European Archives of Oto-Rhino-Laryngology (2019) 276:2565–2576
https://doi.org/10.1007/s00405-019-05525-2

HEAD AND NECK

Comprehensive analysis of gene expression and DNA methylation for human nasopharyngeal carcinoma

Hu Li[1] · Fu-Ling Wang[2] · Liang-peng Shan[3] · Jun An[1] · Ming-lei Liu[1] · Wei Li[1] · Jing-E. Zhang[1] · Ping-ping Wu[1]

Received: 7 March 2019 / Accepted: 18 June 2019 / Published online: 25 June 2019

Abstract

Purpose Nasopharyngeal carcinoma (NPC) is one of the most malignant head and neck carcinomas with unique epidemiological features. In this study, we aimed to identify the novel NPC-related genes and biological pathways, shedding light on the potential molecular mechanisms of NPC.

Methods Based on Gene Expression Omnibus (GEO) database, an integrated analysis of microarrays studies was performed to identify differentially expressed genes (DEGs) and differentially methylated genes (DMGs) in NPC compared to normal control. The genes which were both differentially expressed and differentially methylated were identified. Functional annotation and protein–protein interaction (PPI) network construction were used to uncover biological functions of DEGs.

Results Two DNA methylation and five gene expression datasets were incorporated. A total of 1074 genes were up-regulated and 939 genes were down-regulated in NPC were identified. A total of 719 differential methylation CpG sites (DMCs) including 1 hypermethylated sites and 718 hypomethylated sites were identified. Among which, 11 genes were both DEGs and DMGs in NPC. Pathways in cancer, p53 signaling pathway and Epstein–Barr virus infection were three pathways significantly enriched pathways in DEmRNAs of NPC. The PPI network of top 50 DEGs were consisted of 191 nodes and 191 edges.

Conclusions Our study was helpful to elucidate the underlying mechanism of NPC and provide clues for therapeutic methods.

Keywords Nasopharyngeal carcinoma · Integrated analysis · Differentially expressed genes · PPI network · Methylation

Introduction

Nasopharyngeal carcinoma (NPC) is a unique malignant head and neck carcinomas with marked racial and geographical differences. NPC is common in east Africa and Asia, especially in southern China, with an incidence rate of 0.2% [1, 2]. Base on epidemiological studies, the progression of NPC may be closely related to genetic factors, environmental factors, and Epstein–Barr Virus (EBV) infection [2, 3]. Advanced NPC with high mortality due to late stage diagnosis and metastasis is still the main cause of clinical failure [4]. Therefore, elucidating the mechanisms and screening biomarkers for early diagnosis have important clinical implications.

DNA methylation is one of the most widely studied epigenetic mechanisms regulating gene expression. It is well known that differences in DNA methylation and gene regulation are causal. Hypomethylation usually leads to gene expression, and hypermethylation leads to gene silencing. It is reported that the significance of DNA methylation in regulating development and progression of various cancers, including lung cancer [5], colorectal cancer [6], gastric cancer [7] and head and neck cancer [8]. Nevertheless, aberrant methylation in NPC have not been elucidated systemically.

Here, we identified the DEGs by integrated analysis of multiple gene-expression profiles in NPC compared to normal control. Function annotation and protein–protein interaction (PPI) network construction were constructed to understand the biological function of differentially expressed

✉ Jing-E. Zhang
jinge_zhang_doctor@126.com

✉ Ping-ping Wu
pingping_wu_doctor@126.com

[1] Department of Otolaryngology, The First People's Hospital of Jining, No. 6 Jiankang Road, Jining 272011, Shandong, China

[2] Department of Gynecology, Maternal and Child Health Hospital of RenCheng District, Jining, Jining, Shandong, China

[3] Department of Laboratory, The First People's Hospital of Jining, Jining, Shandong, China

Table 1 Gene expression datasets used in this study

GEO accession	Study type	Platform	Samples (N:P)	Year	Tissue
GSE64634	mRNA array	PL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	4:12	2017	Tissue
GSE40290	mRNA array	GPL8380 Capitalbio 22 K Human oligo array version 1.0	25:8	2016	Tissue
GSE53819	mRNA array	GPL6480 Agilent-014850 Whole Human Genome Microarray 4×44 K G4112F (Probe Name version)	18:18	2014	Tissue
GSE13597	mRNA array	GPL96 [HG-U133A] Affymetrix Human Genome U133A Array	3:25	2009	Tissue
GSE12452	mRNA array	GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	10:31	2008	Tissue
GSE52068	Methylation array	GPL13534 Illumina HumanMethylation450 BeadChip (HumanMethylation450_15017482)	24:24	2016	Tissue
GSE62336	Methylation array	GPL13534 Illumina HumanMethylation450 BeadChip (HumanMethylation450_15017482)	25:25	2016	Tissue

genes (DEGs) and identify the pathways enriched in NPC. We also obtained the differentially methylated genes (DMGs) in NPC according to gene expression omnibus (GEO) database and we performed a comprehensive analysis using a bioinformatics approach to discover the regulatory roles of DEGs and DMGs in the development of NPC. Hopefully, the DEGs, DMGs and pathways found in this study may advance the knowledge about the cellular and molecular events occurring in NPC, and raise new strategies of treatment for NPC.

Materials and methods

Gene expression profile

We searched datasets from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) with the keywords "Nasopharyngeal carcinoma, NPC"[MeSH Terms] OR "Nasopharyngeal carcinoma, NPC" [All Fields]) AND "Homo sapiens"[porgn] AND "gse"[Filter]. The inclusion criteria for the present study were: (1) the type of dataset was described as "expression profiling by array and methylation profiling by array. (2) Dataset should be whole-genome mRNA expression profile by array. (3) Datasets were obtained by tissue samples of NPC and normal control group. (4) The datasets should be normalized or original, and five sets of mRNA data and two sets methylation data of NPC were selected.

Identification of DEGs between NPC and normal controls

GSE64634, GSE40290, GSE53819, GSE13597 and GSE12452 contained both NPC samples and normal control samples. We screened the DEGs between NPC and normal samples across the five data sets using MetaMA package. MetaMA was used to combine data from multiple microarray datasets, and we obtained the individual p values. The Benjamini and Hochberg method was introduced to adjust the p values into a false discovery rate (FDR) to avoid the multi-test problem. The FDR < 0.05 was set as the thresholds for identifying DEGs. Heat map of top 100 DEGs was obtained by pheatmap package.

Protein–protein interaction (PPI) network construction

The top 50 DEGs of NPC were used to construct a PPI network using the Biological General Repository for Interaction Datasets (BioGRID) (https://thebiogrid.org/) and Cytoscape (https://www.cytoscape.org/). The nodes and edges represent the proteins and interactions between two proteins, respectively.

Identification of differentially methylated CpG sites

We deleted the CpG sites for which beta value was not available in more than 80% samples, and obtained 476,993 CpG sites. COHCAP package in R was used to evaluate the differentially methylated CpG sites (DMCs) between NPC and normal tissues. The $|\Delta\beta| > 0.2$, FDR < 0.05 was set as the thresholds for identifying DMCs. Manhattan plot was built to study the distribution of CpG sites according to FDRby qqman package in R. Heat map of the DMCs was obtained by pheatmap package.

