Heliyon 11 (2025) e42581

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Blood-based biomarkers derived from tumor-informed DNA methylation analysis for lung adenocarcinoma

Pitaksin Chitta [a,b], Timothy M. Barrow [c], Atchara Dawangpa [a,b], David C. Christiani [d], Naravat Poungvarin [b], Chanachai Sae-Lee [a,b,*]

[a] Research Division, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[b] Clinical Pathology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand
[c] School of Life Sciences, University of Essex, Colchester, United Kingdom
[d] Harvard T H Chan School of Public Health, Massachusetts General Hospital/Harvard Medical School, Boston, MA, USA

ARTICLE INFO

Keywords:
DNA methylation
Blood
Tissue
Lung adenocarcinoma
Biomarkers
Molecular epidemiology

ABSTRACT

Objective: To identify robust markers of lung adenocarcinoma (LUAD) using DNA methylation profiles from blood samples informed by tissue lung adenocarcinoma.
Methods: This study analyzed 56 LUAD blood samples from patients attending clinic at Siriraj Hospital, Thailand and 51 samples from healthy participants, using 644 tumor and 59 normal tissue methylome datasets from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases for candidate gene identification. We performed comparative analysis to identify DNA methylation (DNAm) changes present in tumors that are also observable in blood, to be taken forward for validation.
Results: DNAm profiling of lung tumor datasets identified 59,639 differentially methylated positions (DMPs), of which 17,251 exhibited a negative correlation with gene expression. In blood samples, 46,680 DMPs were identified among LUAD patients, which were enriched in pathways associated with the ribosome, spliceosome, cell cycle, ubiquitin mediated proteolysis and nucleocytoplasmic transport. Comparative analysis revealed a two DMP epigenetic signature of matching changes in both tissue and blood. This signature offered high diagnostic performance in distinguishing LUAD from normal lung tissue (AUC: 0.77–0.91) and in blood samples from LUAD patients (AUC:0.92–0.96). Similarly high performance was observed in two independent tissue validation datasets (AUC:0.90–0.92).
Conclusions: Our novel two DMP signatures offer robust performance in both lung tissue and blood for the identification of LUAD.

Abbreviations: AGER, Advanced glycosylation end product-specific receptor; AHRR, Aryl hydrocarbon receptor repressor; CDKN2A, Cyclin-dependent kinase inhibitor 2A; cfDNA, Cell-free DNA; CI, Confidence intervals; DMPs, Differentially methylated positions; DNAm, DNA methylation; EGFR, Epidermal growth factor receptor; F2RL3, Coagulation factor II receptor-like 3; FUT7, Fucosyltransferase 7; GEO, Gene Expression Omnibus; GO, Gene Ontology; HOXA9, Homeobox A9; HOXD12, Homeobox D12; KEGG, Kyoto Encyclopedia of Genes and Genomes; KRAS, Kirsten rat sarcoma viral oncogene homolog; LUAD, Lung adenocarcinomas; MPO, Myeloperoxidase; NSCLC, Non-small cell lung cancer; PNPLA2, Patatin-like phospholipase domain-containing protein 2; PTGER2, Prostaglandin E receptor 2; RASSF1A, Ras association domain-containing protein 1 isoform A; ROC, Receiver operating characteristic; SHOX2, Short stature homeobox 2; SOX1, (SRY, sex determining region Y)-box 1; STX8, Syntaxin 8; TAF4, Transcription initiation factor TFIID subunit 4; TCGA, The Cancer Genome Atlas; TP53, Tumor protein p53; ZIC1, Zinc finger protein ZIC 1; ZIC4, Zinc finger protein ZIC 4.

* Corresponding author. Research Division, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.
E-mail address: Chanachai.sae@mahidol.edu (C. Sae-Lee).

https://doi.org/10.1016/j.heliyon.2025.e42581
Received 28 June 2024; Received in revised form 7 February 2025; Accepted 7 February 2025
Available online 11 February 2025

1. Introduction

Lung cancer accounts for approximately 1.8 million deaths worldwide every year [1], which constitute 25 % of all cancer-related deaths [2]. The high mortality associated with this disease is in part due to its diagnosis at an advanced stage, with more than half of cases diagnosed at Stage IV where the five-year survival rate is less than 6 %, in comparison to 59 % for those diagnosed with Stage I or II disease [3]. It is therefore an urgent clinical need to develop novel screening methodologies to detect early tumors and enhance the chances for successful treatment. However, this has been inhibited by the invasiveness of most approaches, such as biopsy, which preclude their use in screening. Subsequently there has been significant interest in the development of non-invasive methodologies to detect early genetic or epigenetic changes. Blood-based tests have been developed for lung [4] and colorectal cancers [5], while cell free DNA (cfDNA) originating from the tumor is present in the blood and has become a promising avenue of research for hard-to-biopsy tissues [6].

