

Overexpression of Sideroflexin1 (SFXN1) Predicts Worse Overall Survival in Lung Adenocarcinoma

Wenting Liu

Tianjin Medical University

Kaiting Jiang

Tianjin Medical University

Jingya Wang

Tianjin Medical University

Ting Mei

Tianjin Medical University

Xuehan Fan

Tianjin Medical University

Yujia Zhang

Tianjin Medical University

Dingzhi Huang (✉ dingzhih72@163.com)

Tianjin Tumor Hospital

Research

Keywords: Sideroflexin1 (SFXN1), Lung Adenocarcinoma (LUAD), The Cancer Genome Atlas (TCGA), Prognostic Biomarker

Posted Date: January 23rd, 2020

DOI: <https://doi.org/10.21203/rs.2.21665/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Mitochondria are the core organelle of cellular energy and metabolism, also closely related to the growth and survival of cancer cells. This research aimed to discover and evaluate a promising prognostic biomarker, a novel mitochondrial gene—Sideroflexin1 (SFXN1), for lung adenocarcinoma (LUAD).

Results

Analysis of 594 samples from The Cancer Genome Atlas (TCGA) Lung Cancer Cohort, this present study indicated that SFXN1/2/4/5 were overexpressed in LUAD tissue compared with normal lung tissue whereas only the SFXN1 was associated significantly with the overall survival (OS) from the Kaplan-Meier survival analysis ($p=0.00093$), suggesting that SFXN1 could be a potential prognostic biomarker. Furthermore, the logistic regression of the correlation of clinicopathological features and the expression status of SFXN1 from 522 patients with LUAD in TCGA revealed that the expression level of SFXN1 was related to Eastern Cooperative Oncology Group (ECOG), clinical stage, tumor size and lymph nodes invasion (all $p<0.05$). Additionally, overexpression of SFXN1 remained associated positively and independently with a worse prognosis after the multivariate Cox regression ($HR=1.486$, $p=0.022$). Functionally, SFXN1 involved in the cell cycle, specifically in DNA replication and meiosis, ATPase activity and folate-dependent one-carbon metabolism from the outcomes of Eastern Cooperative Oncology Group (KEGG) and Gene Ontology (GO) enrichment.

Conclusion

SFXN1 might promote the proliferation and metabolism of cancer cells and function as a prognostic indicator for LUAD, which may contribute to the development of targeted therapy and the emergence of metabolism-based treatment in clinical practice.

Background

Lung cancer remains the leading cause of cancer-related morbidity worldwide in the light of the data released in 2018[1]. Non-small cell lung cancer (NSCLC) accounts for more than 85%[2], consisting of adenocarcinoma, squamous cell carcinoma and large cell carcinoma in subtype histology. Among NSCLC, adenocarcinoma is the most prevalent subtype with an increasing trend of incidence[3, 4]. Fortunately, the enormous development and extensive application of genetic diagnosis and the molecular targeted therapy improved the whole effects of treatment, especially in LUAD, such as the advent of epidermal growth factor receptor-tyrosine kinase inhibitors (EGFR-TKIs)[5], which leads to the emergence of “precision medicine” based on tumor molecular alteration profile rather than tumor histology or tumor site in anatomy[6]. Nevertheless, the 5-year overall survival rate is less than 20%[7] owing to the resistance to these targeted drugs and the lack of recognition of more mutant driver genes. Therefore, at

present, it is imperative to identify other essential genes and prognostic biomarkers for early diagnosis, customized treatment and better prognosis.

Oncologists have never stopped their step and passion for scientific research of cancer metabolism since almost a century ago when Otto Warburg first put forward the aberrant metabolic phenomenon of cancer cells, so-called “The Warburg Effect”[8, 9]. Moreover, in 2011, Hanahan and Weinberg[10] updated the version of “Hallmarks of Cancer” to include four more “emerging hallmarks”, of which “deregulating cellular energetics” is one. Mitochondria are pivotal regulators of cellular energy and metabolism so that they play a crucial role in sustaining the growth and survival of cancer cells[11]. Obviously, mitochondrial dysfunction is a characteristic of most cancers and strongly related to the abnormal metabolism of tumor cells[4] but it has not been completely clear and no relevant targeted medicines, so a lot of work is still needed to be down, for instance, the mitochondria-related genes, proteins and their interactions. Given the widespread presence of mitochondrial malfunction in lung cancer[4, 12]but the void of the application of relevant treatment in patients, identification of the mitochondria-related genetic altered profile is not only a must to comprehensively understand the mechanism of tumor metabolism, but also a base for developing mitochondrial metabolism-based therapy in clinical practice in the near future.

The SFXN family are conservative human mitochondrial proteins[13], comprised of SFXN1 (also namely Tricarboxylate carrier protein[14]), SFXN2, SFXN3, SFXN4 and SFXN5. SFXN2 is expressed mainly in kidney, SFXN3 in retina, SFXN4 in pancreas and SFXN5 primarily in brain while SFXN1 is expressed ubiquitously as a multi-pass protein, existing on the inner membrane of mitochondria [15]. Recently, Nora Kory et al. verified that SFXN1 serves as a mitochondrial serine transporter in the process of one-carbon metabolism and expresses high in many cancers[15]. However, to our best knowledge, there is no article about the research of SFXN1 or SFXN family in any type of cancer.

To help elucidate the possible relevance between SFXN family and LUAD, we performed an integral analysis of SFXN family, particularly SFXN1, through TCGA, GEO, and Oncomine databases, in the hope of providing useful insights into the mitochondrial genic etiology of LUAD.

Table 1 Information of GEO series included in this study				
GEO series	Sources	Normal(N)	Tumor(N)	Platforms
GSE10072	Landi Lung,2008	49	58	[HG-U133A] Affymetrix Human Genome U133A Array
GSE31210	Okayama Lung, 2012	20	226	[HG-U133-Plus-2]Affymetrix Human Genome U133 Plus 2.0 Array
GSE32863	Selamat Lung, 2012	58	58	Illumina HumanWG-6 v3.0 expression beadchip
GSE7670	Su Lung, 2007	31	32	[HG-U133A] Affymetrix Human Genome U133A Array

Results

Expression levels of SFXN complex and associations between SFXN family and OS in TCGA-LUAD cohort

Expression levels of five members of SFXN family were individually assessed in the LUAD patients from TCGA database. The results demonstrated that all SFXN complex components, other than SFXN3, were significantly overexpressed in LUAD tissues compared with normal lung tissues (Fig. 1, A-E).