Functional annotation of DEGs and DMGs

Using CPDB (https://cpdb.molgen.mpg.de/CPDB), Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment

Table 2 The top 40 DEGs in NPC

ID	Symbol	FDR	Up/down
4585	MUC4	0	Down
94025	MUC16	0	Down
50853	VILL	1.10E–12	Down
56673	C11orf16	8.67E–12	Down
9071	CLDN10	1.02E–11	Down
9940	DLEC1	1.76E–11	Down
8382	NME5	2.15E–11	Down
9576	SPAG6	9.29E–11	Down
8938	BAIAP3	1.51E–10	Down
79838	TMC5	1.18E–09	Down
27285	TEKT2	1.57E–09	Down
26150	RIBC2	1.87E–09	Down
1770	DNAH9	9.83E–09	Down
3753	KCNE1	1.67E–08	Down
6236	RRAD	1.67E–08	Down
51364	ZMYND10	2.06E–08	Down
64446	DNAI2	2.06E–08	Down
27019	DNAI1	2.24E–08	Down
4246	SCGB2A1	2.94E–08	Down
5858	PZP	3.17E–08	Down
1434	CSE1L	1.87E–09	Up
51053	GMNN	2.47E–09	Up
1728	NQO1	3.46E–09	Up
29980	DONSON	7.34E–09	Up
4925	NUCB2	9.83E–09	Up
55388	MCM10	9.83E–09	Up
9833	MELK	1.67E–08	Up
51203	NUSAP1	1.67E–08	Up
11169	WDHD1	1.67E–08	Up
6624	FSCN1	1.84E–08	Up
11130	ZWINT	2.12E–08	Up
3223	HOXC6	2.59E–08	Up
84823	LMNB2	3.31E–08	Up
9493	KIF23	3.42E–08	Up
11098	PRSS23	4.61E–08	Up
8537	BCAS1	5.46E–08	Up
	KIAA0101	5.46E–08	Up
11343	MGLL	6.95E–08	Up
83990	BRIP1	6.95E–08	Up
9805	SCRN1	8.59E–08	Up

analysis were performed to uncover biological functions of DEGs. FDR < 0.01 was considered as statistically significant.

Correlation analysis between DMGs and DEGs

The genes which were not only differentially expressed, but also differentially methylated were identified to evaluate the association between gene expression and DNA methylation. Since DNA methylation makes negative influences on gene expression, the DEGs whose expression and DNA methylation plays opposite regulation in NPC were identified.

Results

DEGs in NPC

Five datasets (GSE64634, GSE40290, GSE53819, GSE13597 and GSE12452) were downloaded from GEO (Table 1). Compared with the normal controls, 2013 DEGs in NPC were obtained with FDR < 0.05, among which, 1074 genes were up-regulated and 939 genes were down-regulated. Top 40 DEGs between NPC and normal controls are demonstrated in Table 2. Heat map of top 100 DEGs is shown in Fig. 1.

Functional annotation of DEGs.

According to the GO enrichment analysis with FDR < 0.01, protein binding (FDR = 1.69E–54), response to organic substance (FDR = 2.31E–40), cellular response to chemical stimulus (FDR = 1.91E–39) and cytoplasm (FDR = 3.61E–52) were most significantly enriched GO terms. The top 15 GO terms of DEGs in NPC are demonstrated in Fig. 2a–c. After the KEGG pathway enrichment analysis (FDR < 0.01), we found that Cell cycle (FDR = 1.93E–12), Pathways in cancer (FDR = 3.39E–11), p53 signaling pathway (FDR = 9.58E-–06) and Epstein–Barr virus infection (FDR = 9.88E–06) were significantly enriched pathways in NPC. Top 15 most significantly enriched KEGG pathways of DEGs in NPC are displayed in Fig. 2d.

PPI network construction

The PPI network of top 50 DEGs in NPC was consisted of 191 nodes and 191 edges (Fig. 3). Among them, the top 15 DEG with high degree are CSE1L (degree = 22), ZWINT (degree = 16), GMNN (degree = 15), NUSAP1 (degree = 11), MCM10 (degree = 11), KIF23 (degree = 10), FSCN1 (degree = 9), NUCB2 (degree = 8), LMNB2 (degree = 8), DNAI2 (degree = 7), PRSS23 (degree = 7), ATP12A (degree = 6), ABAT (degree = 6), NQO1 (degree = 5) and MELK (degree = 5).

