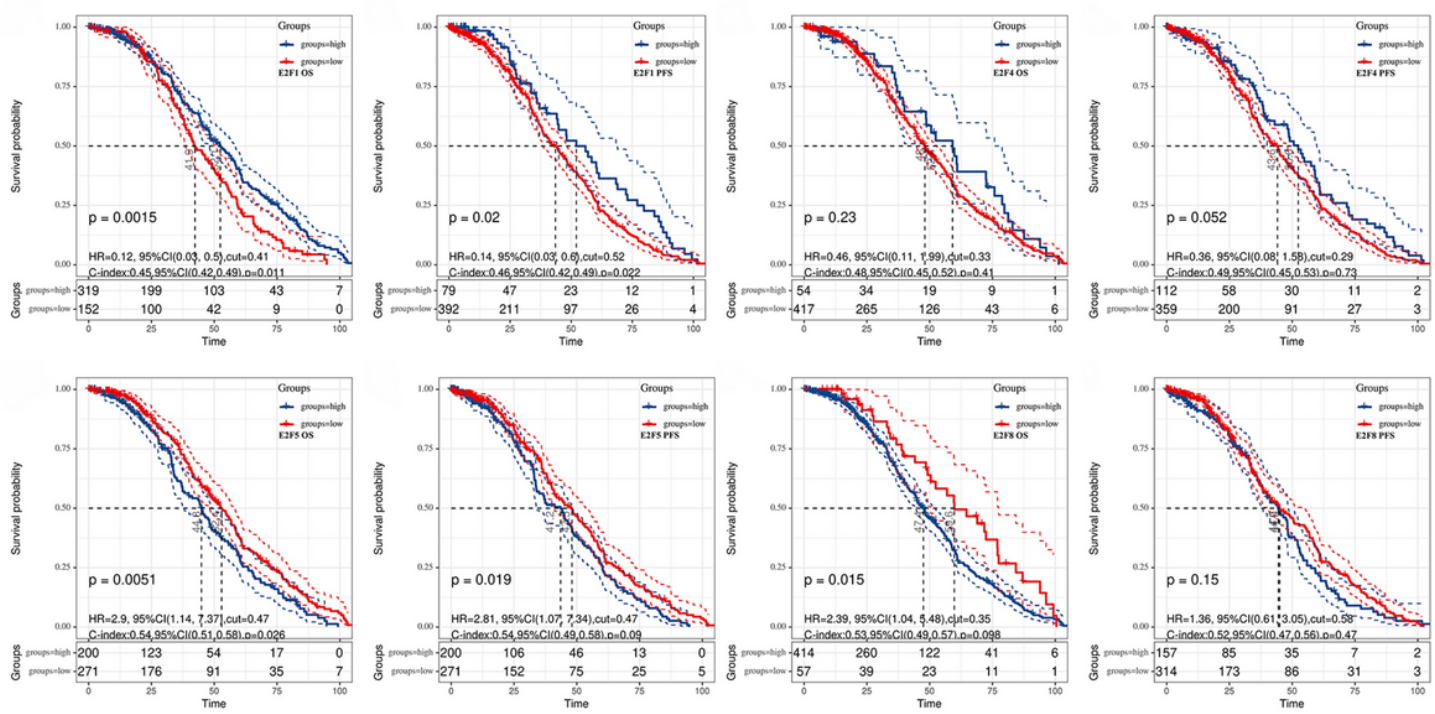

Figure 4

The OS time and PFS time survival curves of DLBCL patients stratified by E2F1/4/5/8 expression level. p<0.05 indicates that the results are statistically significant.

