Clinical Epigenetics

RESEARCH

Open Access

DNA methylation heterogeneity correlates with field cancerization and prognosis in lung adenocarcinoma patients

Ying Zhou[1], Jing Zhang[1], Yang He[2], Yun Wang[1], Bing Li[3], Tengfei Zhu[3] and Yanjun Su[4*]

Abstract

Background Lung adenocarcinoma (LUAD) is the most common histological subtype of lung cancer. The distinctive genetic and epigenetic modifications in tumors and paired non-malignant samples, such as adjacent peri-tumor and tumor-distant normal lung tissues, have not been adequately studied.

Methods We recruited 57 patients with resectable stage I-III LUAD and collected matched samples of the primary tumor, peri-tumoral tissues, and tumor-distant normal lung tissue. We performed bisulfite sequencing using a custom methylation panel to profile DNA methylation levels and obtained somatic variation landscape through targeted next-generation sequencing (NGS). We attempted to identify differential methylation blocks (DMBs) between the tumor, peri-tumor, and normal tissues.

Results We analyzed the DNA methylation patterns of matched tumor, peri-tumor, and normal lung tissue samples from 57 LUAD patients. No significantly different methylation blocks were found between peri-tumoral and normal tissues, while they both exhibited distinct methylation profiles compared to tumor tissues. A total of 1329 tumor-specific DMBs, which are potentially associated with aberrant gene expression in LUAD, were identified. Utilizing a consensus clustering algorithm, we classified the tumor samples into two subgroups (C1 and C2) based on distinct methylation profiles, independent of the patient's sex, tumor stage, smoking history, and tumor cell fraction. The C2 subgroup exhibited a higher malignancy density ratio (MD ratio), suggesting a more pronounced degree of field cancerization, while the C1 subgroup was characterized by a higher frequency of *EGFR* mutations. The DMBs between the two subgroups were enriched in the calcium signaling pathway. Notably, *P2RX2* shows significant hypermethylation in the C2 subgroup, and its low expression in the external The Cancer Genome Atlas (TCGA) cohort may correlate with reduced overall survival in LUAD patients.

Conclusion Our findings revealed distinct methylation patterns between tumor and pre-malignant samples, such as peri-tumor and normal tissues. Moreover, our study suggests that distinct clustering based on DNA methylation may indicate different prognoses in LUAD patients.

Keywords Lung adenocarcinoma, DNA methylation, Field cancerization, Calcium signaling, *P2RX2*

*Correspondence:
Yanjun Su
yanjunsu@yeah.net
Full list of author information is available at the end of the article

Introduction

Lung cancer remains the leading cause of cancer-related deaths worldwide. Lung adenocarcinoma (LUAD) is the most common histological subtype of lung cancer, accounting for 50% of all lung cancer cases [1]. Epigenetic aberrations, as well as genetic mutations, are known to be crucial in the onset and progression of cancer [2]. The combined effect of both genetic and epigenetic alterations facilitates the development of human cancer [3, 4]. DNA methylation, a well-studied epigenetic mechanism, involves the addition of a methyl group to the fifth carbon of cytosine in DNA. Aberrant DNA methylation, characterized by global hypomethylation and dense hypermethylation of gene regulatory CpG islands, is regarded as a hallmark of carcinogenesis [5].

There is an emerging recognition that genetic and epigenetic alterations in cancer are interconnected rather than isolated. Hypermethylation of tumor suppressor genes, such as *CDKN2A* encoding p16, has been recognized as an alternative mechanism of tumor-suppressor inactivation [6–8]. Conversely, oncogene promoter hypomethylation can upregulate gene expression, similar to genomic amplification or oncogene translocation. Several tumor-related genes, including *CDKN2A, RASSF1A, RARbeta,* and *MGMT*, have been reported to be methylated in lung cancer [6, 9–11]. DNA methylation profiling in lung cancer revealed extensive hypomethylation and tumor-specific hypermethylation of CpG islands [12–14]. Despite the promising findings of these studies in lung cancer, comparative studies with matched normal tissues are scarce. Furthermore, most studies have focused on genetic or epigenetic aspects in isolation, with limited exploration of their interaction.

Field cancerization refers to the phenomenon where normal-appearing tissues or cells are affected by carcinogenic alterations, thereby becoming more susceptible to malignant transformation [15]. Genetic and epigenetic alterations associated with tumors can be identified in non-tumor cells surrounding the tumor, blurring the definition of "normal tissue" in the context of tumor proximity [16, 17]. Previous studies have demonstrated that aberrant DNA methylation is associated with the early development of lung cancer, suggesting that epigenetic analysis of the tumor-adjacent region may serve as a potential tool for assessing field cancerization and evaluating the risk of malignant transformation [15, 18, 19]. A study by Yang et al. developed a novel individualized scoring system (known as the malignant density ratio, MD ratio) based on methylation detection to quantify the degree of field cancerization in tumor-adjacent tissues, and they found that the MD ratio is an independent predictor of recurrence risk in early-stage lung adenocarcinoma [20]. This quantitative analysis of field cancerization offers better reproducibility, can be extended to more clinical scenarios, and better reflects tumor heterogeneity, aiding in understanding the impact of tumors on the peri-tumoral environment.

In this study, we evaluated and compared the genomic and DNA methylation landscapes of matched primary tumor, peri-tumor, and tumor-distant normal lung tissues from untreated LUAD patients and calculated the MD ratio to quantify the degree of field cancerization. We characterized subgroups with distinct tumoral DNA methylation patterns and analyzed the biological differences between subgroups, aiming to discover the prognostic signature of LUAD.

Method

Patients

Patients were enrolled following the inclusion criteria: (1) newly diagnosed with lung adenocarcinoma via histological examination at Tianjin Cancer Hospital Airport Hospital (Tianjin, China) from December 2017 to June 2021; (2) underwent lung surgery; (3) provided matched samples of primary tumor, peri-tumor (about 2 cm from the tumor), and tumor-distant normal lung tissue (about 5 cm from the tumor); (4) the tumor cell proportion (tumor percent) in tumor tissues, assessed by pathologists following hematoxylin and eosin (H&E) staining, should meet the detection criteria, while the peri-tumoral and normal tissues were confirmed to be cancer-free.

DNA extraction and methylation profiling

DNA samples from formalin-fixed, paraffin-embedded (FFPE) tissues were extracted using the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany) and quantified using the Qubit dsDNA assay (Life Technologies, Carlsbad, CA, USA). DNA methylation profiling was performed as previously described [21]. Briefly, the bisulfite sequencing (BS-seq) libraries were prepared via the brELSA™ method. Custom-designed methylation profiling RNA baits targeting 80,672 CpG sites spanning over 1.05 megabases of the human genome were utilized for the enrichment of regions of interest. The libraries were quantified and sequenced on a NovaSeq 6000 (Illumina, San Diego, CA, USA) with an average depth of 1000×. Further bioinformatic analyses included adaptor sequence removal, low-quality base filtering, paired-end read alignment and merging, and methylation block construction.