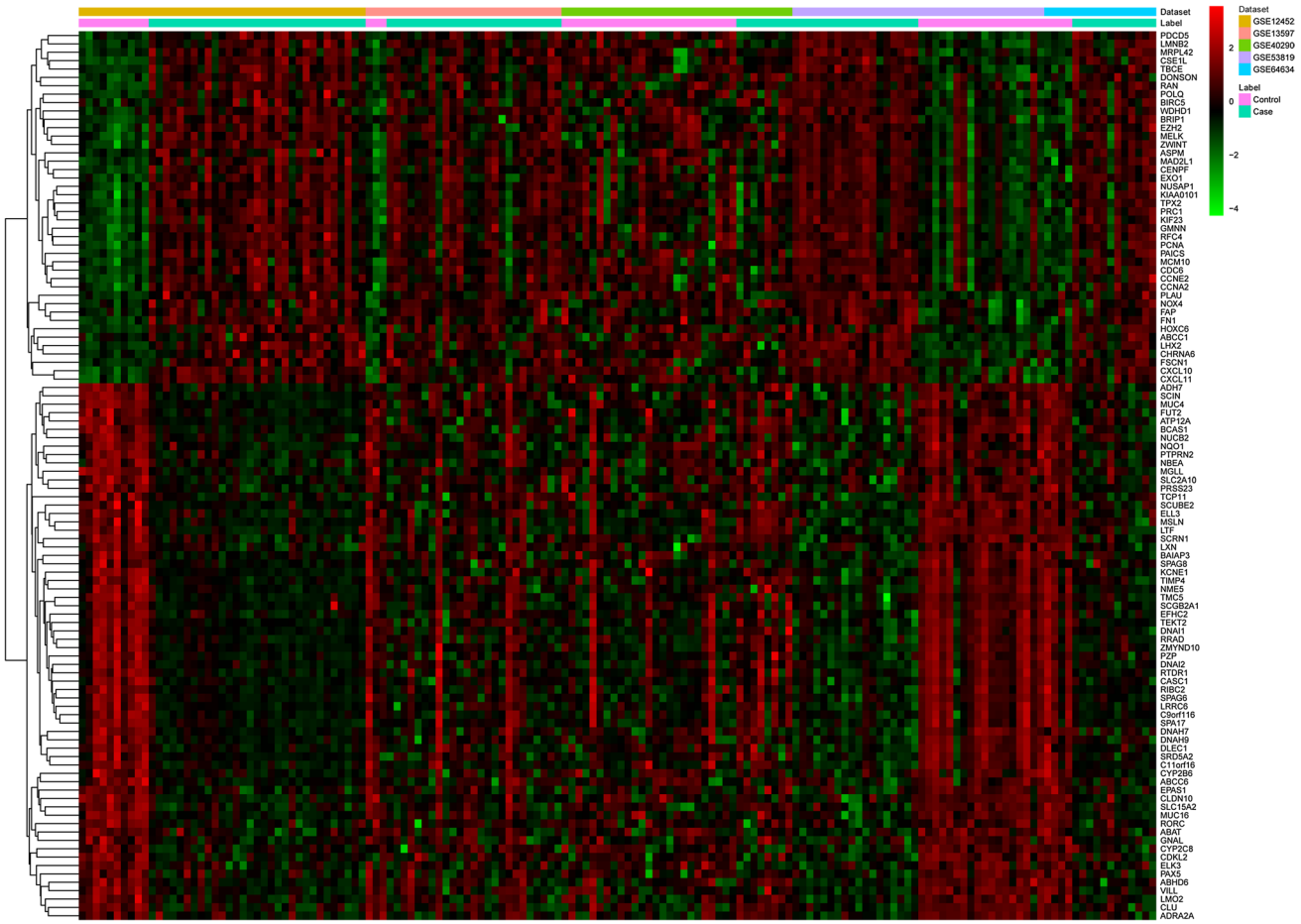

Fig. 1 Heat map of top 100 DEGs. Row and column represented DEGs and tissue samples, respectively. The color scale represented the expression levels

DMCs in NPC

A total of 719 DMCs including 1 hypermethylated sites and 718 hypomethylated sites were identified with $|\Delta\beta| > 0.2$ and FDR < 0.05. As is displayed in Manhattan plot (Fig. 4a), those CpG sites distributed in all chromosomes. Heat map was performed using the 379 DMCs, which showed that the NPC was in methylated state compared to the normal group (Fig. 4b). The top 20 differential methylation CpG sites are listed in Table 3. The results of correlation analysis between DMGs and DEGs is displayed in Table 4. Among which, 11 DEGs were both up-regulated and hypomethylated in NPC compared to normal control.

Functional annotation of DMGs

According to the GO enrichment analysis with FDR < 0.01, sequence-specific DNA binding 45 (FDR = 8.54E–07), DNA-binding transcription factor activity, RNA polymerase II-specific (FDR = 3.47E–06), cellular developmental process (FDR = 2.39E–10) and cell–cell signaling (FDR = 9.93E–09) were most significantly enriched GO terms. The top 15 GO terms of DEGs in NPC were displayed in Fig. 5a–c. After the KEGG pathway enrichment analysis (FDR < 0.01), we found that csignaling pathway (FDR = 0.003102242), Rap1 signaling pathway (FDR = 0.007804562) and Phosphatidylinositol signaling system (FDR = 0.008031404) were significantly enriched pathways in NPC. Top 15 most significantly enriched KEGG pathways of DEGs in NPC were demonstrated in Fig. 5d. Five datasets including GSE64634, GSE40290, GSE53819, GSE13597 and GSE12452 were selected in this study. Only two datasets (GSE53819 and GSE13597) have clinical information. In the GSE53819 dataset, there is a gender classification for each sample, and plots of the five genes of DLEC1, CSE1L, NQO1, FSCN1, and CDKN1A in different genders (Fig. 6). The GSE13597 dataset contains clinical staging information,

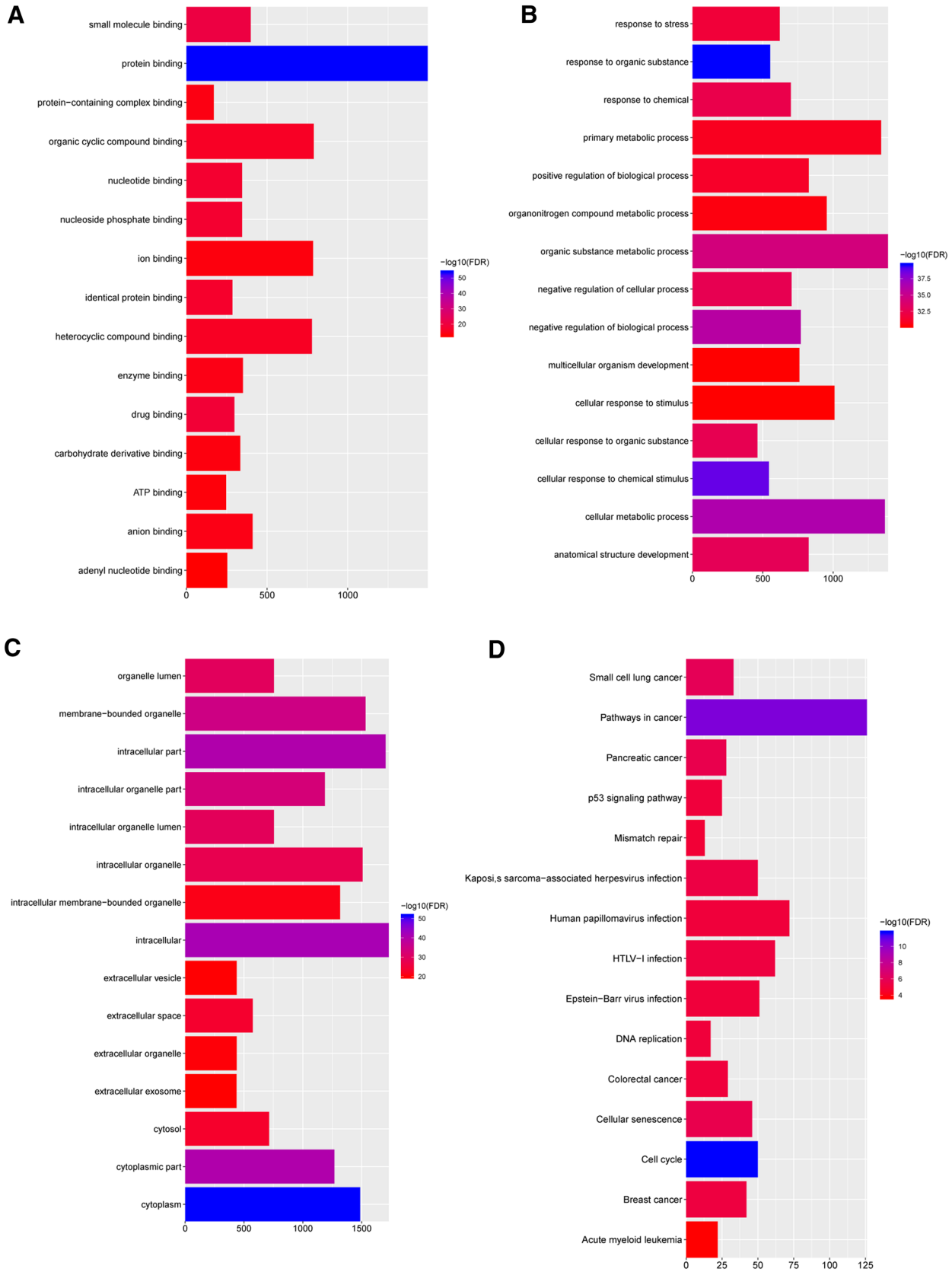

Fig. 2 The top 15 significant enrichment GO terms and KEGG pathways of DEGs in NPC. The *x* axis shows -log FDR and *y* axis shows GO terms or KEGG pathways. **a** Molecular function. **b** Biological process. **c** Cellular component. **d** KEGG pathways