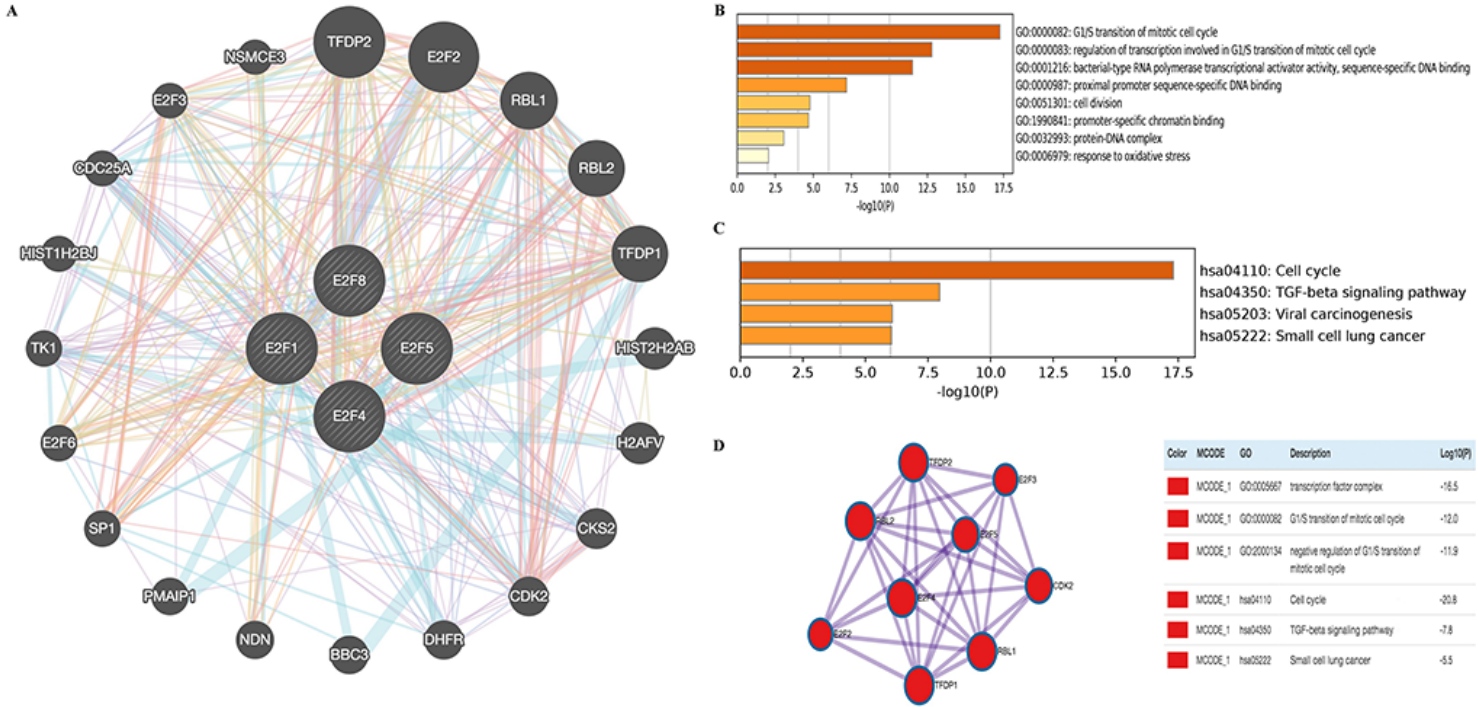

Figure 5

Functional enrichment analysis of E2F1/4/5/8 and co-expression genes in patients with DLBCL (A) Gene–gene interaction network among E2F1/4/5/8. Each node represents a gene. (B) GO analysis of E2F1/4/5/8 and co-expression genes in Patients with DLBCL using Metascape. (C) KEGG analysis of E2F1/4/5/8 and co-expression genes in Patients with DLBCL using Metascape. (D) The MCODE networks identified for individual gene lists using Metascape.