Somatic variation sequencing

We performed targeted next-generation sequencing (NGS) with a 520 cancer-related gene panel (OncoScreen plus, Burning Rock Biotech, Guangzhou, China, Supplemental Table S1). Somatic mutations were called using

optimized bioinformatics pipelines that can accurately report various cancer-related genetic alterations, including single-nucleotide variants (SNVs), insertion-deletion variants (indels), copy number variants (CNVs), and genomic rearrangements, as described previously [22]. The tumor mutational burden (TMB) per patient was calculated as the ratio of non-synonymous mutations to the panel's total coding region size [23]. Tools utilized are detailed in Supplemental Table S2.

Identification of differential methylation blocks (DMBs) and sample clustering

The 80,672 CpG sites included in the panel were grouped into 8312 methylation blocks as described previously [24]. We applied a region-defined algorithm considering the co-methylation effect among adjacent CpG sites. Tools utilized are detailed in Supplemental Table S2. To estimate the predefined coefficients of the algorithm, a series of methylation data of different tissues were used with the same panel in this study. Methylation blocks were defined as genomic regions consisting of neighboring CpG sites that were close in distance and correlated in methylation levels. Briefly, the methylation frequency differences between each pair of CpG sites were calculated by Pearson's correlation analysis and then normalized against genomic distance and methylation level variance. Blocks with a |fold change| > 2 and a false discovery rate (FDR, Benjamini-Hochberg-corrected) < 0.05 between comparative groups were defined as DMBs. Tumor-specific DMBs were ranked according to the median absolute deviation (MAD) across diverse patient tumor samples. The clustering of tumor samples was conducted with the ConsensusClusterPlus R package based on the top 100 DMBs with the highest MAD.

Pathway enrichment analysis

For KEGG terms, c2.cp.kegg.v7.4.entrez.gmt and c2.cp.v7.4.entrez.gmt were used separately. We employed the R package "clusterProfiler" for Gene Ontology (GO) annotations and visualization of enriched GO terms, including biological process, cellular component, and molecular function analysis.

Quantification of the field cancerization in tumor-adjacent tissues

We used a previously reported method to quantify the degree of field cancerization in adjacent tissues [20]. Briefly, the baseline methylation signatures of normal tissues and tumors were estimated via maximum likelihood estimation, and the malignancy density ratio (MD ratio) of peri-tumoral tissues was calculated with a mixed beta-binomial model, reflecting the proportion of malignant methylation signals in the peri-tumoral tissue shared by the corresponding tumor tissue.

Statistical analysis

Statistical analysis was performed using R version 4.1.0. The Fisher's exact and nonparametric tests were used to compare categorical data, and the Wilcoxon test was used to analyze differences in the TMB and gene expression between groups. P values < 0.05 were considered to indicate statistical significance.

Results

Patient characteristics

57 LUAD patients who underwent lung resection were enrolled in this study. The median age was 60 years (range 51.5–65 years), and 47.4% of patients were male. Twenty-three patients (40.4%) were smokers, and the majority ($n = 48$, 84.2%) had stage I disease. The clinicopathological characteristics of all patients are summarized in Table 1.

The methylation profile of LUAD tumors is different from that of non-tumor tissues

We assessed the DNA methylation patterns in 171 FFPE samples across 8312 blocks and compared the methylation profiles among the tumor, peri-tumor, and normal lung tissues of each patient. Generally, 84% of the blocks were gene-associated, with 59% in promoters, 7% in exons, and 18% in introns (Fig. 1A). Principal component analysis (PCA) indicated that the peri-tumor and normal

Table 1 Clinicopathological characteristics of the enrolled patients

Characteristic	Total Cohort ($n = 57$)	
	Number	%
Age		
Median (IQR)	60.0 (51.5, 65.0)	
Sex		
Male	27	47.4
Female	28	49.1
N/A	2	3.5
History of smoking		
Yes	23	40.4
No	33	57.9
Unknown	1	1.8
Staging		
IA	43	75.4
IB	5	8.8
IIB	3	5.3
III	6	10.5