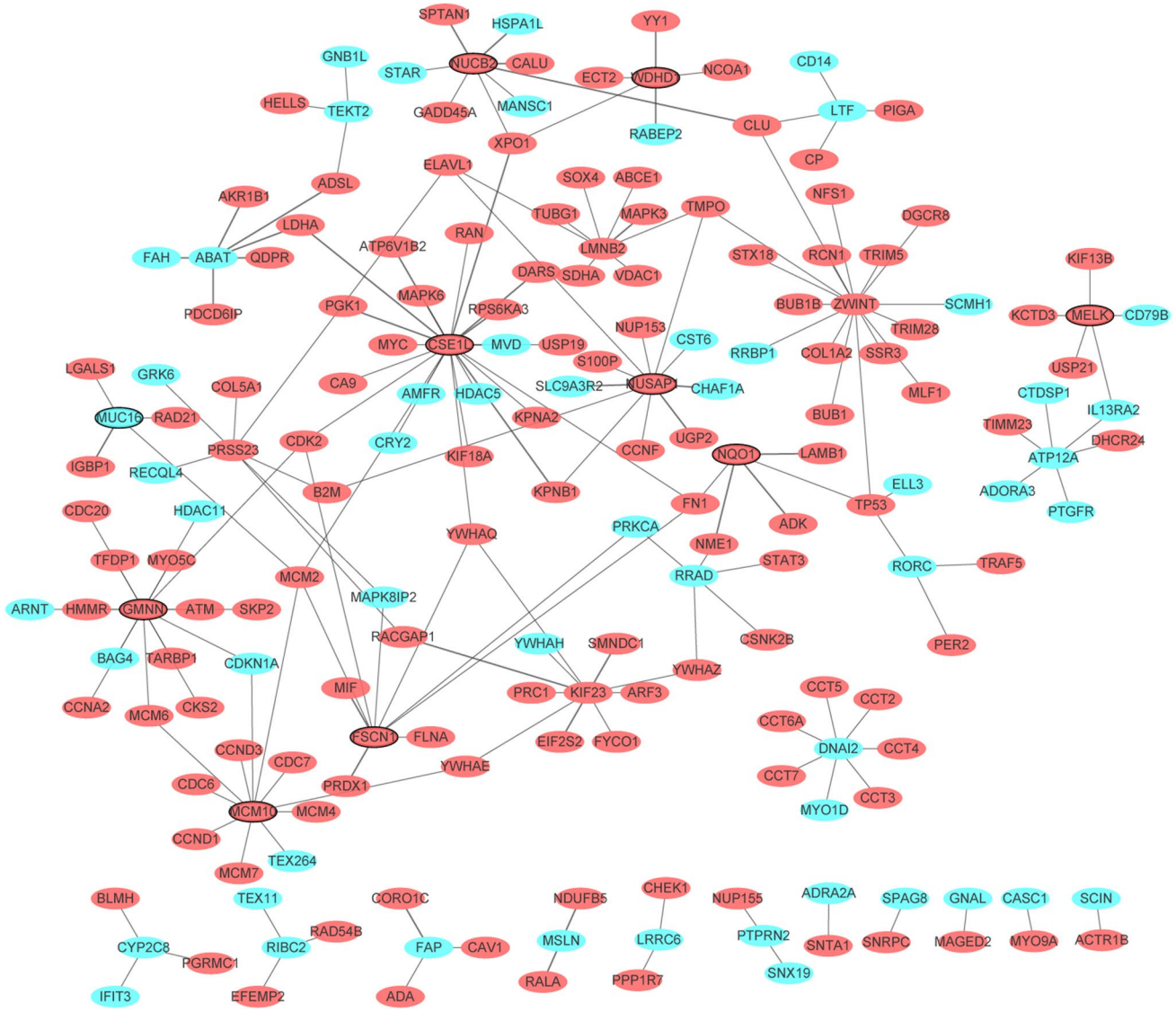

Fig. 3 The PPI network. Red and green represented up- and down-ward adjustments, respectively. The black border indicates top 10 Up/Down

and plots the five genes of DLEC1, CSE1L, NQO1, FSCN1, and CDKN1A in different stages (Fig. 6).

Discussion

The incidence of NPC differs from other malignancies by race and geography. NPC is a common malignancy in east Africa and Asia, especially in southern China. Although NPC's radiotherapy and radiotherapy and chemotherapy have greatly improved, the results of patients with local advanced NPC are still unsatisfied [9]. Based on the integrated analysis of bioinformatics analysis, our study identified the DEGs and DMCs in NPC. The genes which were both differentially expressed and methylated in NPC were identified as well. After functional annotation and PPI network, we identified several genes associated with NPC (see Fig. 7).

It is well known that differences in DNA methylation and gene regulation are causal. Hypomethylation usually leads to gene expression, and hypermethylation leads to gene silencing. Tian et al. found that DNA methylation of DLEC1 was found to be significantly higher in 25% of NPC patients than that in normal controls [10]. In the current study, DLEC1was down-regulated in NPC. Our integrated analysis results may provide evidence for the

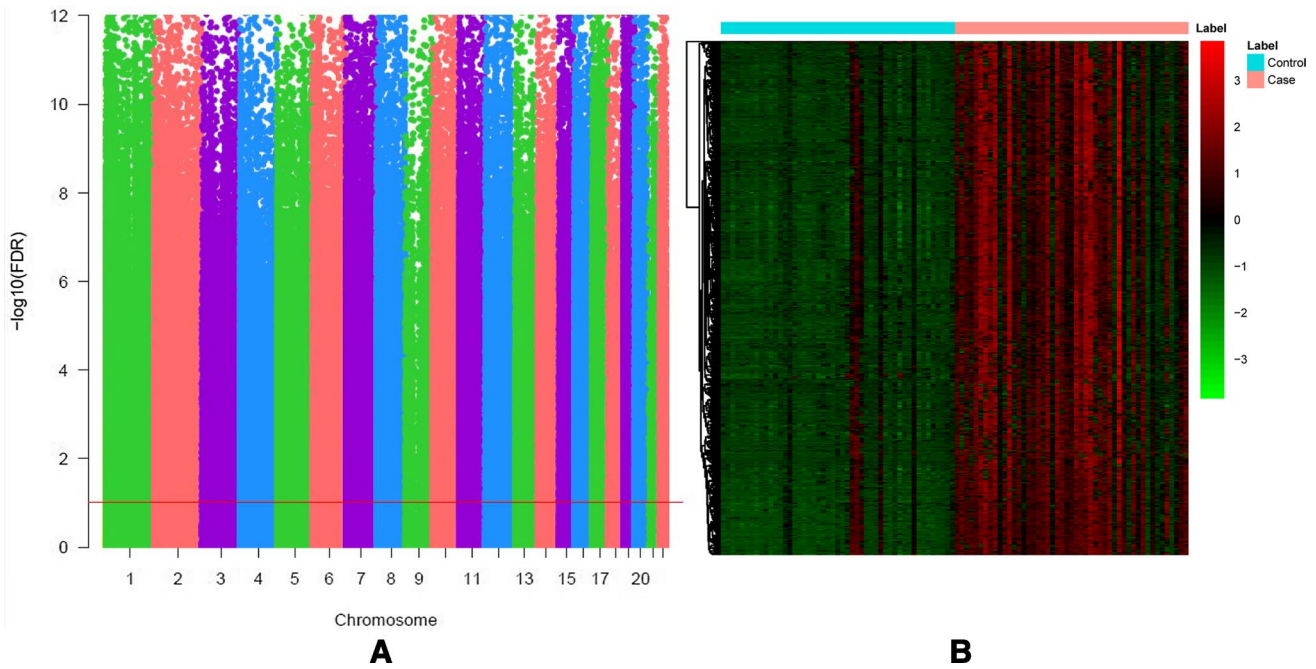

Fig. 4 **a** Manhattan plot of NPC. **b** Heat map of 379 DMCs. Row and column represented DMCs and tissue samples, respectively. The color scale represented the expression levels

hypermethylation of the promoter of DLEC1 as a means of detecting NPC.