To identify the promising prognostic biomarkers among SFXN family, the associations between the SFXN family and overall survival in the TCGA-LUAD cohort were examined. The outcomes of Kaplan-Meier survival analysis (Fig. 2, A-E) illustrated that only SFXN1-high was significantly correlated to the decrease of survival times compared with SFXN1-low ($p=0.00093$). Hence, SFXN1 was considered as a potential marker and focused on in the following research.

Validation of the overexpression of SFXN1 in LUAD from GEO

For further validation, the different expressions of SFXN1 between LUAD and normal lung tissues are analyzed respectively from GSE7670, GSE10072, GSE31210, GSE32863. The details of aforementioned GEO series were summarized in Table 1. As is shown in Figure 3 A-D, SFXN1 was significantly upregulated among LUAD patients in all cohorts (all $P<0.01$). Meanwhile, a meta-analysis of SFXN1 expression in the above cohorts was performed with the threshold by $p\text{-Value} \leq 1E-4$, fold change ≥ 2 and top 10% gene rank in the Oncomine database (Fig.3E).

Table 2 Clinicopathologic of 522 LUAD patients from TCGA

Clinical characteristics	Total(n=522)	%	Clinical characteristics	Total(n=522)	%
Age/years			Clinical stage		
Mean (SD)	65.3(10.0)		Stage I	279	53.4
Median [Min,Max]	66[33,88]		Stage II	124	23.8
Gender			Stage III	85	16.3
Female	280	53.6	Stage IV	27	5.2
Male	242	46.4	Missing	7	1.3
Ethnic origin			Tumor size		
White	392	75.1	T1	172	33.0
Non-White	62	11.9	T2	281	58.8
Missing	68	13.0	T3	47	9.0
Smoking history			T4	19	3.6
Non-smoker	75	14.3	Missing	3	0.6
Current or former smoker	433	83.0	Lymph nodes invasion		
Missing	14	2.7	Negative	335	64.2
ECOG			Positive	175	33.5
≤1	195	37.4	Missing	3	2.3
≥1	24	4.6	Distant metastasis		
Missing	303	58.0	Negative	353	67.6
Cancer status			Positive	25	4.8
Tumor free	314	60.1	Missing	144	27.6
With tumor	110	21.1			
Missing	98	18.8			

Patient characteristics from TCGA

522 LUAD patients with both clinical characteristics and gene expression data were downloaded from TCGA database. As is shown in Table 2, among these cases, 53.6%(n=280) were females and 46.4%(n=242) are males. The average age at diagnosis was 65.3 years old and the vast majority of these cases (n=433, 83.00%) have a smoking history. The number of patients with ECOG performance status less than or equal 1 is 195(37.4%) while more than 1 is 24(4.6%). In this cohort, one third (n=175, 33.5%)

cases had lymph nodes invasion and only 25 patients (4.8%) metastasized; most research objects (67.6%, n=353) have tumors in T1 or T2. Accordingly, about half of the entire cases (n=279, 53.4%) was found in stage I, 124 (23.8%) in stage II, 85 (16.3%) in stage III and merely 27 (5.2%) in stage IV.

Association with SFXN1 expression and clinicopathologic variables

We analyzed 535 LUAD and 59 normal samples of SFXN1 expression data across all patients' clinicopathologic characteristics from TCGA. Fig.4(A-G) showed that increased expression of SFXN1 associated significantly with the age ($p=0.04$), ECOG ($p=0.015$), clinical stage ($p=0.0014$), tumor size ($p=0.027$) and lymph nodes invasion ($p=0.014$) while no obvious difference was observed with distant organ metastasis ($p=0.054$) and smoking history ($p=0.071$).

Univariate logistic regression analysis (Table 3) suggested that SFXN1 expression, as a categorical dependent variable (based on median expression value of 6.064), was correlated to poor prognostic clinicopathologic characteristics, especially the ECOG performance status (Odds Ratio, OR=2.722 for $\text{ECOG} \leq 1$ vs. >1 , $p=0.034$), clinical stage (OR=2.013 for III vs. I, $p=0.006$; OR = 1.563 for II vs. I, $p=0.042$), tumor size (OR=1.626 for T2 vs. T1, $p=0.014$; OR=2.960 for T4 vs. T1, $p=0.036$) and lymph nodes invasion (OR=1.516 for negative vs. positive, $p=0.028$). These results elucidated that LUAD patients with SFXN1-high is inclined to progress to a more advanced stage than those with SFXN1-low.

Table 3 The association of SFXN1 expression levels ^a with clinicopathologic characteristics (logistic regression)

Clinical characteristics	Comparsion items	Number of patients [N]	Odds ratio(95%CI)	p- value
Age years	Age≤60 vs.Age>60	494	0.727[0.496-1.063]	0.101
Gender	Male vs.Female	513	1.358[0.959-1.925]	0.085
Ethnic origin	White vs.Non-White	446	1.000[0.579-1.723]	1.000
Smoking history	Non-smoker vs.Smoker	500	1.382[0.842-2.287]	0.202
ECOG performance status	ECOG≤1 vs.ECOG>1	217	2.722[1.120-7.313]	0.034
Cancer status	Tumor free vs.With tumor	424	1.260[0.813-1.957]	0.302
Clinical stage	II vs.I	395	1.563[1.018-2.409]	0.042
	III vs.I	358	2.013[1.228-3.338]	0.006
	IV vs.I	300	2.084[0.925-4.912]	0.081
Tumor size	T2 vs.T1	444	1.626[1.106-2.400]	0.014
	T3 vs.T1	215	1.103[0.571-2.114]	0.767
	T4 vs.T1	187	2.960[1.113-8.773]	0.036
Lymph nodes invasion	Negative vs.Positive	501	1.516[1.047-2.204]	0.028
Distant metastasis	Negative vs.Positive	369	1.862[0.817-4.509]	0.148

^a: Categorical dependent variable, greater or less than the median expression level (6.064). The bold numbers are p values less than 0.05.

Cox regression analysis association with clinicopathologic variables and OS

At univariate analysis, over expressions of SFXN1 was associated positively with a worse prognosis (Hazard Ratio, HR: 1.582; 95%CI:1.148–2.180; p=0.005). Other related clinicopathologic variables included clinical stage, tumor size and lymph nodes invasion (Table 4). After multivariate analysis

(Table4 & Fig.5), the expression status of SFXN1 remained independently correlated to the overall survival, with an HR of 1.486 (CI:1.059–2.085, p=0.022), along with clinical stage.