IQR interquartile range, *N/A* not available

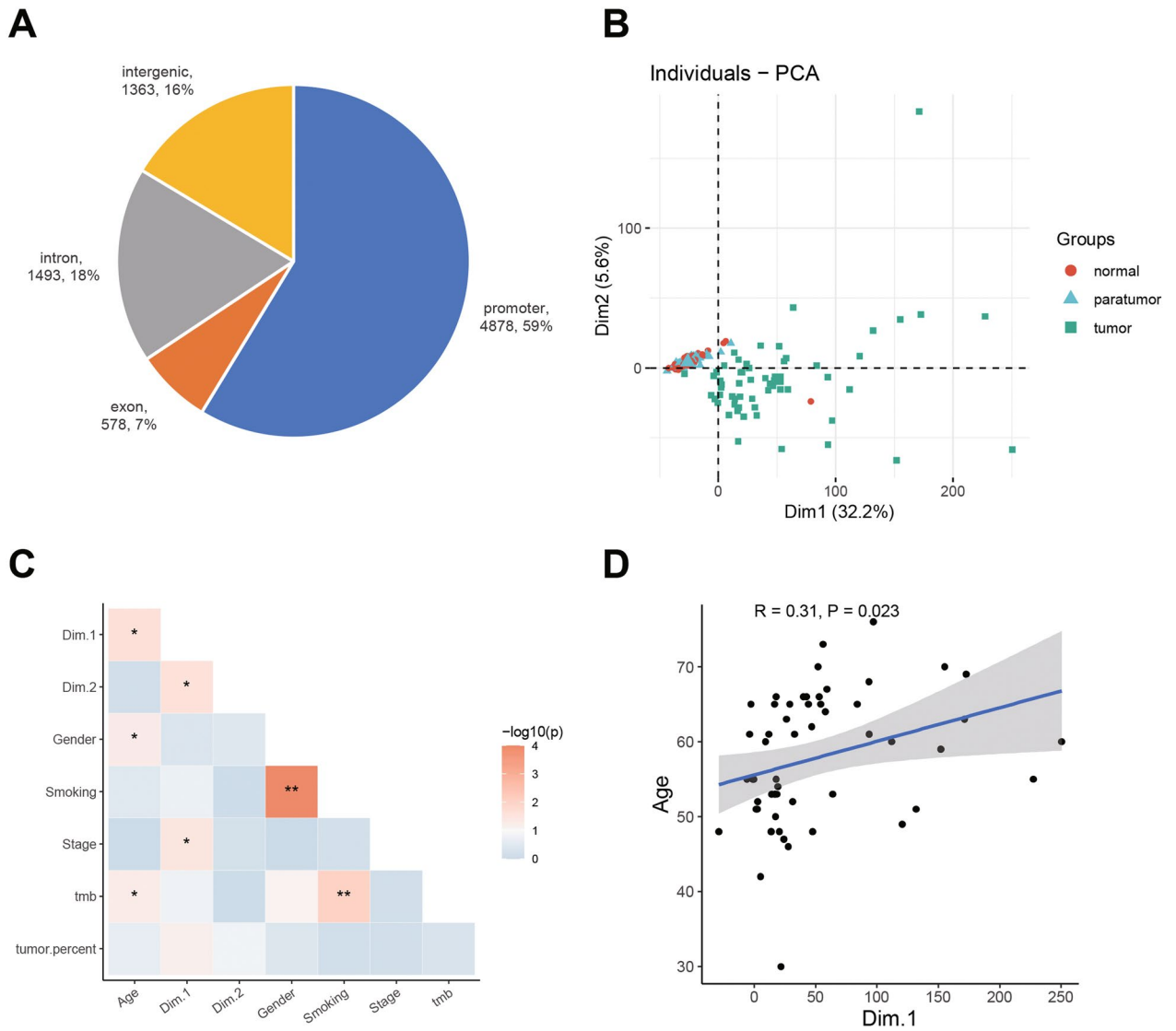

Fig. 1 Analysis of DNA methylation in LUAD patient tissues. **A** Distribution of DNA methylation blocks across the gene body. **B** Principal component analysis (PCA) of methylation profiles of various samples. **C** Analysis of the relationships between the top two principal components of the methylation profiles (Dim.1, Dim.2) and clinical characteristics. **D** Correlation between Dim.1 and patient age. *: $P < 0.05$, **: $P < 0.01$

tissues had similar methylation profiles, while there was a clear distinction between these two groups and tumor tissues (Fig. 1B). Heterogeneous methylation profiles were observed within tumor tissues (the first principal component [Dim.1] = 32.2%). We also evaluated the correlations between the first two principal components (Dim.1, and Dim.2) and patient clinical characteristics (Fig. 1C). There was a significant but weak positive correlation between patient age and Dim.1 (Fig. 1D).

We then performed a receiver operating characteristic (ROC) analysis based on Dim.1 to assess the efficacy of DNA methylation as a tumor identification biomarker. DNA methylation could distinguish tumor tissues from peri-tumor or normal tissues with significant sensitivity and specificity, as demonstrated by the area under the curve (AUC) values of 0.971 and 0.984, respectively (Supplemental Figure S1A, S1B). However, methylation signals could hardly distinguish between peri-tumor and normal tissues (AUC = 0.477, Supplemental Figure S1C).

Identification and functional analysis of differential methylation blocks in tumors

We identified 1329 tumor-specific differential methylation blocks (DMBs), including 1311 hypermethylated and 18 hypomethylated blocks, in contrast to normal tissues (Fig. 2A) In comparison to peri-tumoral tissues, the

numbers of hyper- and hypomethylated DMBs in tumors were 1350 and 18, respectively (Fig. 2B). The majority of DMBs identified when comparing tumors to normal tissues (1195 hypermethylated and 15 hypomethylated blocks) are also observed as DMBs when comparing tumors to peri-tumoral tissues (Fig. 2C). No DMBs were found between the peri-tumor and normal tissues (data not shown).

To explore the related functions and pathways of these tumor-specific hypermethylated DMBs, we conducted enrichment analysis with Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway information (Fig. 2D). The most significantly enriched biological process (BP) terms were associated with cell differentiation, encompassing pattern specification process, regionalization, embryonic organ development/morphogenesis, and cell fate commitment. Regarding cellular component (CC), the most enriched terms were mostly related to synapses and ion channels. The enriched molecular function (MF) terms extended this trend, including several ion channel-related entries and terms associated with DNA-binding transcription repressors. KEGG pathway analysis to identify activated and suppressed pathways in tumor versus normal tissues indicated that hypermethylated DMBs were enriched in neuroactive ligand-receptor interactions and calcium signaling pathways (Fig. 2E). Pathways related to MAPK signaling and cell adhesion were markedly enriched with hypermethylated DMBs, suggesting potential dysregulation of these processes. The analysis of hypermethylated DMBs located in promoters revealed a similar pattern (Supplemental Fig. 1C).

Additionally, we investigated whether the tumor-specific hypermethylation of DMBs might impact gene expression. The hypermethylated DMBs affected 396 genes. By analyzing the expression of these genes in the TCGA-LUAD dataset, we discovered that 81 genes (20.5%), such as *CDO1* and *ITGA8*, exhibited downregulated expression in tumor samples (Fig. 2F, Supplemental Table S3).

Subtype classification in lung adenocarcinoma revealed by differential methylation blocks

We divided the tumor samples into two distinct subtypes based on the DMBs exhibiting the most significant heterogeneity across patients (Fig. 3A). The comparison of clinical characteristics between the two subtypes did not reveal significant differences (Fig. 3B, Supplemental Figure S3A-E), suggesting that the clustering predominantly reflects the inherent heterogeneity of the tumor. Notably, Cluster 2 (C2) exhibited a significantly higher malignancy density ratio (MD ratio) compared to Cluster 1 (C1, $P=0.013$), suggesting a greater degree of tumor aggressiveness in C2 (Fig. 3C).

We performed a comprehensive analysis of genomic alterations in 36 tumor samples, including 20 from the C1 subtype and 16 from the C2 subtype. Mutations were detected in 33 samples (91.7%, 19 from C1, 14 from C2), including *EGFR* ($n=22$, 61%), *TP53* ($n=16$, 44%), and *KRAS* ($n=5$, 14%, Fig. 3D). Notably, a greater prevalence of *EGFR* mutations was observed in the C1 subtype than in the C2 subtype (17/20 vs. 5/16, $P<0.05$, Fig. 3E). The tumor mutational burden (TMB) also showed an increasing trend in the C2 subtype (4.99 vs. 1.99 mut/MB), although the difference did not reach statistical significance ($P=0.080$, Supplemental Figure S3F).

To further investigate the role of methylation in tumor progression, we compared the DMBs between subtypes. 151 hypermethylated and 2 hypomethylated DMBs were found in the C2 subtype compared to the C1 subtype (Fig. 4A), and KEGG pathway analysis indicated enrichment in the calcium signaling pathway (Fig. 4B). Among the DMBs associated with this pathway (Supplemental Table S4), the *P2RX2* promoter (block br1029) was hypermethylated in the C2 subtype ($P<0.001$, Fig. 4C). Analysis of the public database revealed significantly lower expression of *P2RX2* in tumor tissues (Fig. 4D), with this downregulation becoming more pronounced in advanced stages (Fig. 4E). This finding indicates a potential role for *P2RX2* in oncogenesis and progression. In the TCGA-LUAD cohort, patients with lower *P2RX2* expression had significantly shorter overall survival than those with higher *P2RX2* expression (HR 0.64, $P<0.01$, Fig. 4F), suggesting that the downregulation of *P2RX2*, including promoter methylation, may be associated with more aggressive tumor features.

(See figure on next page.)