Chromosome segregation 1-like (CSE1L), is considered to be an oncogene candidate and has shown high frequencies of gain in NPC, glioblastoma multiforme and prostate cancer. Fang et al. reported that CSE1L was overexpressed in NPC [11], and suggested that high expression of CSE1L may occur mainly in highly malignant NPC cells and maintain

Table 3 The top 20 of the differentially methylated CpG sites

SiteID	Gene	Island	$\Delta\beta$	*p* value	FDR
cg10210594		chr1:208132327–208133117	0.314477	4.12E–22	4.38E–17
cg04482794	ITPKB	chr1:226924560–226926553	0.270915	2.48E–22	4.38E–17
cg00803088	RET	chr10:43600146–43601165	0.356699	4.47E–22	4.38E–17
cg07990843		chr17:25676304–25676603	0.408869	4.59E–22	4.38E–17
cg02018277	PIF1	chr15:65116013–65116567	0.324356	1.39E–21	1E–16
cg20387815	POMC	chr2:25391072–25391875	0.222942	1.47E–21	1E–16
cg11299854	CCNI2	chr5:132082873–132083911	0.363718	2.33E–21	1.24E–16
cg00800353		chr12:53273231–53273498	0.293051	3.29E–21	1.43E–16
cg16195091	LOC283999	chr17:76228110–76228380	0.228186	4.19E–21	1.54E–16
cg15579587	LRRFIP1	chr2:238599857–238601430	0.33526	4.08E–21	1.54E–16
cg02616947	KAZALD1	chr10:102820488–102822874	0.282918	4.95E–21	1.69E–16
cg15603424	ARNTL	chr11:13298796–13300735	0.286282	1.59E–20	4.47E–16
cg04878000		chr1:32237827–32238661	0.24621	1.98E–20	5.26E–16
cg08545268	TGM2	chr20:36793549–36793867	0.269524	2.34E–20	5.87E–16
cg21101720	ANKRD13B	chr17:27939298–27940770	0.398497	2.68E–20	6.38E–16
cg27436264	TGM2	chr20:36793549–36793867	0.276017	2.88E–20	6.53E–16
cg07143451		chr17:25676304–25676603	0.205586	3.69E–20	7.65E–16
cg05175020	TSC22D4	chr7:100075303–100075551	0.287379	4.02E–20	8E–16
cg27058257	VSTM2B	chr19:30015781–30021367	0.325064	6.69E–20	1.23E–15
cg04158367	GFI1	chr1:92945907–92952609	0.221172	8.15E–20	1.31E–15

Table 4 The genes both differentially expressed and differentially methylated in NPC

Gene	Island	$\Delta\beta$	FDR	Hyper/hypo methylation
LRRFIP1	chr2:238599857–238601430	0.335259606	1.53859E–16	Hypo
BCAT1	chr12:25055599–25056246	0.288916843	4.22752E–15	Hypo
RAPGEFL1	chr17:38347533–38347765	0.324117562	1.00398E–14	Hypo
RAPGEFL1	chr17:38347533–38347765	0.246660522	3.00228E–14	Hypo
PLCL2	chr3:16925381–16926849	0.257923134	3.35129E–14	Hypo
CD1D	chr1:158150620–158151503	0.246680901	5.81639E–14	Hypo
PPFIA4	chr1:203044722–203045390	0.2043036	1.36398E–13	Hypo
PTPRU	chr1:29585897–29586598	0.268008457	1.55869E–13	Hypo
FKBP4	chr12:2903389–2905189	0.220853274	1.97969E–13	Hypo
ICAM5	chr19:10405924–10406398	0.228596527	2.16947E–13	Hypo
PPFIA4	chr1:203044722–203045390	0.208203082	2.84344E–13	Hypo
CCNA1	chr13:37005581–37006453	0.246992902	3.1168E–13	Hypo
RAPGEFL1	chr17:38347533–38347765	0.338674556	3.74192E–13	Hypo
ALDH1A3	chr15:101419261–101421133	0.217854307	5.23247E–13	Hypo
EPHA4	chr2:222436034–222438941	0.24356032	5.34376E–13	Hypo
TACC2	chr10:123922850–123923542	0.255694342	7.42524E–13	Hypo
PTPRU	chr1:29585897–29586598	0.269342619	1.41123E-12	Hypo
RAPGEFL1	chr17:38347533–38347765	0.267254183	1.87706E–12	Hypo
TACC2	chr10:123922850–123923542	0.254598551	3.66038E–12	Hypo
MEIS1		0.211586121	4.79862E–12	Hypo
PPP3CC	chr8:22298112–22299142	0.226109517	4.79917E–12	Hypo
TRIM36	chr5:114514716–114516220	0.213817107	4.82898E–12	Hypo
RHCG	chr15:90039464–90039984	0.212918837	7.46022E–12	Hypo
RPS6KA2	chr6:167275720–167276626	0.251640516	8.26616E–12	Hypo
MEIS2	chr15:37390175–37390380	0.238734976	9.14978E–12	Hypo
MEIS1	chr2:66672431–66673636	0.230772767	9.50236E–12	Hypo
FAIM2	chr12:50297580–50297988	0.223173565	1.12521E–11	Hypo
ALDH1A3	chr15:101419261–101421133	0.245845005	1.17646E–11	Hypo
MEIS1		0.209810649	1.87679E–11	Hypo
F2R	chr5:76011120–76012292	0.243913913	1.97985E–11	Hypo
ME3	chr11:86382695–86383586	0.2194241	2.79883E–11	Hypo
GREM1	chr15:33009530–33011696	0.212291736	3.28016E–11	Hypo
LHX2	chr9:126785021–126785225	0.241631276	3.4579E–11	Hypo
ALDH1A3	chr15:101419261–101421133	0.226826129	3.54097E–11	Hypo
SLCO5A1	chr8:70744098–70747441	0.232060483	8.85692E–11	Hypo
FAIM2	chr12:50297580–50297988	0.226727858	1.49908E–10	Hypo
MEIS2	chr15:37387386–37387614	0.226186023	1.60738E–10	Hypo
CR2	chr1:207627531–207628280	0.200462584	1.62821E–10	Hypo
CCND2	chr12:4383193–4384405	0.24369245	1.68615E–10	Hypo
NR2E1	chr6:108485671–108490539	0.24786707	2.44664E–10	Hypo
H2AFY	chr5:134734096–134735261	0.211897669	2.80863E–10	Hypo
TRIM36	chr5:114514716–114516220	0.239740658	6.13188E–10	Hypo
PDGFB	chr22:39638219–39640966	0.255870451	6.33758E–10	Hypo
ALDH1A3	chr15:101419261–101421133	0.220314628	7.1204E–10	Hypo
ITGA5	chr12:54811981–54812202	0.255721499	1.04256E–09	Hypo
CH25H	chr10:90967555–90967900	0.205594501	1.39809E–09	Hypo
MEIS2	chr15:37387386–37387614	0.215936979	1.52667E–09	Hypo
RTDR1	chr22:23483997–23484309	0.224116387	1.98848E–09	Hypo
MEIS2	chr15:37390175–37390380	0.206145901	4.27244E–09	Hypo
SIX2	chr2:45235511–45237792	0.206948946	4.6741E–09	Hypo