Table 4 The correlation of clinicopathologic characteristics with OS (Cox regression)

Parameters	Univariate analysis			Multivariate analysis		
	HR	95%CI	p-value	HR	95%CI	p-value
Age	1.001	0.982-1.020	0.929	1.010	0.991-1.029	0.297
Gender	1.001	0.699-1.434	0.996	0.831	0.574-1.204	0.329
Clinical stage	1.645	1.397-1.937	0.000	1.676	1.136-2.474	0.009
Tumor size	1.623	1.310-2.011	0.000	1.191	0.938-1.512	0.152
Lymph nodes invasion	2.721	1.893-3.911	0.000	1.365	0.775-2.405	0.281
Distant metastasis	1.681	0.924-3.060	0.089	0.542	0.201-1.463	0.227
SFXN1 expression	1.582	1.148-2.180	0.005	1.486	1.059-2.085	0.022

KEGG/GO identify SFXN1-related signaling pathways

To further explore the underlying function of SFXN1 in LUAD, we conducted KEGG and GO analysis on SFXN1 interactive genes to screen closely relevant signaling pathways and biological functions. The enrichment bubble plot figure from GO (Fig. 6A) demonstrated that each of the three GO terms shows the top ten significant items associated with SFXN1 related genes. Additionally, the enrichment of KEGG (Fig.6B) exhibited that cell cycle, oocyte meiosis, RNA transport, DNA replication, progesterone-mediated oocyte maturation, ribosome biogenesis in eukaryotes, Fanconi anemia pathway, homologous recombination, nucleotide excision repair, mismatch repair and one carbon pool by folate and so forth were the most enriched pathways(all p<0.05).

Discussion

With the expanding global population and longer life expectation, the staggering incidence of cancer in 2030 was predicted up to 20.3 million worldwide, which is set to become the main cause of morbidity and mortality in the next decade in every area of the world[16]. Specifically, the highest lifetime cumulative risk of incidence of cancers is lung cancer, the most common type of cancer death in 94 countries, in both high and low HDI (Human Development Index) regions[16]. What’s worse, despite many years of efforts, the metabolic reprogramming of cancer cells is far beyond fully understanding and there has been no treatment towards this mechanism in clinical practice[17] due to the complicated feature of aberrant metabolism of cancer. Recently, it has been reported that compared with normal tissues, mRNA levels of SFXN1 was much higher in virtually all kinds of human cancers[15]. Yet, so far as we know, there has

been no published paper regarding the differential expression levels of SFXN family and the role of SFXN1 in cancer. Therefore, data and paper mining were applied to identify the potential prognostic biomarkers within the complex in LUAD in this research.

In this present study, comprehensive and comparative analysis of the compound genes is a more effective method to explore the interrelations among individual genes and evaluate novel biomarkers. Targeting one of these genes not only meets the principle of "Precision Medicine", but also may lead to a synergistic effect maximally in the real world. Our results indicated that the expression of all SFXN complex components, other than SFXN3, was increased markedly in LUAD but SFXN1 was the only one that affected OS significantly to LUAD patients. Previously, it has been reported that all SFXN family members are present in pancreatic islet cells while SFXN3 is the most important carrier molecule in the differentiation and regeneration of pancreatic β -cells[18]. Similarly, although many of the family members are expressed high in LUAD, the one exerts vital influence during the development and progression of LUAD is SFXN1. Thus, our attention was fixed on SFXN1 for further exploration as it appeared to be a potential prognostic biomarker.

Furthermore, our investigation revealed that SFXN1 expression was related to age, ECOG performance status, clinical stage, tumor size and lymph nodes invasion while it showed that SFXN1 made no evident difference in metastasis, merely a little bit above the threshold ($p=0.054$, the cutoff $p=0.05$). As SFXN1-high was correlated to advanced stage, large tumor size and positive lymph nodes invasion, it is reasonable to attribute this negative statistical outcome to the root cause of small sample size of metastasis ($n=25$, 4.8%). Additionally, the univariate and multivariate Cox regression proved that SFXN1 remained associated with OS, which affirmed that this gene is an independent prognostic indicator for LUAD. Besides, the KEGG pathway and GO biological process of SFXN1 interactive genes in this research illustrated that SFXN1 involves in cell cycle, including DNA replication and meiosis. Absolutely, due to their own DNA (mtDNA) in eukaryotes, mitochondria also participate in regulating the throughout cell cycle and mitosis[4]. As tumor cells proliferate rapidly, the mtDNA mutations generate the more favorable metabolic profile to accelerate the progression and metastasis of cancer[19]. Hence, SFXN1 exacerbates the progression of LUAD by means of promoting the proliferation of the cancer, resisting the programmed death of cells and activating the pathways of invasion and metastasis. Consequently, the alteration of SFXN1 expression levels in LUAD may be a critical factor and it is worthwhile to probe whether it is a driver or a passenger oncogene.

Functionally, the GO term of molecular function manifested that SFXN1 has an extensive function in ATPase activity on the membrane of mitochondria. According to other reports, the SFXN1 gene mutated flexed-tail (f/f) mice are susceptible to pathologic intra-mitochondrial iron accumulation in erythrocytes, which implies that SFXN1 facilitates the transportation of iron into mitochondria[20]—Mingzhu Tang et.al. also discovered SFXN1, a mitochondria iron transporting protein, carries the overload free iron to mitochondria, causing the impairment of mitochondria and the subsequent cardiomyocytes hypertrophy in the long run[21]. Additionally, Sousa L and colleagues demonstrated that iron overload patients exhibit higher ATPase activity in their erythrocytes[22]. Accordingly, SFXN1 may promote the transportation of

iron by virtue of enhancing the ATPase activity although further experimental validation is required. Moreover, metabolic reprogramming is the major part observed and discovered amongst pathways of central carbon metabolism during the last decade in cancer cells[23]. In our study, KEGG enrichment also demonstrated that SFXN1 involved in the folate-dependent one-carbon metabolism, consistent with the elegant work that SFXN1 is the main SFXN carrier protein for transferring serine into mitochondria during the process of one-carbon metabolism[15].