Epigenetics is the study of heritable changes in cellular and physiological traits that are passed down to daughter cells without altering the underlying DNA sequence [7]. These changes are primarily driven by mechanisms such as DNA methylation, histone modifications, and non-coding RNAs. Epigenetic modifications play a crucial role in development, cellular differentiation, and adaptation to environmental stimuli [8], and they are implicated in a wide range of diseases, including cancer, neurological disorders, and metabolic conditions. Unlike genetic mutations, which are permanent alterations in the DNA sequence, epigenetic changes are often reversible and can be influenced by factors such as diet [9]. While the transgenerational inheritance of 'epimutations' remains a topic of debate, DNA methylation can impact upon evolutionary genetics. Highly methylated regions of the genome are prone to mutation through spontaneous deamination of 5-methylcytosine [10], resulting in decreased abundance of CpG sites in the human genome to only 20 % of their expected frequency [11].

As with most malignancies, the epigenome shows extensive disruption in the initiation and progression of lung cancer. DNA methylation (DNAm) changes occur early in lung tumor development and are observable in Stage I disease [12], although many changes in localized tumors are not clonal [13]. Key genes implicated in cancer development are known to be epigenetically disrupted, such as the hypermethylation of *RASSF1A* [13,14], *CDKN2A* [13,15], *HOXA9* and *SOX1* [16]. At the epigenome-wide level, DNAm patterns are associated with the occurrence of genetic aberrations within key genes such as *EGFR*, *KRAS* and *TP53* mutations [17,18] and the smoking history of the patient [18]. Aberrant epigenetic marks may influence response to chemotherapeutic agents through silencing of genes conferring sensitivity [19], and subsequently, epigenetic profiling can predict patient prognosis [12,18,20].

Epigenomic aberrations have unique features which make them highly suitable to serve as diagnostic biomarkers. They can occur very early in disease development [21], and even single-gene tests can offer very high sensitivity and specificity when assessed in lung tissue [22]. However, an inherent challenge remains in identifying appropriate targets for non-invasive testing. Many of the epigenetic differences reported between the blood of cancer patients and healthy subjects are representative of changes in leukocyte populations, rather than epigenetic changes present in tumors [23]. Risk factors for cancer can also be associated with epigenetic changes in the blood, such as *AHRR* and *F2RL3* hypomethylation in response to tobacco smoke exposure [24,25]. Here, we sought to identify robust markers of lung adenocarcinoma (LUAD), the most common subtype of non-small cell lung cancer (NSCLC) worldwide, that are detectable in non-tumor tissue. Our analysis identified novel epigenetic aberrations that can serve as diagnostic biomarkers with high sensitivity and specificity.

2. Materials and methods

2.1. Examination of DNA methylation in lung tumors and normal tissue

To investigate DNA methylation patterns and signature loci in lung tissues, we utilized methylation data obtained from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases. In the available datasets employed in this study: GSE85845, GSE66836, GSE141682, GSE56044, and GSE75008 datasets were approved by the review board and regional ethics committee of the Fujian Medical University, China [26], the Oslo University Hospital-Rikshospitalet, Norway (S-05307) [27], Tongji Medical College,

Table 1
Clinical characteristics of the methylome datasets.

Dataset	Total (N)	Sample type (N (%))		Gender (N (%))		Age (year) Mean ± SD	Smoking (N (%))		
		LUAD	Normal	Male	Female		Smoker	Non-smoker	Unknown
Tissue									
GSE66836	180	161 (89.44)	19 (10.56)	80 (44.44)	100 (55.56)	NA	180 (100)	0 (0)	0 (0)
GSE85845	16	8 (50)	8 (50)	NA	NA	NA	16 (100)	0 (0)	0 (0)
TCGA	507	475 (93.69)	32 (6.31)	238 (46.94)	269 (53.06)	64.98 ± 10.27	416 (82.05)	72 (14.2)	19 (3.75)
Blood									
Thai	107	56 (52.34)	51 (47.66)	23 (21.50)	84 (78.50)	60.06 ± 14.64	20 (18.70)	87 (81.30)	0 (0)
Han Chinese	28	0 (0)	28 (100)	14 (50)	14 (50)	49.93 ± 10.46	22 (78.57)	6 (21.43)	0 (0)
Validation									
GSE56044	136	124 (91.18)	12 (8.82)	56 (41.18)	80 (58.82)	NA	110 (80.88)	26 (19.12)	0 (0)
GSE75008	80	40 (50)	40 (50)	NA	NA	NA	NA	NA	NA

N = number of samples; LUAD = lung adenocarcinoma; NA = no data available.