Table 4 (continued)

Gene	Island	Δβ	FDR	Hyper/hypo methylation
MYO15B	chr17:73583838–73586337	0.257722387	4.73617E–09	Hypo
SULF1		0.211337721	6.01072E–09	Hypo
AGRN	chr1:967966–970238	0.208576801	4.45476E–08	Hypo
SIM1	chr6:100912071–100913337	0.20366358	4.90972E–08	Hypo
AGRN	chr1:975496–980305	0.200635082	1.79984E–07	Hypo
MYO15B	chr17:73583838–73586337	0.215727584	5.55534E–07	Hypo

The bold represents the up-regulated differentially expressed gene

its strongly malignant phenotype of cells. Herein, CSE1L, one of top 10 DEGs, was up-regulated in NPC. The results of our integrated analysis further confirmed that CSE1L is an oncogene in NPC.

NAD(P)H quinone dehydrogenase 1 (NQO1), is an enzyme that has been displayed to detoxify many natural and synthetic compounds. Instead, it activates certain anticancer agents [12]. The lower enzyme activity of NQO1 has been shown to be associated with increased risk of myeloid leukemia and bladder carcinoma [12, 13]. It is reported that NQO1 genes does not contribute to overall NPC risk in a Han Chinese in southern China [14]. In this study, NQO1 was up-regulated in NPC. These findings suggested that further research is needed to determine the role of NQO1 in NPC.

FSCN1 is up-regulated and involved in cell movement and tumor invasiveness in breast and colon cancer [15, 16]. Wu et al. found that FSCN1 was overexpression in NPC and its expression predicts poor prognosis in patients with NPC and correlates with tumor invasion [17]. It is reported that inhibiting FSCN1 expression significantly inhibited NPC cell proliferation and invasion [18]. In this study, we found that FSCN1 was up-regulated in NPC. The above results suggested that FSCN1 is a key gene in regulating the occurrence and progression of NPC.

CDKN1A, a cyclin-dependent kinase inhibitor, regulates the cell cycle progression at G1 [19]. Its expression is tightly controlled by the tumor suppressor protein p53, through which CDKN1A mediates the p53-dependent cell cycle G1 phase arrest in response to many of stress stimuli [20]. In this study, our results showed that CDKN1A was down-regulated in NPC. Base on the results of functional annotation, CDKN1A was significantly enriched pathway of Epstein–Barr virus infection. The Epstein–Barr virus infection pathway is associated with a wide variety of human tumor. The closest association with Epstein–Barr virus infection is seen in NPC [21]. Therefore, we presumed that CDKN1A might involve development of NPC by regulating Epstein–Barr virus infection pathway.

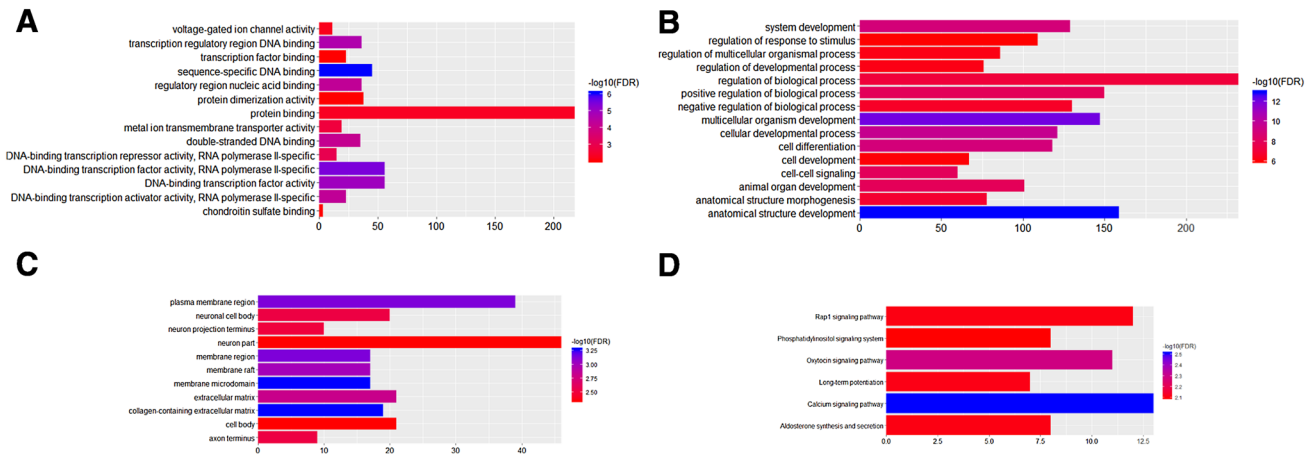

Fig. 5 The top 15 significant enrichment GO terms and KEGG pathways of DMGs in NPC. The *x* axis shows -log FDR and *y* axis shows GO terms or KEGG pathways. **a** Molecular function. **b** Biological process. **c** Cellular component. **d** KEGG pathways