Conclusion

For the first time, this present investigation revealed the disparate expression levels of SFXN family between LUAD and normal samples and analyzed the association of SFXN1 with the clinicopathologic characteristics and the overall survival rate of 522 LUAD patient from TCGA, which elucidated SFXN1 plays an indispensable role in the progression of LUAD, especially in tumor proliferation and metabolism. As this study was only conducted via bioinformatics analysis, experimental studies will be performed in order to verify the results abovementioned. Nevertheless, the information herein indeed provides the primary foundation for metabolic prognostic biomarkers of the initiation, progression and metastasis of LUAD and translationally, an anti-cancer strategy targeting at metabolic genes on mitochondria and its pathways in LUAD. In a nutshell, SFXN1 may serve as a prognostic biomarker for LUAD.

Methods

RNA-sequencing data and bioinformatics analysis of patients

We downloaded the mRNA-sequencing data of 522 lung-adenocarcinoma patients, diagnosed from 1991 to 2013, including 59 paracancerous normal lung tissue samples and 535 lung adenocarcinoma samples (594 samples in total, Workflow Type: Htseq-FPKM) from the official website of TCGA (<https://cancergenome.nih.gov/>) in October 2019. Meanwhile, the clinical information of these patients was obtained with the TCGA biolinks package[24]. Then, the comparative analysis of differential expression levels of SFXN family between LUAD and normal tissues was operated by Limma package[25] and the association with SFXN family and OS was performed via Kaplan-Meier methods. The high and low SFXN1 expression levels was classified based on median expression value (6.064).

Four expression microarray series GSE7670[26], GSE10072[27], GSE31210[28] and GSE32863[29] were downloaded from GEO (<https://www.ncbi.nlm.nih.gov/geo/>) using GEOquery package[30] to verify the disparate expression levels of the SFXN1 gene among tumor and non-tumor samples in LUAD patients. In addition, a meta-analysis of the abovementioned four lung adenocarcinoma cohorts was tested on the Oncomine database (<https://www.oncomine.org/>). Platforms and samples of these GEO series were described in Table1.

SFXN1 gene related pathway/biological process enrichment analysis

With the spearman correlation analysis, the top 500 SFXN1 correlated genes were screened from the downloaded TCGA lung adenocarcinoma gene expression data, according to the absolute p value ($p < 0.05$) of the correlation coefficient. To understand the biological function of SFXN1, the clusterProfiler[31] package was utilized to conduct KEGG and GO enrichment analysis on the relevant 500 genes. GO terms and KEGG pathways with P value < 0.05 were considered statistically significant.

Statistical analysis

Statistical analysis of all data was performed with R (v.3.6.1). Statistical methods to analyze the correlation between SFXN1 expression and the main clinicopathological features of LUAD patients depend on the number of subgroups: samples divided into 2 groups were analyzed using the Wilcoxon signed rank sum test whilst more than 2 groups with the Kruskal-Wallis test. Logistic regression was conducted to test the association of clinicopathological characteristics and the expression levels of SFXN1; Cox regression was operated to compare and determine the independent factors that play a direct and significant role in the prognosis of LUAD patients. A two-tailed $P < 0.05$ was considered statistically significant for all tests.

Declarations

Ethics approval and consent to participate

Not available

Consent for publication

Not available

Availability of data and materials

Please contact the author for data requests.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Nature Science Foundation of China (grant number 81572321) awarded to DZ. Huang.

Authors' contributions

LWT and JKT designed the study. LWT collected and analyzed the data; JKT wrote the manuscript. WJY and MT participated in analysis and FXH and ZYJ revised the manuscript. HDZ directed and supervised the whole work. All participants have read and approved for this manuscript.

Acknowledgments

We thank the TCGA Research Network: <http://cancergenome.nih.gov/> and the GEO Network <https://www.ncbi.nlm.nih.gov/geo/> as our original data was based upon these platforms.

Abbreviations

SFXN	Sideroflexin	LUAD	Lung Adenocarcinoma
TCGA	The Cancer Genome Atlas	GEO	Gene Expression Omnibus dataset
OS	Overall Survival	KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology	ECOG	Eastern Cooperative Oncology Group
NSCLC	Non-Small Cell Lung Cancer	CI	Confidence Interval

References

1. RL, S., M. KD, and J. A, *Cancer statistics, 2018*. CA: a cancer journal for clinicians, 2018. **68**(1): p. 7-30.
2. Molina, J.R., et al., *Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship*. Mayo Clin Proc, 2008. **83**(5): p. 584-94.
3. X, Z., et al., *Prognostic significance of the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification of stage I lung adenocarcinoma: A retrospective study based on analysis of 110 Chinese patients*. Thoracic cancer, 2017. **8**(6): p. 565-571.
4. FE, L. and S. R, *Mitochondrial dynamics: biology and therapy in lung cancer*. Expert opinion on investigational drugs, 2014. **23**(5): p. 675-92.
5. Nan, X., et al., *EGFR TKI as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer*. Oncotarget, 2017. **8**(43): p. 75712-75726.
6. Jackson, S.E. and J.D. Chester, *Personalised cancer medicine*. Int J Cancer, 2015. **137**(2): p. 262-6.
7. Boolell, V., et al., *The Evolution of Therapies in Non-Small Cell Lung Cancer*. Cancers (Basel), 2015. **7**(3): p. 1815-46.
8. O., W., *on the origin of cancer cells*. Science, 1956. **123**(3191): p. 309-14.
9. Warburg, O., F. Wind, and E. Negelein, *THE METABOLISM OF TUMORS IN THE BODY*. J Gen Physiol, 1927. **8**(6): p. 519-30.
10. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell, 2011. **144**(5): p. 646-74.