Huazhong University of Science and Technology, China (IEC #S099/2018) [28], Lund, Sweden (Registration no. 2004/762 and 2008/702) [17], and the University of Lübeck (Ref. 12–220 and Ref. 12–238) [29], respectively. For the TCGA dataset, Illumina HumanMethylation450K BeadChip (450K) DNA methylation microarray data from 475 LUAD tumors and 32 associated normal tissues were obtained from the Genomic Data Commons data portal (https://portal.gdc.cancer.gov/) using the TCGAbiolinks R package [30]. For the combined GEO dataset, 450K data from further two studies, GSE85845 [26] and GSE66836 [27], were obtained from GEO (https://www.ncbi.nlm.nih.gov/geo/) using GEOquery [31]. A total of 196 samples were included from GSE85845 (8 LUAD and 8 matched non-LUAD tissues) and 180 from GSE66836 (161 LUAD and 19 matched normal tissues). Validation of markers was performed using data from two 450K data sets (GSE56044 (124 LUAD and 12 matched normal tissues) and GSE75008 (40 LAUD and 40 matched normal tissues). Further details can be found in Table 1.

2.2. Epigenetic analysis of blood samples from LUAD patients

Two versions of the Illumina HumanMethylation BeadChip (850K and 900K) DNA methylation microarray were employed to measure DNA methylation in whole blood samples from 56 Thai LUAD patients and compared with methylation profiles of 51 Thai healthy individuals elderly cohort or 28 Han Chinese healthy individuals (GSE141682). The 850K and 900K microarray include more than 90 % of the probes present on the 450K array, enabling direct comparison of the blood and tumor datasets. They also provide excellent coverage of both gene promoters and enhancer regions, as well as non-coding regions important for gene regulation. Samples from LUAD patients and elderly cohort participants attending the clinical molecular pathology laboratory at Siriraj Hospital were collected, DNA extracted using QIAamp Blood Mini Kit (Germany), and DNA methylation analyzed at the Genomics Centre at University of Minnesota, USA. Ethical approval for this study (COA no. Si701/2021) was granted by the Siriraj Institutional Review Board, Siriraj Hospital, Mahidol University, Bangkok, Thailand. The required sample size for the Thai DNA samples was calculated using G*Power [32], based on a *t*-test for means with the statistical power set at 0.95.

2.3. Differentially methylated positions (DMPs)

All intensity data (.idat) files were analyzed in R Studio using the Bioconductor package, except GSE75008 and GSE83842 datasets for which DNAm levels were calculated from methylated and unmethylated signal intensities. Poor-quality probes (detection *p*-value >0.01) were removed from the analysis. For normalization, the single-sample normalized (ssNoob) method, which is part of the minfi package, was utilized. This normalization procedure is well-suited for the incremental preprocessing of individual methylation arrays, especially when integrating data from multiple generations of Infinium methylation arrays [33]. The remaining probes were then adjusted for probe-type bias for Infinium type I and type II probes [34]. SNP probes, cross-reactive probes, and probes mapping to the sex chromosomes were removed. To mitigate the potential influence of batch effects that could compromise biological findings, our high-throughput data were incorporated into adjustment and removed batch effects by the ComBat function within the sva package [35]. Furthermore, correction for cell-type heterogeneity in blood was performed with the reference-based RefbaseEWAS [36] within the ChAMP package. RefbaseEWAS uses pre-defined reference DNA methylation data to infer cell-type composition and adjust for confounding factors, particularly in blood samples, to enhance the detection of true epigenetic changes. A total of 327,667 (GEO, 450k), 320,520 (TCGA, 450k), 709,611 probes (blood (Thai), EPIC v.2), and 782,501 probes (blood (Thai and Han Chinese), EPIC) remained after quality control and filtering steps. The methylation level of each CpG site was determined by calculating the β value between 0 and 1 based on the ratio of fluorescent intensities. DMPs were identified using the limma package [37]. This package performs a pairwise comparison of β value between groups. It achieves this by applying the same general linear model to each probe individually, calculating a moderated *t*-statistic, and determining an unadjusted p-value. The moderated *t*-statistic is derived from the ratio of the beta-value for a CpG to a pooled standard error (SE). The pooling of information from all CpGs allows limma to moderate SE at individual probes, enhancing the accuracy of inference for each CpG. Probes with Benjamini-Hochberg adjusted *p*-values <0.05 (adj. *P* Value) were considered significant.