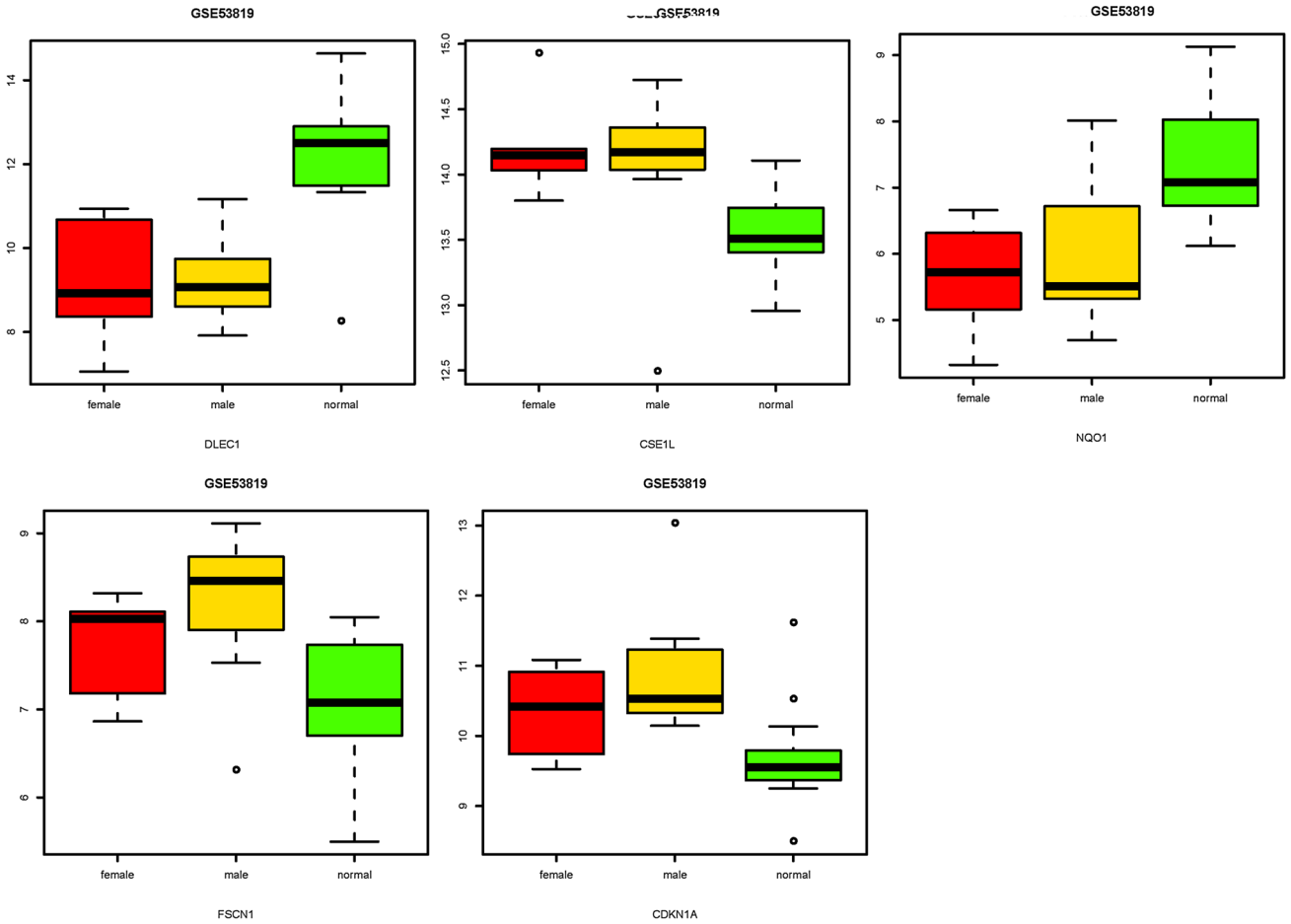

Fig. 6 Box-plot displayed the expression levels of DLEC1, CSE1L, NQO1, FSCN1, and CDKN1A in NPC of different genders

DNA methylation is a key event of epigenetic modification with heritable and stable. Promoter silencing or down-regulation of hypermethylation-inducible genes, particularly for tumor suppressor genes, while global DNA hypomethylation leads to genomic instability [22]. Emerging evidences have demonstrated that aberrant DNA methylation profiles were involved the occurrence and development of NPC [23, 24]. Herein, our integrated analysis results showed that a total of 719 DMCs including one hypermethylated sites and 719 hypomethylated sites were identified in NPC. According to the results of correlation analysis between DMGs and DEGs, 11 DEGs were both up-regulated and hypomethylated in NPC compared to normal control. Our further results may provide evidence suggested that the aberrant DNA methylation profiles were involved the occurrence and development of NPC.

In this study, we identified the DEGs associated with NPC by integrated analysis of many gene-expression profiles in NPC compared to normal control. Function annotation and PPI network construction were performed to interpret the biological function of DEGs and identify the pathways enriched in NPC. We also obtained the DMGs in NPC according to GEO. Our study highlights the importance of DEGs and DMGs in NPC, and may advance the knowledge about the molecular events occurring in NPC.

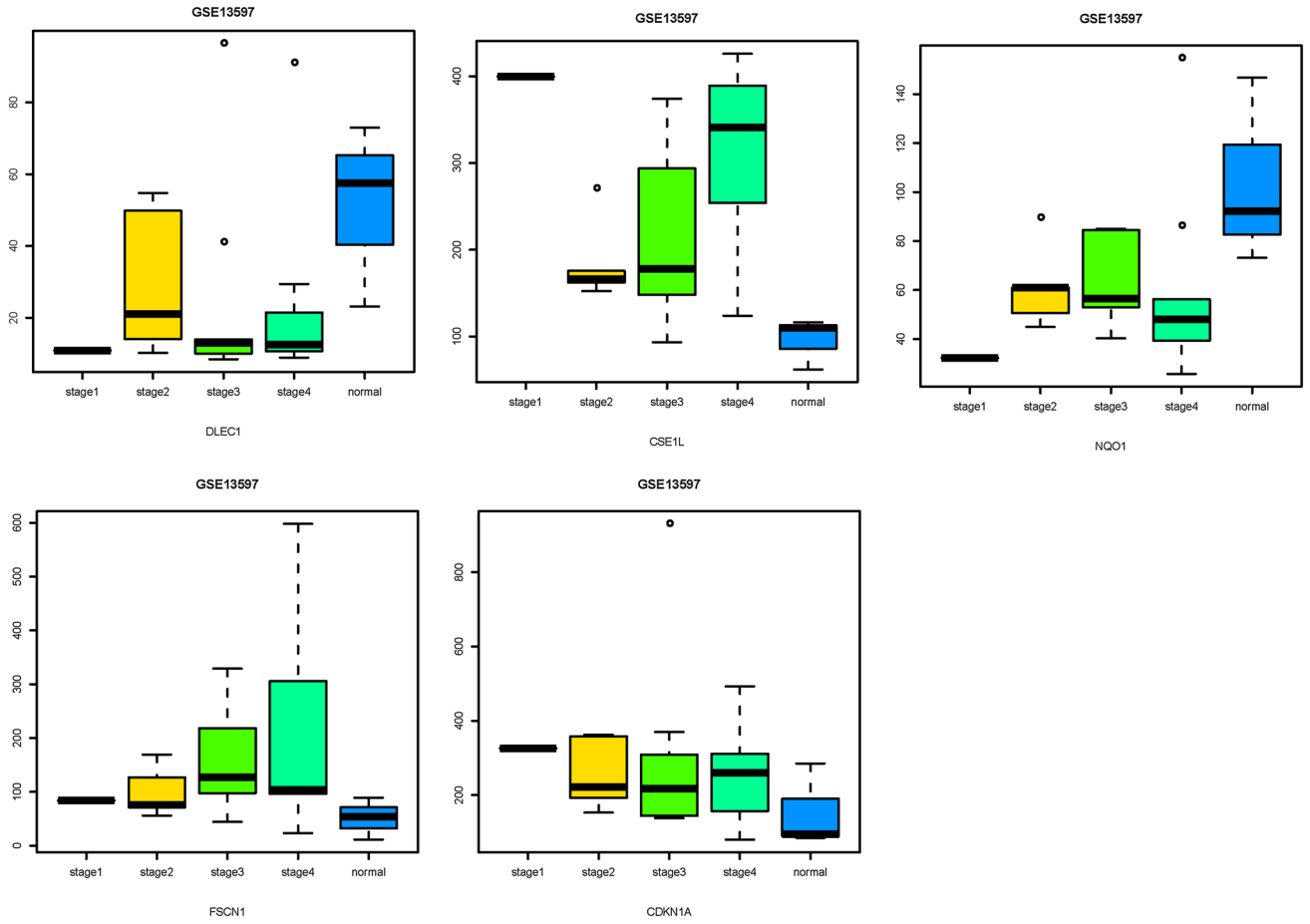

Fig. 7 Box-plot displayed the expression levels of DLEC1, CSE1L, NQO1, FSCN1, and CDKN1A in NPC of different stages

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest.