Fig. 2 Differentially methylated blocks (DMBs) in tumor tissues compared with peri-tumor and normal tissues. **A** Volcano plot of DMBs between tumor and peri-tumoral tissues. **B** Volcano plot of DMBs between tumor tissues and normal lung tissues. **C** Venn diagram depicting the distribution of DMBs across various tissues. **D** Gene Ontology enrichment of genes linked to tumor-specific hypermethylated DMBs. **E** KEGG pathway enrichment for genes linked to tumor-specific hypermethylated DMBs. **F** Expression of selected genes associated with hypermethylated DMBs in LUADs and normal tissues within the public datasets. *: $P<0.05$

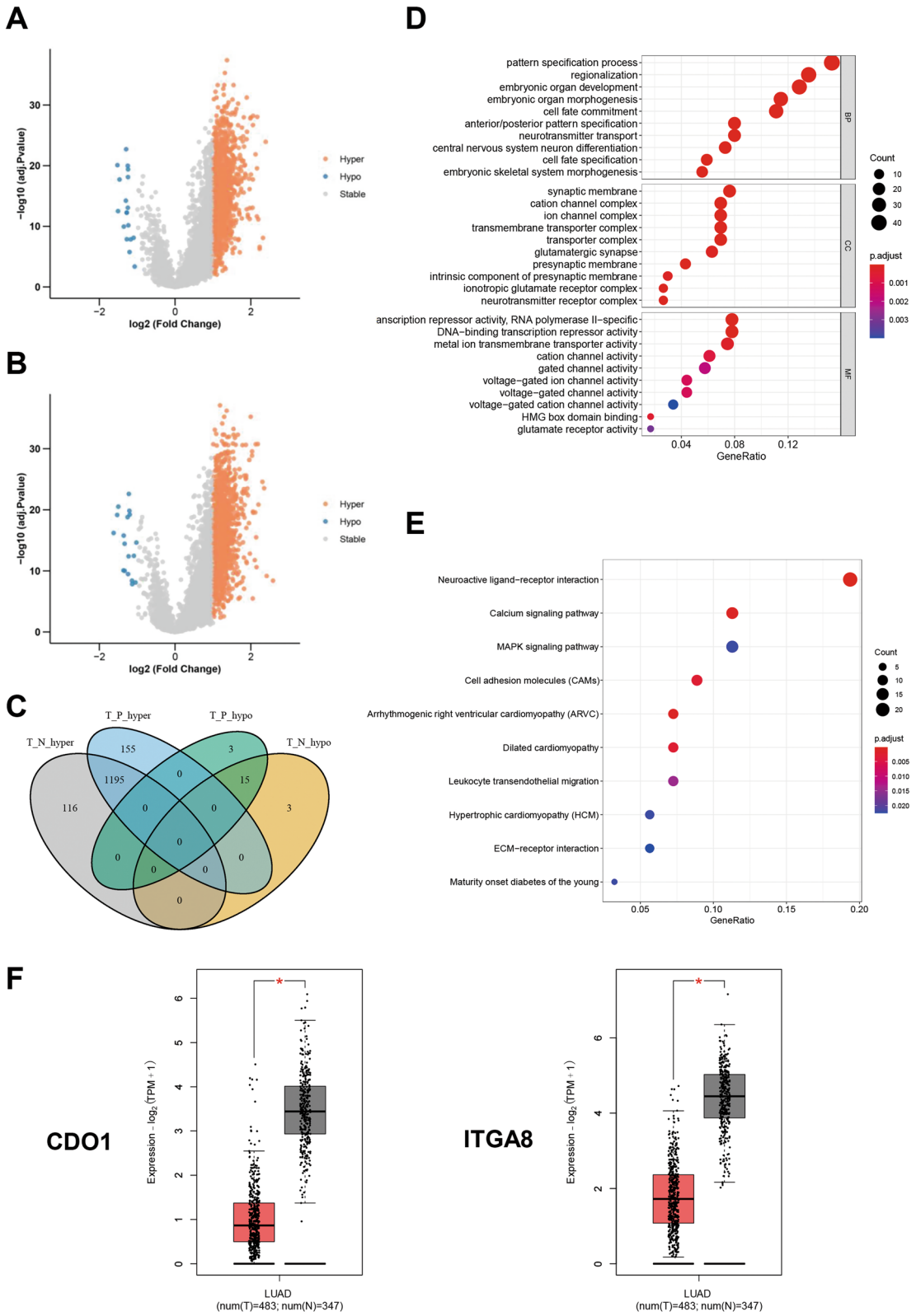

Fig. 2 (See legend on previous page.)