11. Momcilovic, M., et al., *In vivo imaging of mitochondrial membrane potential in non-small-cell lung cancer*. Nature, 2019. **575**(7782): p. 380-384.
12. HS, N., et al., *Mitochondria in chronic obstructive pulmonary disease and lung cancer: where are we now?* Biomarkers in medicine, 2017. **11**(6): p. 475-489.
13. Zheng, H., et al., *Molecular cloning and characterization of a novel human putative transmembrane protein homologous to mouse sideroflexin associated with sideroblastic anemia*. DNA Seq, 2003. **14**(5): p. 369-73.
14. Xi, D., et al., *Molecular cloning, sequence identification and tissue expression profile of three novel genes Sfxn1, Snai2 and Cno from Black-boned sheep (Ovis aries)*. Mol Biol Rep, 2011. **38**(3): p. 1883-7.
15. N, K., et al., *SFXN1 is a mitochondrial serine transporter required for one-carbon metabolism*. Science (New York, N.Y.), 2018. **362**(6416): p. undefined.
16. Bray, F., et al., *Global cancer transitions according to the Human Development Index (2008-2030): a population-based study*. Lancet Oncol, 2012. **13**(8): p. 790-801.
17. Frezza, C., *Metabolism and cancer: the future is now*. Br J Cancer, 2019.
18. Yoshikumi, Y., et al., *Roles of CTPL/Sfxn3 and Sfxn family members in pancreatic islet*. J Cell Biochem, 2005. **95**(6): p. 1157-68.
19. Whitehall, J.C. and L.C. Greaves, *Aberrant mitochondrial function in ageing and cancer*. Biogerontology, 2019.
20. Fleming, M.D., et al., *A mutation in a mitochondrial transmembrane protein is responsible for the pleiotropic hematological and skeletal phenotype of flexed-tail (f/f) mice*. Genes Dev, 2001. **15**(6): p. 652-7.
21. Tang, M., et al., *Ferritinophagy activation and sideroflexin1-dependent mitochondria iron overload is involved in apelin-13-induced cardiomyocytes hypertrophy*. Free Radic Biol Med, 2019. **134**: p. 445-457.
22. Sousa, L., et al., *Effects of Iron Overload on the Activity of Na,K-ATPase and Lipid Profile of the Human Erythrocyte Membrane*. PLoS One, 2015. **10**(7): p. e0132852.
23. Seth Nanda, C., et al., *Defining a metabolic landscape of tumours: genome meets metabolism*. Br J Cancer, 2019.
24. Colaprico, A., et al., *TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data*. Nucleic Acids Res, 2016. **44**(8): p. e71.
25. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies*. Nucleic Acids Res, 2015. **43**(7): p. e47.
26. Su, L.J., et al., *Selection of DDX5 as a novel internal control for Q-RT-PCR from microarray data using a block bootstrap re-sampling scheme*. BMC Genomics, 2007. **8**: p. 140.
27. Landi, M.T., et al., *Gene expression signature of cigarette smoking and its role in lung adenocarcinoma development and survival*. PLoS One, 2008. **3**(2): p. e1651.

28. Okayama, H., et al., *Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas*. *Cancer Res*, 2012. **72**(1): p. 100-11.
29. Selamat, S.A., et al., *Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression*. *Genome Res*, 2012. **22**(7): p. 1197-211.
30. Davis, S. and P.S. Meltzer, *GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor*. *Bioinformatics*, 2007. **23**(14): p. 1846-7.
31. Yu, G., et al., *clusterProfiler: an R package for comparing biological themes among gene clusters*. *Omics*, 2012. **16**(5): p. 284-7.

Figures

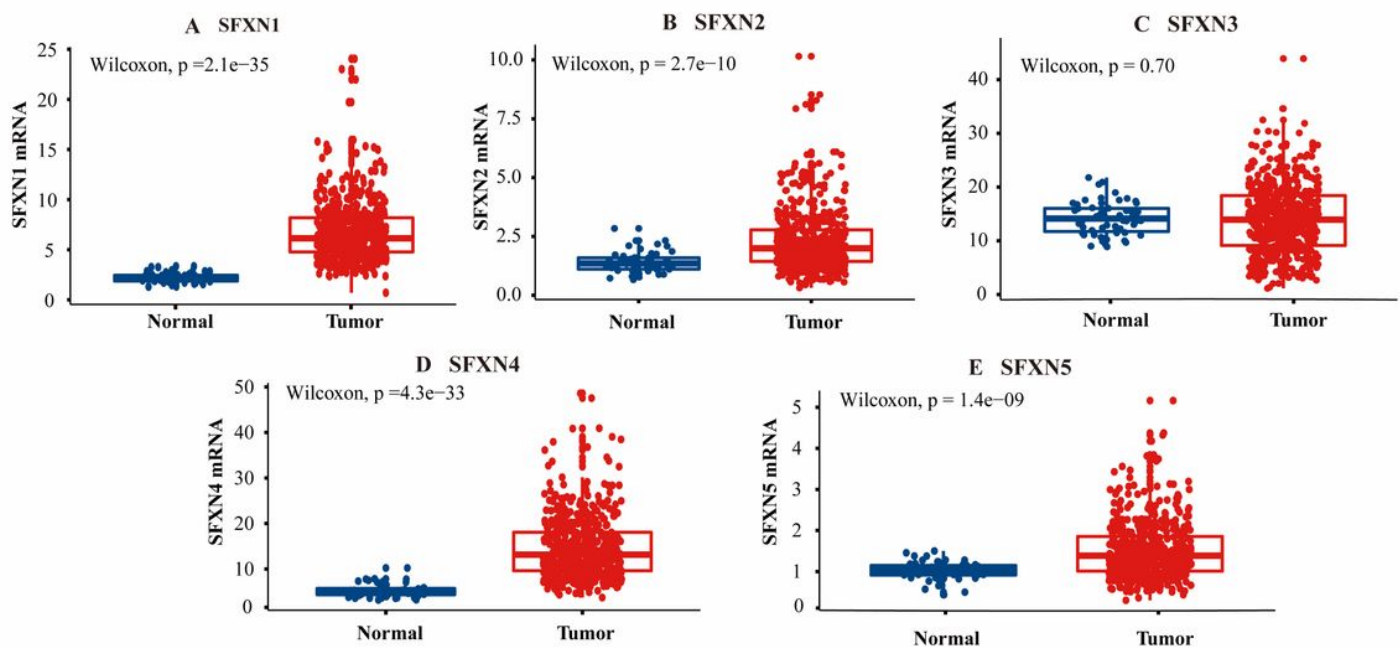


Figure 1

The different mRNA expression of SFXN complex between tumor and normal samples. Each plot is one sample of the LUAD and normal tissues from TCGA. SFXN family includes A SFXN 1, B SFXN 2, C SFXN3, D SFXN4, and E SFXN5.