References

1. Tham IW, Lu JJ (2010) Controversies and challenges in the current management of nasopharyngeal cancer. Expert Rev Anticancer Ther 10:1439–1450
2. Zong J, Huang Q, Guo Q, Pan J (2016) Evolution of the Chinese staging system for nasopharyngeal carcinoma. Chin Clin Oncol 5:19
3. Fang W, Li X, Jiang Q, Liu Z, Yang H, Wang S, Xie S, Liu Q, Liu T, Huang J, Xie W, Li Z, Zhao Y, Wang E, Marincola FM, Yao K (2008) Transcriptional patterns, biomarkers and pathways characterizing nasopharyngeal carcinoma of Southern China. J Transl Med 6:32
4. Cao SM, Simons MJ, Qian CN (2011) The prevalence and prevention of nasopharyngeal carcinoma in China. Chin J Cancer 30:114–119
5. He W, Ju D, Jie Z, Zhang A, Xing X, Yang Q (2018) Aberrant CpG-methylation affects genes expression predicting survival in lung adenocarcinoma. Cancer Med 7:5716–5726
6. Kordowski F, Kolarova J, Schafmayer C, Buch S, Goldmann T, Marwitz S, Kugler C, Scheufele S, Gassling V, Nemeth CG, Brosch M, Hampe J, Lucius R, Roder C, Kalthoff H, Siebert R, Ammerpohl O, Reiss K (2018) Aberrant DNA methylation of ADAMTS16 in colorectal and other epithelial cancers. BMC Cancer 18:796
7. Song SH, Jeon MS, Nam JW, Kang JK, Lee YJ, Kang JY, Kim HP, Han SW, Kang GH, Kim TY (2018) Aberrant GATA2 epigenetic dysregulation induces a GATA2/GATA6 switch in human gastric cancer. Oncogene 37:993–1004
8. Mochizuki D, Misawa Y, Kawasaki H, Imai A, Endo S, Mima M, Yamada S, Nakagawa T, Kanazawa T, Misawa K (2018) Aberrant epigenetic regulation in head and neck cancer due to distinct EZH2 overexpression and DNA hypermethylation. Int J Mol Sci 19:E3707
9. Feng L, Wang R, Lian M, Ma H, He N, Liu H, Wang H, Fang J (2016) Integrated analysis of long noncoding RNA and mRNA expression profile in advanced laryngeal squamous cell carcinoma. PLoS One 11:e0169232
10. Tian F, Yip SP, Kwong DL, Lin Z, Yang Z, Wu VW (2013) Promoter hypermethylation of tumor suppressor genes in serum as potential biomarker for the diagnosis of nasopharyngeal carcinoma. Cancer Epidemiol 37:708–713

11. Fang WY, Liu TF, Xie WB, Yang XY, Wang S, Ren CP, Deng X, Liu QZ, Huang ZX, Li X, Ding YQ, Yao KT (2005) Reexploring the possible roles of some genes associated with nasopharyngeal carcinoma using microarray-based detection. Acta Biochim Biophys Sin (Shanghai) 37:541–546
12. Larson RA, Wang Y, Banerjee M, Wiemels J, Hartford C, Le Beau MM, Smith MT (1999) Prevalence of the inactivating 609C–%3eT polymorphism in the NAD(P)H:quinone oxidoreductase (NQO1) gene in patients with primary and therapy-related myeloid leukemia. Blood 94:803–807
13. Schulz WA, Krummeck A, Rosinger I, Schmitz-Drager BJ, Sies H (1998) Predisposition towards urolithiasis associated with the NQO1 null-allele. Pharmacogenetics 8:453–454
14. Guo X, Zeng Y, Deng H, Liao J, Zheng Y, Li J, Kessing B, O'Brien SJ (2010) Genetic polymorphisms of CYP2E1, GSTP1, NQO1 and MPO and the risk of nasopharyngeal carcinoma in a Han Chinese population of Southern China. BMC Res Notes 3:212
15. Rodriguez-Pinilla SM, Sarrio D, Honrado E, Hardisson D, Calero F, Benitez J, Palacios J (2006) Prognostic significance of basal-like phenotype and fascin expression in node-negative invasive breast carcinomas. Clin Cancer Res 12:1533–1539
16. Vignjevic D, Schoumacher M, Gavert N, Janssen KP, Jih G, Lae M, Louvard D, Ben-Ze'ev A, Robine S (2007) Fascin, a novel target of beta-catenin-TCF signaling, is expressed at the invasive front of human colon cancer. Cancer Res 67:6844–6853
17. Wu D, Chen L, Liao W, Ding Y, Zhang Q, Li Z, Liu L (2010) Fascin1 expression predicts poor prognosis in patients with nasopharyngeal carcinoma and correlates with tumor invasion. Ann Oncol 21:589–596
18. Li YQ, Lu JH, Bao XM, Wang XF, Wu JH, Hong WQ (2015) MiR-24 functions as a tumor suppressor in nasopharyngeal carcinoma through targeting FSCN1. J Exp Clin Cancer Res 34:130
19. Gartel AL, Radhakrishnan SK (2005) Lost in transcription: p21 repression, mechanisms, and consequences. Cancer Res 65:3980–3985
20. Rodriguez R, Meuth M (2006) Chk1 and p21 cooperate to prevent apoptosis during DNA replication fork stress. Mol Biol Cell 17:402–412
21. Tsao SW, Tsang CM, Lo KW (2017) Epstein-Barr virus infection and nasopharyngeal carcinoma. Philos Trans R Soc Lond B Biol Sci 372:20160270
22. Robertson KD (2005) DNA methylation and human disease. Nat Rev Genet 6:597–610
23. Zhang Y, Fan J, Fan Y, Li L, He X, Xiang Q, Mu J, Zhou D, Sun X, Yang Y, Ren G, Tao Q, Xiang T (2018) The new 6q27 tumor suppressor DACT2, frequently silenced by CpG methylation, sensitizes nasopharyngeal cancer cells to paclitaxel and 5-FU toxicity via beta-catenin/Cdc25c signaling and G2/M arrest. Clin Epigenet 10:26
24. Dai W, Cheung AK, Ko JM, Cheng Y, Zheng H, Ngan RK, Ng WT, Lee AW, Yau CC, Lee VH, Lung ML (2015) Comparative methylome analysis in solid tumors reveals aberrant methylation at chromosome 6p in nasopharyngeal carcinoma. Cancer Med 4:1079–1090