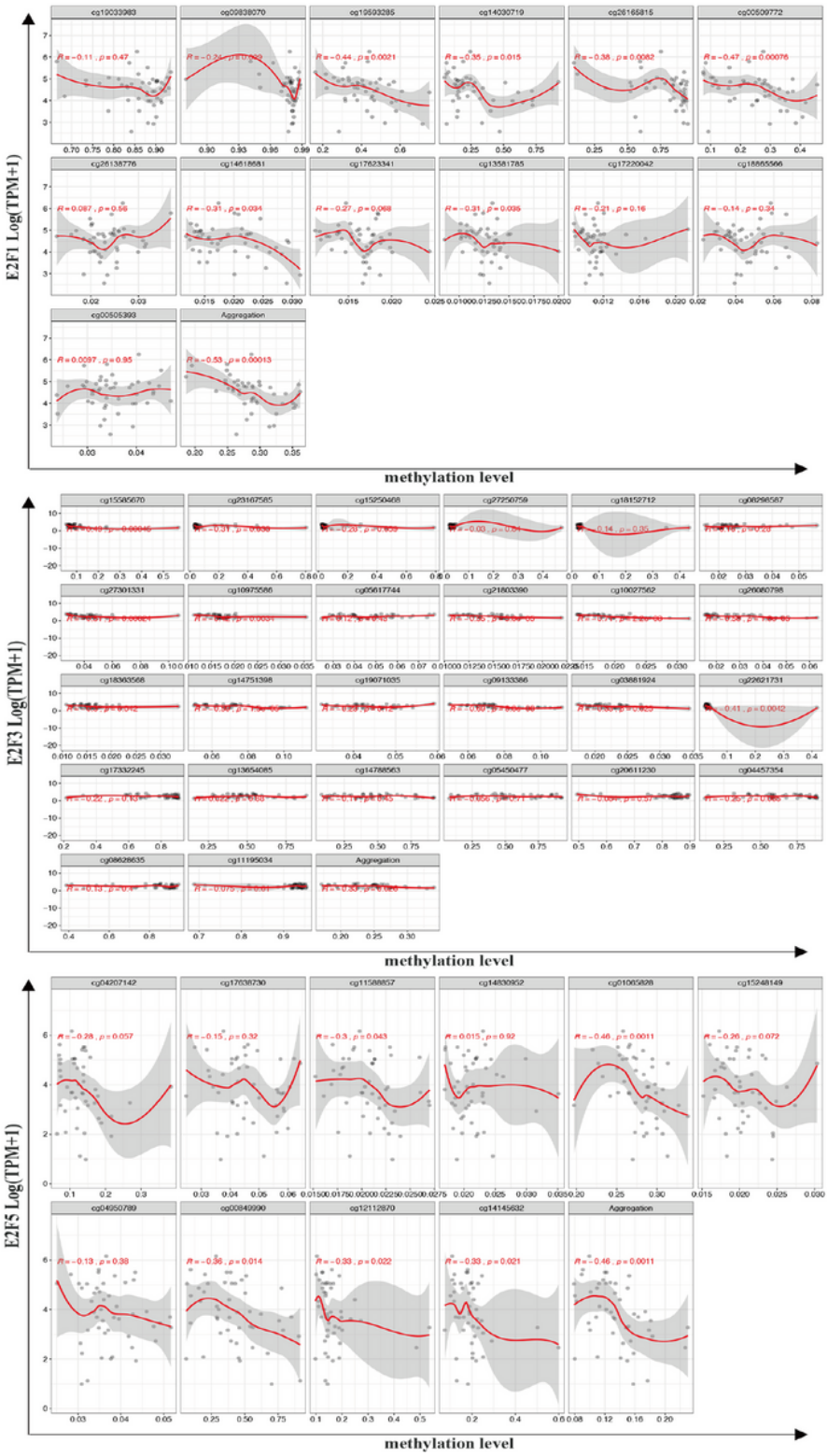

Figure 6

The methylation of E2F1/3/5 in the SMARTAPP database.