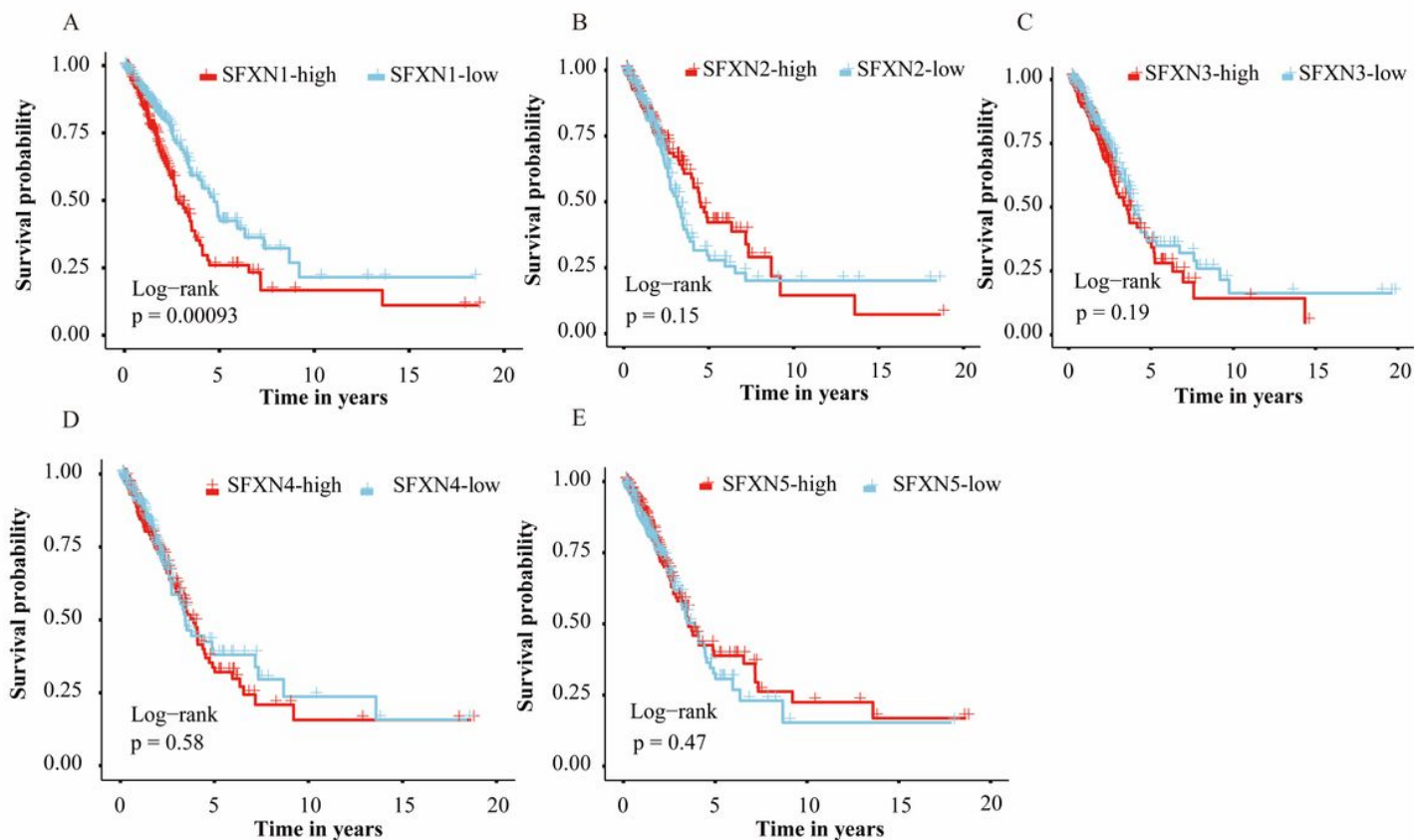


Figure 2

Association between SFXN family and OS in TCGA-LUAD cohort patients with Kaplan-Meier analysis. The samples were divided into high and low expression groups according to the median expression level of individual SFXNs: A SFXN1, B SFXN2, C SFXN3, D SFXN4, and E SFXN5.