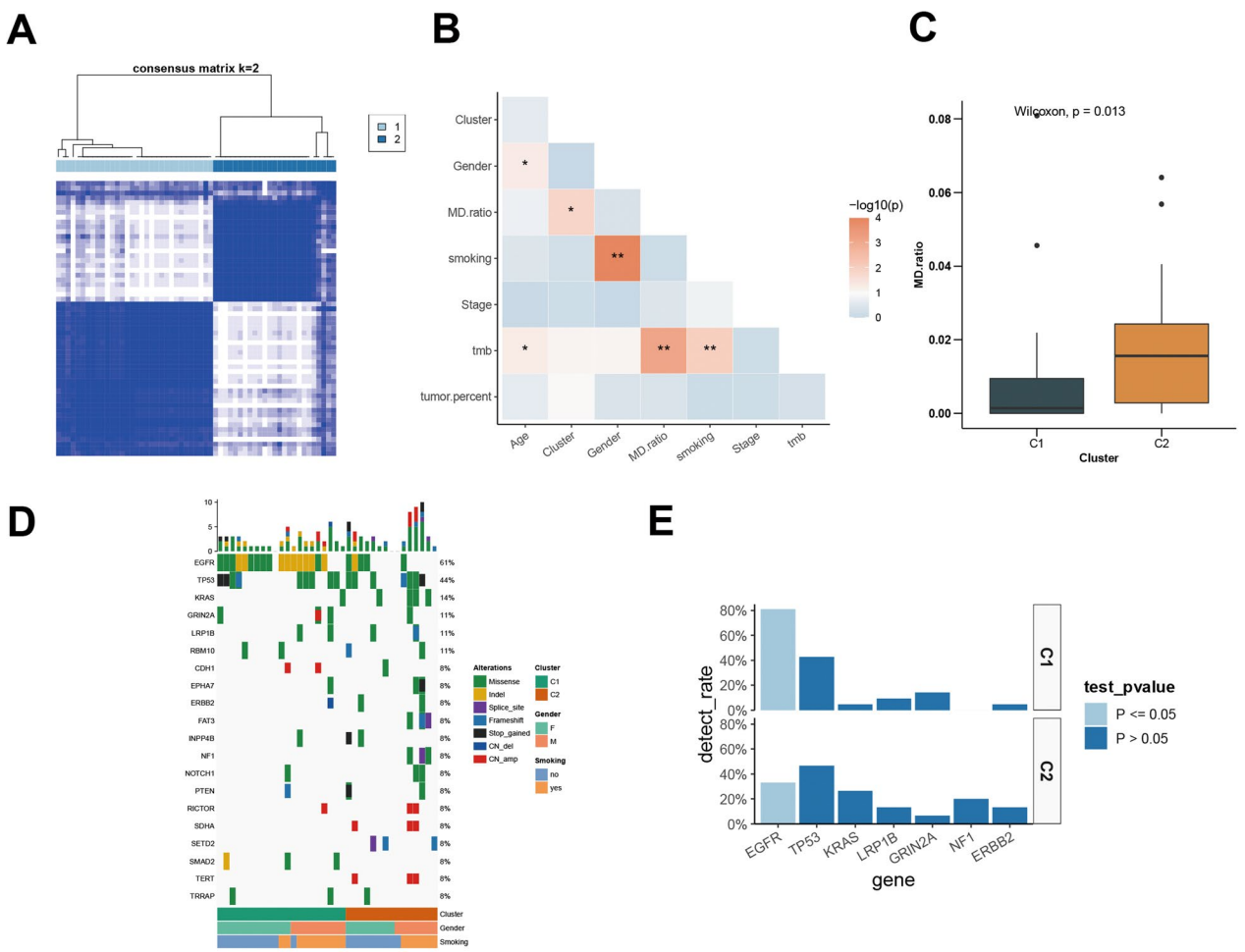

Fig. 3 DMB-based binary clustering of tumor tissues and somatic mutation contributions in subtypes. **A** Consensus clustering of tumor tissue samples. **B** Relationships between subtypes and patient clinical features/molecular markers. **C** Distribution of the malignancy density ratio (MD ratio) across subtypes. **D** Somatic mutation landscape in the tested tumors ($n = 44$). **E** Detection ratios of driver gene mutations in the two subtypes. *: $P < 0.05$, **: $P < 0.01$

Discussion

Because of its stability, reversibility, and easy detectability, changes in DNA methylation have attracted clinical attention as powerful diagnostic, prognostic, and predictive biomarkers, including in lung cancer [24]. LUAD is the most common subtype of lung cancer, and it is also a highly heterogeneous disease. Previous research has shown that methylation profiles in LUAD patients exhibit variations related to demographic and clinical characteristics [25]. In the presented study, we identified the DNA methylation landscape of tumors, adjacent non-cancerous tissues, and normal lung tissues from 57 early-stage LUAD patients. The tumor tissues exhibited specific methylation characteristics distinct from those of adjacent or normal tissues and demonstrated notable inter-patient heterogeneity. This heterogeneity may be partly attributed to differences in patient age and tumor stage and may represent a "DNA methylation age" in tumor tissues elucidated in previous studies [25, 26].

The methylation blocks we detected were primarily located in cancer-related, CpG-rich gene promoter regions. Most tumor-specific DMBs were hypermethylated compared to those in peri-tumor or normal tissues, consistent with previous findings on DNA methylation in lung cancer [12, 13]. The hypermethylation of promoters and CpG islands may inhibit the expression of nearby genes, thereby regulating the biological processes of cancer [27]. Among the genes affected by the hypermethylated DMBs we identified, some were significantly downregulated in TCGA-LUAD tumors. *CDO1* is a representative tumor suppressor gene that undergoes promoter hypermethylation and downregulated expression

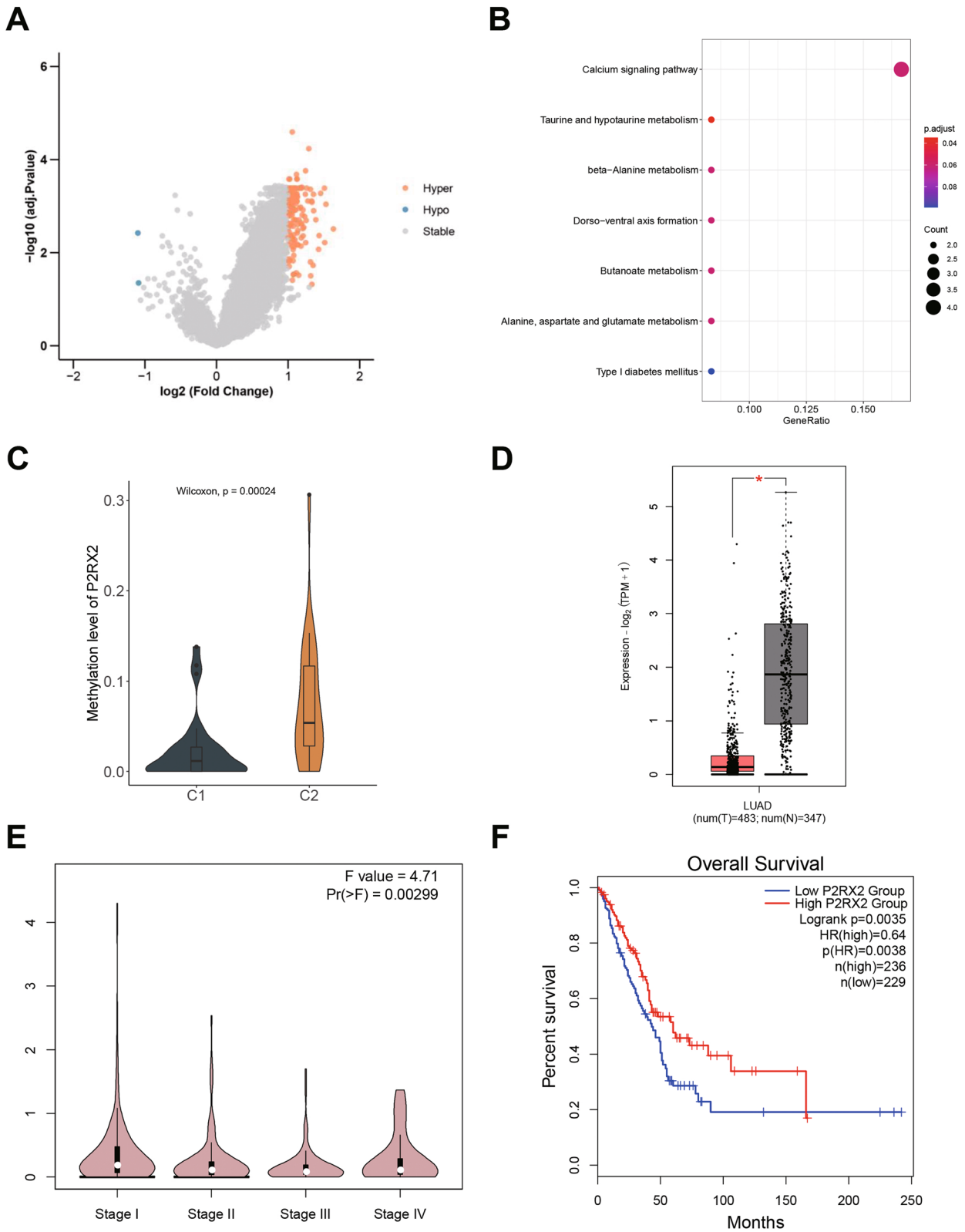

Fig. 4 Analysis of DMBs and associated genes between subtypes. **A** Volcano plot of DMBs between tissues of different subtypes. **B** KEGG enrichment of genes associated with DMBs between subtypes. **C** Distribution of *P2RX2* promoter (block br1029) methylation levels in the two subtypes. **D** Expression of *P2RX2* in LUAD tissues and normal tissues within public datasets. **E** Expression of *P2RX2* in LUADs at different stages. **F** Overall survival analysis of TCGA-LUAD patients grouped based on tumor *P2RX2* expression. *: $P < 0.05$

in various cancers[28–33]. In lung cancer, methylation of the *CDO1* promoter region is considered a notable biomarker [34–36], potentially influencing tumor proliferation through the regulation of cysteine metabolism [37]. *ITGA8*, which encodes an integrin subunit, is implicated in clear cell renal carcinoma and colorectal cancer [38, 39]. A recent study by Li et al. suggested that lower expression of *ITGA8* can alter the immune microenvironment and stemness of lung adenocarcinoma [40]. Our research did not assess the efficacy of LUAD-specific methylation features as cancer detection biomarkers, nor did it explore the functions of the affected genes in cancer progression. However, these findings may represent promising avenues for future investigations.