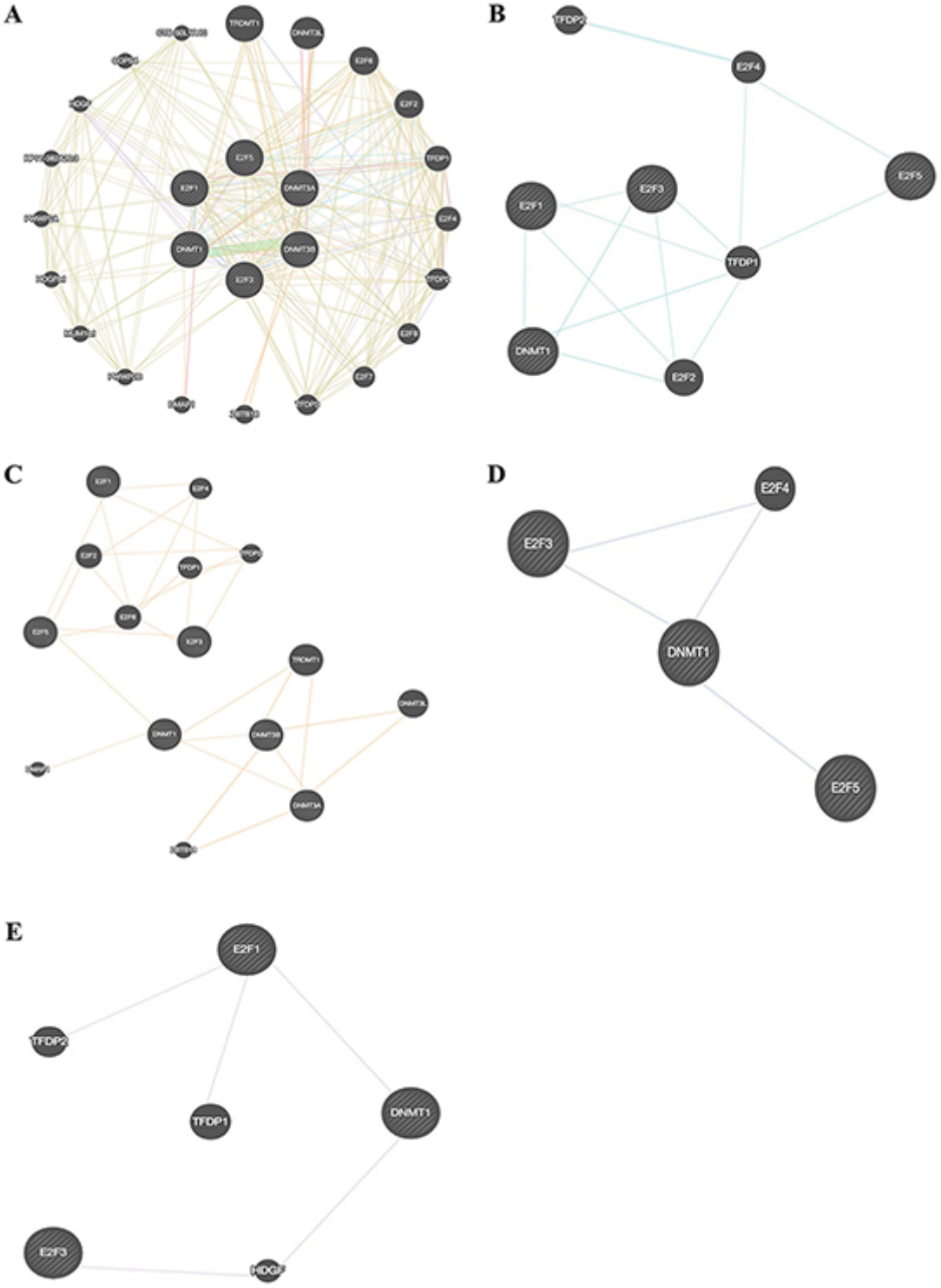

Figure 7

Interaction analysis of the E2F1/3/5 with the DNA methylation associated genes (A) A gene–gene interaction network for the E2F1/3/5, DNMT1/3A/3B in the GeneMANIA database. (B) The genes participate in the same reaction within a pathway. (C) The genes predict functional relationships between

protein interactions. (D) The genes share the same localization. (E) The genes correlate in terms of expression.

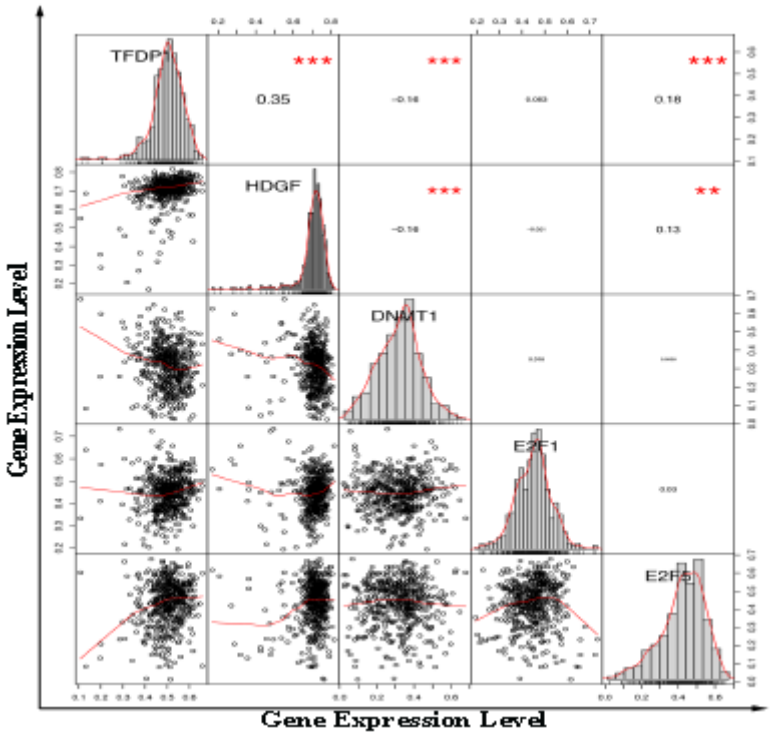

Figure 8

Pearson correlations between the TFDP1, HDGF, DNMT1, E2F1 and E2F5 in GSE31312.*p < 0.05**p < 0.01***p <0 .001

Supplementary Files

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

- SupplementaryFigure1.pdf
- SupplementaryFigure2.pdf
- SupplementaryFigure3.pdf