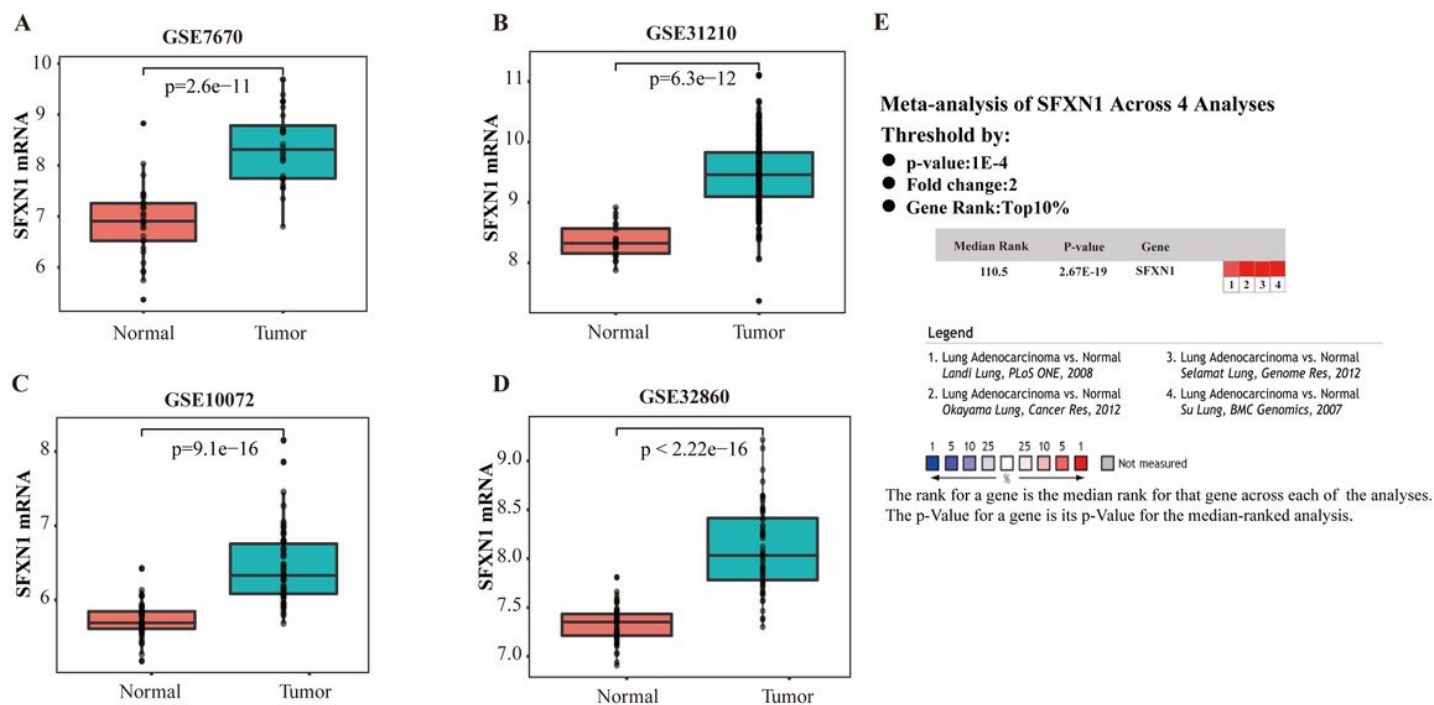


Figure 3

Verification of SFXN1 mRNA expression levels in LUAD from four GEO series and its Meta-analysis The four GEO series, including A GSE7670, B GSE31210, C GSE10072, and D GSE32860, all showed different expression between LUAD and normal tissue. A meta-analysis of SFXN1 expression in these 4 analyses in Oncomine database (E) was illustrated.

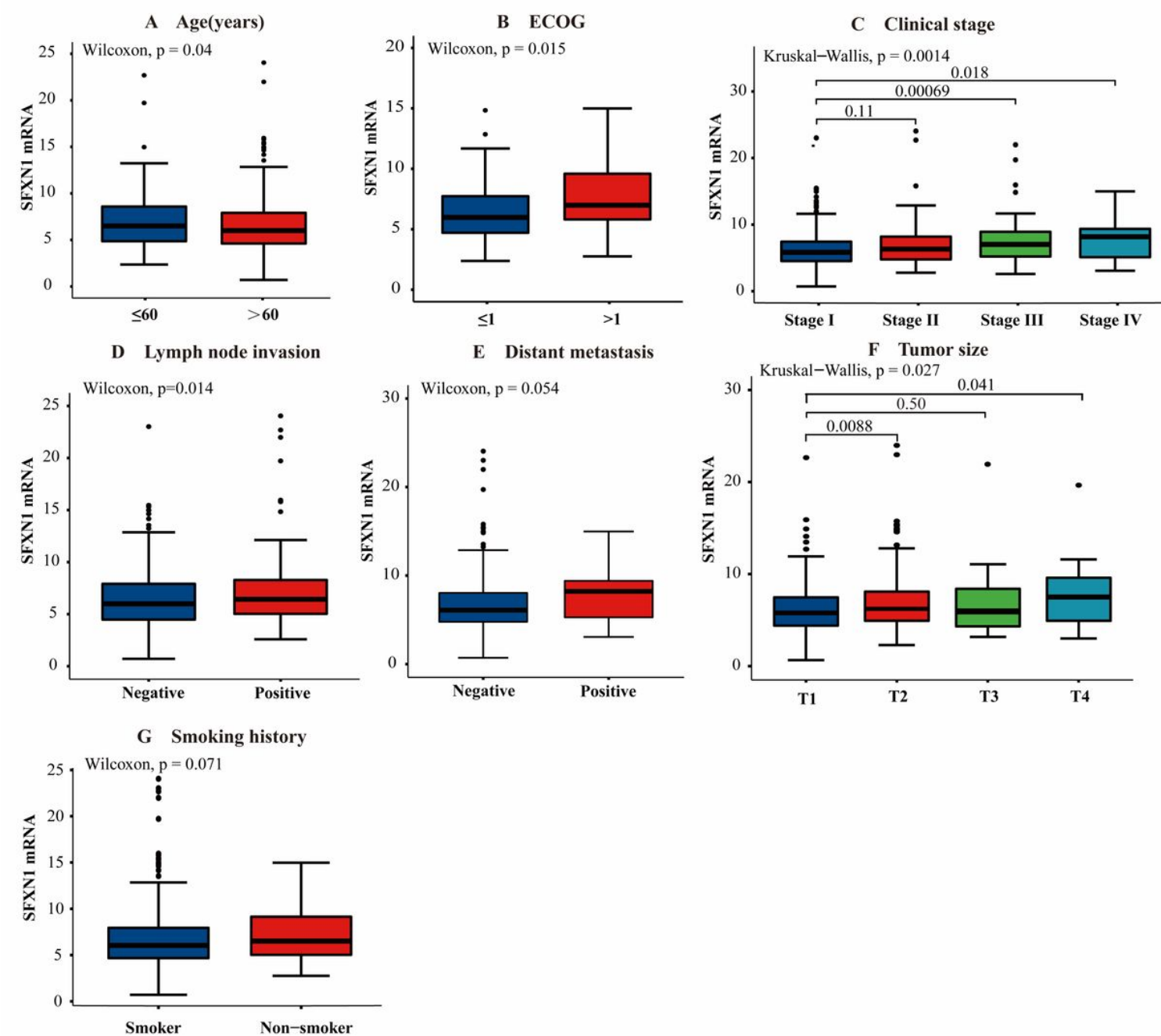


Figure 4

Association with SFXN1 expression and clinicopathologic characteristics The clinicopathologic variables are A Age, B ECOG, C Clinical stage, D Lymph node invasion, E Distant metastasis, F Tumor size and G

Smoking history. Among these, A, B, D, E, G was conducted with Wilcoxon whilst C and F via Kruskal-Wallis.