Given the relative ease of DNA sample collection, several studies have attempted molecular subtyping and risk stratification of cancer patients utilizing DNA methylation. Guidry et al. divided 88 LUAD patients into six distinct molecular subtypes based on whole-genome methylation sequencing, revealing variations in clinical characteristics, driver mutations, immune microenvironments, and survival [25]. Xu et al. identified seven molecular subtypes with public methylation microarray data and further distinguished two groups with different prognoses using 33 methylation sites [41]. Yu et al. categorized TCGA LUAD samples into three subgroups based on 40 sites and conducted a multifaceted comparative analysis [42]. Other molecular subtype investigations have integrated DNA methylation data with additional omics data [43, 44]. Our research clustered the cohort based on tumor-specific DMBs and generated a binary classification model independent of clinical features. Aberrations in DNA methylation are considered early events in carcinogenesis and are thus suitable for detecting field cancerization in adjacent tissues, a concept fundamental to cancer progression that was previously applied in research [20]. Employing a similar methodology, we assessed the degree of field cancerization in two clusters and found that the C2 subtype exhibited greater invasiveness. This variation may lead to differences in patient prognosis, which needs to be supplemented with more extensive outcome data.

We discovered that the C2 subtype was enriched with hypermethylation events related to calcium signaling pathway genes compared to the C1 subtype, and this pattern was also evident when comparing tumors with non-tumor tissues. Calcium is a major second messenger in cells and plays significant roles in cell proliferation, apoptosis, and oncogenesis [45, 46]. The activation of receptor tyrosine kinases (RTKs), such as EGFR, mediates the regulation of the calcium signaling pathway. These enzymes can generate inositol 1,4,5-trisphosphate by activating downstream phospholipase Cc isoforms, subsequently triggering the release of calcium ions stored in the endoplasmic reticulum into the cytoplasm [46–48]. Given the higher incidence of *EGFR* mutations in the C1 subtype than in the C2 subtype, the differential methylation within the calcium signaling pathway might suggest distinct strategies employed by tumors in the regulation of cytoplasmic calcium. Additionally, cytoplasmic calcium ions can boost the turnover of peripheral adhesions and facilitate the formation of focal adhesions, thereby enhancing the motility of cancer cells [45, 47]. This mechanism may play a key role in field cancerization. Nevertheless, the impact of the observed methylation differences on protein expression needs to be confirmed, and further validation through cytological experiments is also essential.

Among the top DMBs-associated genes identified, several have been previously reported to be associated with cancer. *RYR3* mutations have been detected in over 20% of NSCLC patients, potentially correlating with younger patient age, smoking history, higher TMB, and distant recurrence [48, 49]. *ADCY2* germline mutations are related to pulmonary function and smoking cessation ability in genome-wide association studies [50, 51]. In colorectal cancer, *ADCY2* mutations are associated with metastasis [52], while its overexpression is related to poor prognosis [53]. As co-family members of *ADCY2*, *ADCY4* and *ADCY8* are subject to hypermethylation-induced suppression in NSCLC [54]. *PDE1C* encodes an enzyme that regulates the proliferation and migration of vascular smooth muscle cells, and it is highly expressed in abdominal aortic aneurysms and can drive the proliferation, migration, and invasion of glioblastoma cells in vitro [55, 56]. The DMB associated with the *P2RX2* promoter exhibited significant differences in methylation levels between the two subtypes. Analysis of public expression datasets also demonstrated its prognostic value, although methylation inhibition of this gene has only been reported in presbycusis [57]. In prostate cancer, *P2RX2* expression may positively correlate with immune cell infiltration and immune checkpoint gene expression, with its low expression associated with poorer survival rates [58]. The roles of these genes in lung cancer have not been fully elucidated, and further research is needed to explore the regulatory effects of methylation on their expression and the biological processes involved.

Our study has certain limitations. First, due to the availability of samples, we were unable to obtain patient tissues for RNA analysis, preventing confirmation of how DNA methylation in our cohort regulates gene expression. Only a subset of genes identified with differential methylation showed corresponding expression level changes in the TCGA-LUAD cohort, which constrains the interpretability of our results. Second, the DMBs

identified through NGS were not validated by additional experimental methods, necessitating a careful and accurate assessment of their validity. Third, we did not obtain sufficient treatment response or survival data in our cohort for analysis. Therefore, the relationship between tumor malignancy and field cancerization or differential methylation in specific genome regions should be regarded with caution. Finally, the limited size of our study cohort means that our conclusions need further validation across a broader spectrum.

Conclusions

In summary, we delineated a unique DNA methylation landscape in lung adenocarcinoma, pinpointing tumor-specific signals potentially serving as biomarkers for early lung cancer detection. Based on methylation exhibiting notable inter-tumoral heterogeneity, we established a binary classification model related to cancer cell invasiveness and identified several potential prognostic biomarkers. Our study contributes to the field of epigenetic studies of lung adenocarcinoma, enhancing our understanding of its pathogenesis and providing biomarkers for lung cancer screening, diagnosis, and patient stratification.

Abbreviations

AUC	Area under the curve
CNV	Copy number variant
DMB	Differential methylation block
FDR	False discovery rate
FFPE	Formalin-fixed paraffin-embedded
GO	Gene Ontology
Indel	Insertion-deletion variant
IQR	Interquartile range
LUAD	Lung adenocarcinoma
MD ratio	Malignancy density ratio
N/A	Not available
NGS	Next-generation sequencing
PCA	Principal component analysis
ROC	Receiver operating characteristic
SNV	Single-nucleotide variant
TCGA	The Cancer Genome Atlas
TMB	Tumor mutational burden

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13148-025-01845-9.

Additional file1

Additional file2

Additional file3

Additional file4

Acknowledgements

Not applicable.

Author contributions

Conceptualization: YZ, YS. Data Curation: YZ, JZ, YH, YW. Formal Analysis: YZ, JZ. Investigation: YZ, YH, YS. Methodology: YZ, YW. Project Administration: YS. Resources: YH, YS. Software: JZ, YW. Supervision: YS. Visualization: JZ. Writing—Original Draft: YZ. Writing – Review and Editing: BL, TZ, YZ, YS.

Funding

Not applicable.

Availability of data and materials

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

Declarations

Ethics approval and consent to participate

This study was approved by the Institutional Review Board of Tianjin Cancer Hospital. Informed consent was obtained from all patients.

Consent for publication

Consent for publication was obtained from all patients.

Competing interests

BL and TZ are employees of Burning Rock Biotech, Guangzhou, China. Other authors declare that they have no competing interests.

Author details

[1]Department of Integrative Oncology, Tianjin Cancer Hospital Airport Hospital, Tianjin, China. [2]Department of Breast Cancer, Tianjin Cancer Hospital Airport Hospital, Tianjin, China. [3]Burning Rock Biotech, Guangzhou, China. [4]Department of Thoracic Oncology, Tianjin Cancer Hospital Airport Hospital, No.99 Dongwudao Road, Tianjin 300308, China.

Received: 13 June 2024 Accepted: 16 February 2025
Published online: 20 March 2025

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63.
2. Papanicolau-Sengos A, Aldape K. DNA methylation profiling: an emerging paradigm for cancer diagnosis. Annu Rev Pathol. 2022;17:295–321.
3. Ushijima T, Clark SJ, Tan P. Mapping genomic and epigenomic evolution in cancer ecosystems. Science. 2021;373(6562):1474–9.
4. Recillas-Targa F. Cancer epigenetics: an overview. Arch Med Res. 2022;53(8):732–40.
5. Esteller M, Dawson MA, Kadoch C, Rassool FV, Jones PA, Baylin SB. The epigenetic hallmarks of cancer. Cancer Discov. 2024;14(10):1783–809.
6. Lee JU, Sul HJ, Son JW. Promoter methylation of CDKN2A, RARbeta, and RASSF1A in non-small cell lung carcinoma: quantitative evaluation using pyrosequencing. Tuberc Respir Dis (Seoul). 2012;73(1):11–21.
7. Tam KW, Zhang W, Soh J, Stastny V, Chen M, Sun H, et al. CDKN2A/p16 inactivation mechanisms and their relationship to smoke exposure and molecular features in non-small-cell lung cancer. J Thorac Oncol. 2013;8(11):1378–88.
8. Zhao R, Choi BY, Lee MH, Bode AM, Dong Z. Implications of genetic and epigenetic alterations of CDKN2A (p16(INK4a)) in cancer. EBioMedicine. 2016;8:30–9.
9. Pankova D, Jiang Y, Chatzifrangkeskou M, Vendrell I, Buzzelli J, Ryan A, et al. RASSF1A controls tissue stiffness and cancer stem-like cells in lung adenocarcinoma. EMBO J. 2019;38(13): e100532.
10. Dutkowska A, Antczak A, Pastuszak-Lewandoska D, Migdalska-Sek M, Czarnecka KH, Gorski P, et al. RARbeta promoter methylation as an epigenetic mechanism of gene silencing in non-small cell lung cancer. Adv Exp Med Biol. 2016;878:29–38.
11. Chen L, Wang Y, Liu F, Xu L, Peng F, Zhao N, et al. A systematic review and meta-analysis: association between MGMT hypermethylation and the clinicopathological characteristics of non-small-cell lung carcinoma. Sci Rep. 2018;8(1):1439.