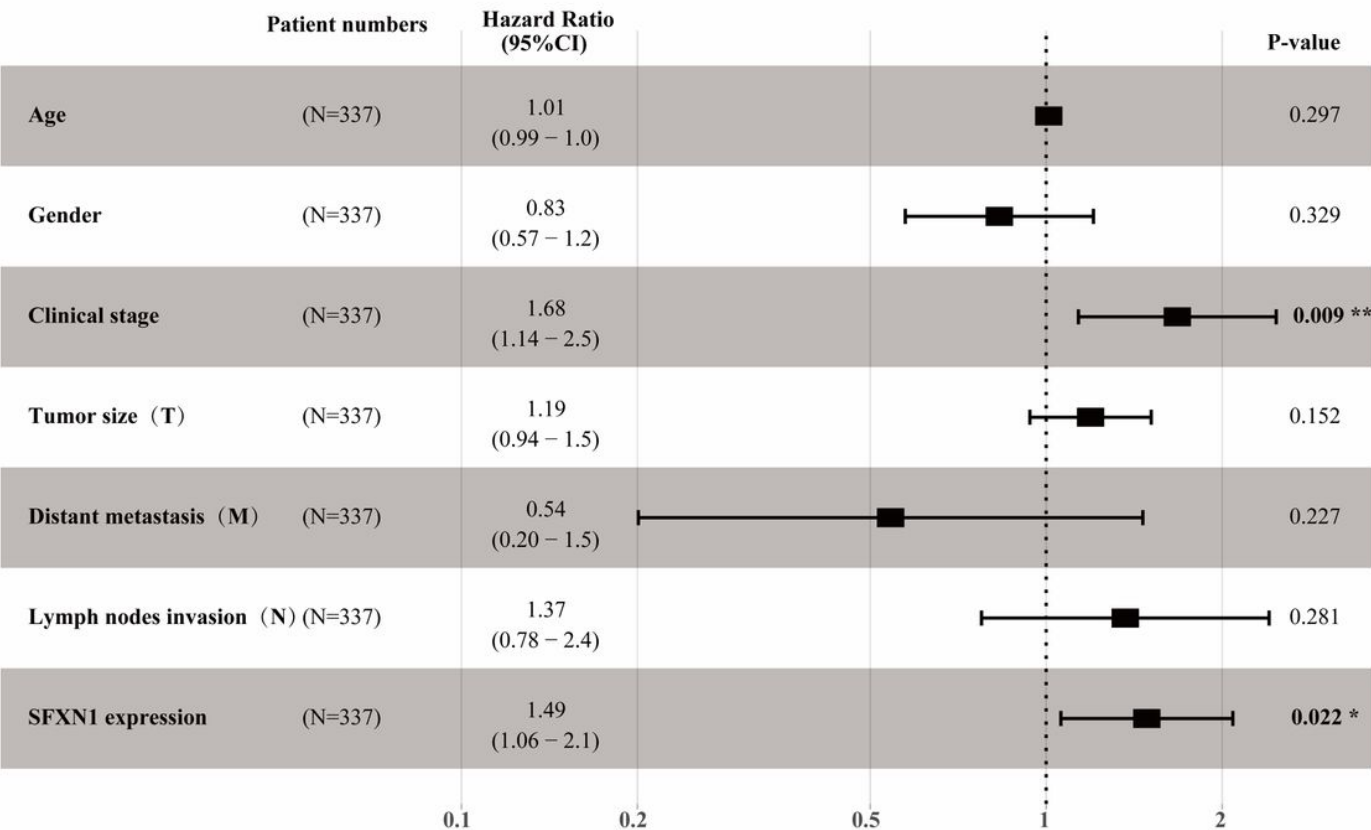


Figure 5

Forest plot of hazard ratios for the association between the clinicopathologic characteristics and OS The location of the black square is the indicator of the value of HR and the span of horizontal line is the range of 95% CI. The meaningful p value is bold. * ≤ 0.05 ; ** ≤ 0.01 .

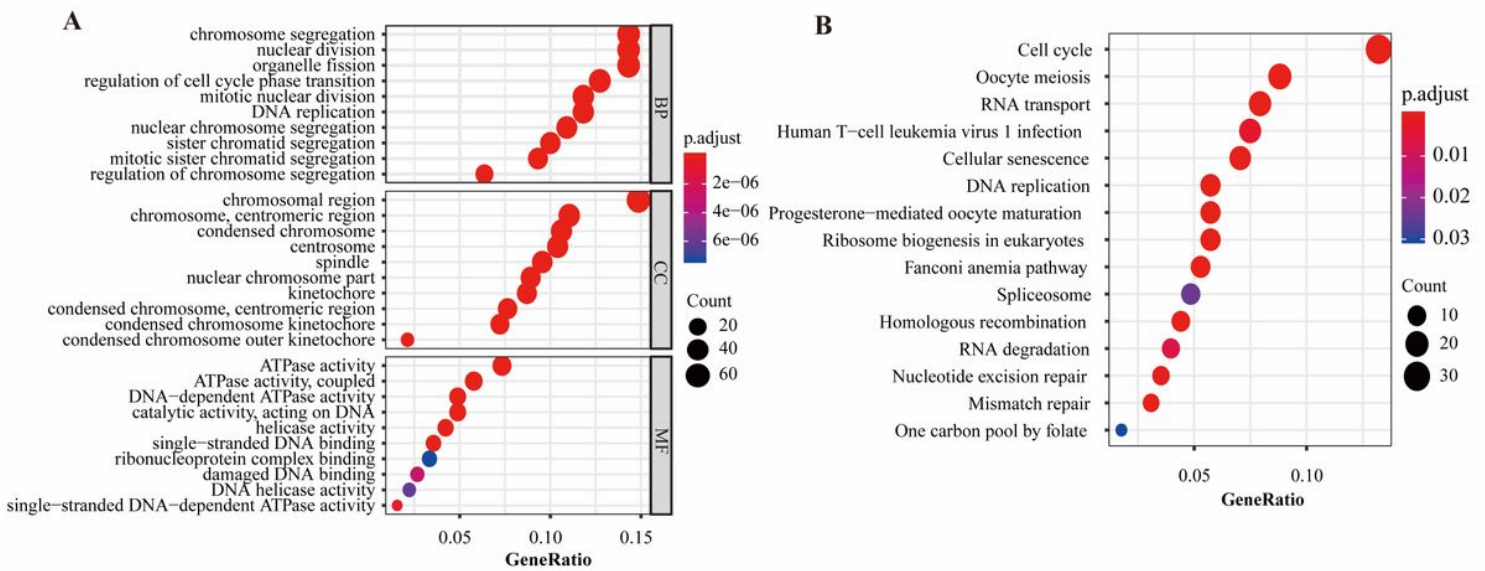


Figure 6

The GO biological process and the KEGG pathway enrichment of SFXN1 interactive genes analysis. The size of the bubbles represents the enrichment of gene numbers in some function or pathway: the bigger the circle, the larger the count of related genes; the color of bubbles stands for the intensity of correlation (p-value) of SFXN1 interactive genes and the corresponding term or pathway: the pure red means the p value is infinitesimal while the blue shows it is larger relatively but still under 0.05. A Three GO terms cover biological processes (BP), cellular components (CC) and molecular functions (MF). B KEGG pathways enrichment.