12. Rauch TA, Zhong X, Wu X, Wang M, Kernstine KH, Wang Z, et al. High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. Proc Natl Acad Sci U S A. 2008;105(1):252–7.
13. Heller G, Babinsky VN, Ziegler B, Weinzierl M, Noll C, Altenberger C, et al. Genome-wide CpG island methylation analyses in non-small cell lung cancer patients. Carcinogenesis. 2013;34(3):513–21.
14. Ramazi S, Daddzadi M, Sahafnejad Z, Allahverdi A. Epigenetic regulation in lung cancer. MedComm (2020). 2023;4(6):e401.
15. Curtius K, Wright NA, Graham TA. An evolutionary perspective on field cancerization. Nat Rev Cancer. 2018;18(1):19–32.
16. Wistuba II, Gazdar AF. Lung cancer preneoplasia. Annu Rev Pathol. 2006;1:331–48.
17. Tang X, Varella-Garcia M, Xavier AC, Massarelli E, Ozburn N, Moran C, et al. Epidermal growth factor receptor abnormalities in the pathogenesis and progression of lung adenocarcinomas. Cancer Prev Res (Phila). 2008;1(3):192–200.
18. Kadara H, Wistuba II. Field cancerization in non-small cell lung cancer: implications in disease pathogenesis. Proc Am Thorac Soc. 2012;9(2):38–42.
19. Kerr KM, Galler JS, Hagen JA, Laird PW, Laird-Offringa IA. The role of DNA methylation in the development and progression of lung adenocarcinoma. Dis Markers. 2007;23(1–2):5–30.
20. Yang L, Zhang J, Yang G, Xu H, Lin J, Shao L, et al. The prognostic value of a methylome-based malignancy density scoring system to predict recurrence risk in early-stage lung adenocarcinoma. Theranostics. 2020;10(17):7635–44.
21. Sui J, Wu X, Wang C, Wang G, Li C, Zhao J, et al. Discovery and validation of methylation signatures in blood-based circulating tumor cell-free DNA in early detection of colorectal carcinoma: a case-control study. Clin Epigenetics. 2021;13(1):26.
22. Li YS, Jiang BY, Yang JJ, Zhang XC, Zhang Z, Ye JY, et al. Unique genetic profiles from cerebrospinal fluid cell-free DNA in leptomeningeal metastases of EGFR-mutant non-small-cell lung cancer: a new medium of liquid biopsy. Ann Oncol. 2018;29(4):945–52.
23. Chen X, Fang L, Zhu Y, Bao Z, Wang Q, Liu R, et al. Blood tumor mutation burden can predict the clinical response to immune checkpoint inhibitors in advanced non-small cell lung cancer patients. Cancer Immunol Immunother. 2021;70(12):3513–24.
24. Liang N, Li B, Jia Z, Wang C, Wu P, Zheng T, et al. Ultrasensitive detection of circulating tumour DNA via deep methylation sequencing aided by machine learning. Nat Biomed Eng. 2021;5(6):586–99.
25. Guidry K, Vasudevaraja V, Labbe K, Mohamed H, Serrano J, Guidry BW, et al. DNA methylation profiling identifies subgroups of lung adenocarcinoma with distinct immune cell composition, DNA methylation age, and clinical outcome. Clin Cancer Res. 2022;28(17):3824–35.
26. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
27. Esteller M. Epigenetics in cancer. N Engl J Med. 2008;358(11):1148–59.
28. Brait M, Ling S, Nagpal JK, Chang X, Park HL, Lee J, et al. Cysteine dioxygenase 1 is a tumor suppressor gene silenced by promoter methylation in multiple human cancers. PLoS ONE. 2012;7(9): e44951.
29. Harada H, Soeno T, Yokoi K, Nishizawa N, Ushiku H, Hosoda K, et al. Prediction of efficacy of postoperative chemotherapy by DNA methylation of CDO1 in gastric cancer. J Surg Res. 2020;256:404–12.
30. Yokoi K, Harada H, Yokota K, Ishii S, Tanaka T, Nishizawa N, et al. Epigenetic status of CDO1 gene may reflect chemosensitivity in colon cancer with postoperative adjuvant chemotherapy. Ann Surg Oncol. 2019;26(2):406–14.
31. Tanaka Y, Kosaka Y, Waraya M, Yokota K, Harada H, Kaida T, et al. Differential prognostic relevance of promoter DNA methylation of CDO1 and HOPX in primary breast cancer. Anticancer Res. 2019;39(5):2289–98.
32. Nishizawa N, Harada H, Kumamoto Y, Kaizu T, Katoh H, Tajima H, et al. Diagnostic potential of hypermethylation of the cysteine dioxygenase 1 gene (CDO1) promoter DNA in pancreatic cancer. Cancer Sci. 2019;110(9):2846–55.
33. Nakamoto S, Kumamoto Y, Igarashi K, Fujiyama Y, Nishizawa N, Ei S, et al. Methylated promoter DNA of CDO1 gene and preoperative serum CA19-9 are prognostic biomarkers in primary extrahepatic cholangiocarcinoma. PLoS ONE. 2018;13(10): e0205864.
34. Liu B, Ricarte Filho J, Mallisetty A, Villani C, Kottorou A, Rodgers K, et al. Detection of promoter DNA methylation in urine and plasma aids the detection of non-small cell lung cancer. Clin Cancer Res. 2020;26(16):4339–48.
35. Wrangle J, Machida EO, Danilova L, Hulbert A, Franco N, Zhang W, et al. Functional identification of cancer-specific methylation of CDO1, HOXA9, and TAC1 for the diagnosis of lung cancer. Clin Cancer Res. 2014;20(7):1856–64.
36. Kwon YJ, Lee SJ, Koh JS, Kim SH, Lee HW, Kang MC, et al. Genome-wide analysis of DNA methylation and the gene expression change in lung cancer. J Thorac Oncol. 2012;7(1):20–33.
37. Kang YP, Torrente L, Falzone A, Elkins CM, Liu M, Asara JM, et al. Cysteine dioxygenase 1 is a metabolic liability for non-small cell lung cancer. Elife. 2019;8.
38. Kok-Sin T, Mokhtar NM, Ali Hassan NZ, Sagap I, Mohamed Rose I, Harun R, et al. Identification of diagnostic markers in colorectal cancer via integrative epigenomics and genomics data. Oncol Rep. 2015;34(1):22–32.
39. Lu X, Wan F, Zhang H, Shi G, Ye D. ITGA2B and ITGA8 are predictive of prognosis in clear cell renal cell carcinoma patients. Tumour Biol. 2016;37(1):253–62.
40. Li X, Zhu G, Li Y, Huang H, Chen C, Wu D, et al. LINC01798/miR-17-5p axis regulates ITGA8 and causes changes in tumor microenvironment and stemness in lung adenocarcinoma. Front Immunol. 2023;14:1096818.
41. Xu F, He L, Zhan X, Chen J, Xu H, Huang X, et al. DNA methylation-based lung adenocarcinoma subtypes can predict prognosis, recurrence, and immunotherapeutic implications. Aging (Albany NY). 2020;12(24):25275–93.
42. Yu R, Huang X, Lin J, Lin S, Shen G, Chen W. Bioinformatics analysis based on DNA methylation data identified in lung adenocarcinoma subgroups with different immune characteristics and clinical outcomes. J Thorac Dis. 2023;15(4):2184–97.
43. Yang K, Wu Y. A prognosis-related molecular subtype for early-stage non-small lung cell carcinoma by multi-omics integration analysis. BMC Cancer. 2021;21(1):128.
44. Zhao Y, Gao Y, Xu X, Zhou J, Wang H. Multi-omics analysis of genomics, epigenomics and transcriptomics for molecular subtypes and core genes for lung adenocarcinoma. BMC Cancer. 2021;21(1):257.
45. Patergnani S, Danese A, Bouhamida E, Aguiari G, Previati M, Pinton P, et al. Various aspects of calcium signaling in the regulation of apoptosis, autophagy, cell proliferation, and cancer. Int J Mol Sci. 2020;21(21).
46. Wu L, Lian W, Zhao L. Calcium signaling in cancer progression and therapy. FEBS J. 2021;288(21):6187–205.
47. Prevarskaya N, Skryma R, Shuba Y. Calcium in tumour metastasis: new roles for known actors. Nat Rev Cancer. 2011;11(8):609–18.
48. Wang Y, Chen Y, Zhang L, Xiong J, Xu L, Cheng C, et al. Ryanodine receptor (RYR) mutational status correlates with tumor mutational burden, age and smoking status and stratifies non-small cell lung cancer patient prognosis. Transl Cancer Res. 2022;11(7):2070–83.
49. Valter A, Luhari L, Pisarev H, Truumees B, Planken A, Smolander OP, et al. Genomic alterations as independent prognostic factors to predict the type of lung cancer recurrence. Gene. 2023;885: 147690.
50. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nat Genet. 2010;42(1):45–52.
51. Uhl GR, Liu QR, Drgon T, Johnson C, Walther D, Rose JE, et al. Molecular genetics of successful smoking cessation: convergent genome-wide association study results. Arch Gen Psychiatry. 2008;65(6):683–93.
52. Fang LT, Lee S, Choi H, Kim HK, Jew G, Kang HC, et al. Comprehensive genomic analyses of a metastatic colon cancer to the lung by whole exome sequencing and gene expression analysis. Int J Oncol. 2014;44(1):211–21.
53. Yu SJ, Yu JK, Ge WT, Hu HG, Yuan Y, Zheng S. SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK are prognosis-related in colorectal cancer. World J Gastroenterol. 2011;17(15):2028–36.
54. Wang XX, Xiao FH, Li QG, Liu J, He YH, Kong QP. Large-scale DNA methylation expression analysis across 12 solid cancers reveals hypermethylation in the calcium-signaling pathway. Oncotarget. 2017;8(7):11868–76.
55. Zhang C, Zhao H, Cai Y, Xiong J, Mohan A, Lou D, et al. Cyclic nucleotide phosphodiesterase 1C contributes to abdominal aortic aneurysm. Proc Natl Acad Sci U S A. 2021;118(31).
56. Rowther FB, Wei W, Dawson TP, Ashton K, Singh A, Madiesse-Timchou MP, et al. Cyclic nucleotide phosphodiesterase-1C (PDE1C) drives cell

proliferation, migration and invasion in glioblastoma multiforme cells in vitro. Mol Carcinog. 2016;55(3):268–79.
57. Bouzid A, Smeti I, Dhouib L, Roche M, Achour I, Khalfallah A, et al. Down-expression of P2RX2, KCNQ5, ERBB3 and SOCS3 through DNA hypermethylation in elderly women with presbycusis. Biomarkers. 2018;23(4):347–56.
58. Li Q, Wu B, Daba M, Gao X, Chen B, Song G, et al. Identification of calcium channel-related gene P2RX2 for prognosis and immune infiltration in prostate cancer. Dis Markers. 2022;2022:8058160.

Publisher's Note