2.4. Gene expression analysis

For the gene expression analysis of LUAD, we included a dataset from TCGA consisting of 59 normal and 539 tumor samples obtained from the Genomic Data Commons data portal (https://portal.gdc.cancer.gov/) using the TCGAbiolinks R package [30]. RNA-seq expression profiles were generated using the Illumina Hiseq platform. To ensure accurate quantification, we applied the trimmed mean of M values (TMM) method through the edgeR R package [38] to normalize the raw RNA-seq read counts. Subsequently, the normalized counts were fitted using a negative binomial generalized log-linear model.

2.5. Gene functions and pathway analysis

To predict biological functions and gene regulatory networks associated with overlapping DMPs between tissue and blood of LUAD patients, a list of genes (6361 genes mapping to the 17,251 DMPs that showed significant correlations with gene expression) was submitted to pathfinder [39] for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and Gene Ontology (GO).

2.6. Statistical analyses

Statistical analyses were performed using GraphPad Prism (version 9.2.0) and R Studio (version April 1, 1103). The Mann-Whitney *U* test was used to compare the differences in DNA methylation by CpG density between the LUAD and non-LUAD. The diagnostic value of methylation levels at signature CpG sites for discriminating LUAD samples from non-LUAD samples was assessed using the area under the ROC curve (AUC), which was computed by plotROC. This evaluation was conducted on combined datasets from GEO (GSE85845 and GSE66836), TCGA, Thai blood samples, Thai and Han Chinese samples, as well as validation datasets from GEO (GSE56044 and GSE75008). The proportions of samples were reported with 95 % confidence intervals, calculated based on binomial distributions. Data are presented as mean ± standard deviation (SD), and $p \leq 0.05$ were considered to be significant.

3. Results

In our study, a total of 644 tumors and 59 normal tissues from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases were examined and used to inform the analysis of whole blood samples from 56 cancer patients and 79 healthy individuals. The clinical characteristics of the study groups are summarized in Table 1, and the study design is summarized in Fig. 1.

3.1. Identification of DMPs in lung adenocarcinoma tissue

We first identified the differentially methylated positions (DMPs) in LUAD versus normal tissue samples. This revealed 83,977 DMPs in the combined GEO dataset (Fig. 2A) and 160,337 DMPs in the TCGA dataset (Fig. 2B). Of the 59,639 DMPs showing significant changes in both datasets, 28,499 (48 %) were hypomethylated and 31,140 (52 %) exhibited hypermethylation (Fig. 2C). Gains of methylation were most commonly observed at CpG-rich regions (CpG islands and shores), while losses were observed at regions of lower CpG density (open seas and shelves) (Fig. 2D). The leading significant DMPs are presented in Table 2. Amongst these, we identified a set of five CpG sites (cg09887059, cg16768018, cg26521404, cg12595013, and cg08328160) mapping to the *ZIC4*, *HOXA9*, *ZIC1*, and *AGER* genes that were strongly hypermethylated in both the TCGA and combined GEO datasets.

3.2. Correlation of DMPs with gene expression changes in lung adenocarcinoma tissue

Among the 59,639 DMPs that intersected between the GEO and TCGA datasets, the correlation between the mean DNA methylation difference and log fold change of gene expression was classified into four groups: hypermethylated–upregulated (hyper–up); hypermethylated–downregulated (hyper–down); hypomethylated–upregulated (hypo–up); and hypomethylated–downregulated (hypo–down) genes (Fig. 2E). We identified 17,251 DMPs mapping to 6361 genes that showed significant correlations with gene expression. Of the 2830 hypermethylated genes, 2597 were upregulated and 233 genes were downregulated. Among the 3531 hypomethylated genes, 3419 genes were upregulated and 112 genes were downregulated.

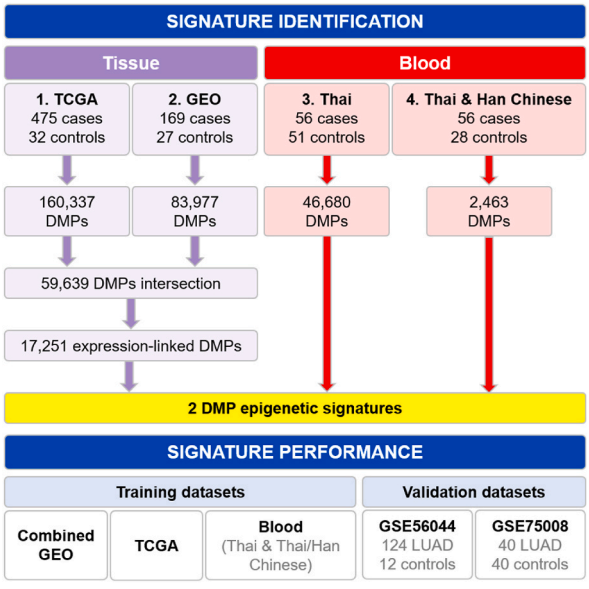

Fig. 1. Flow chart of study design.

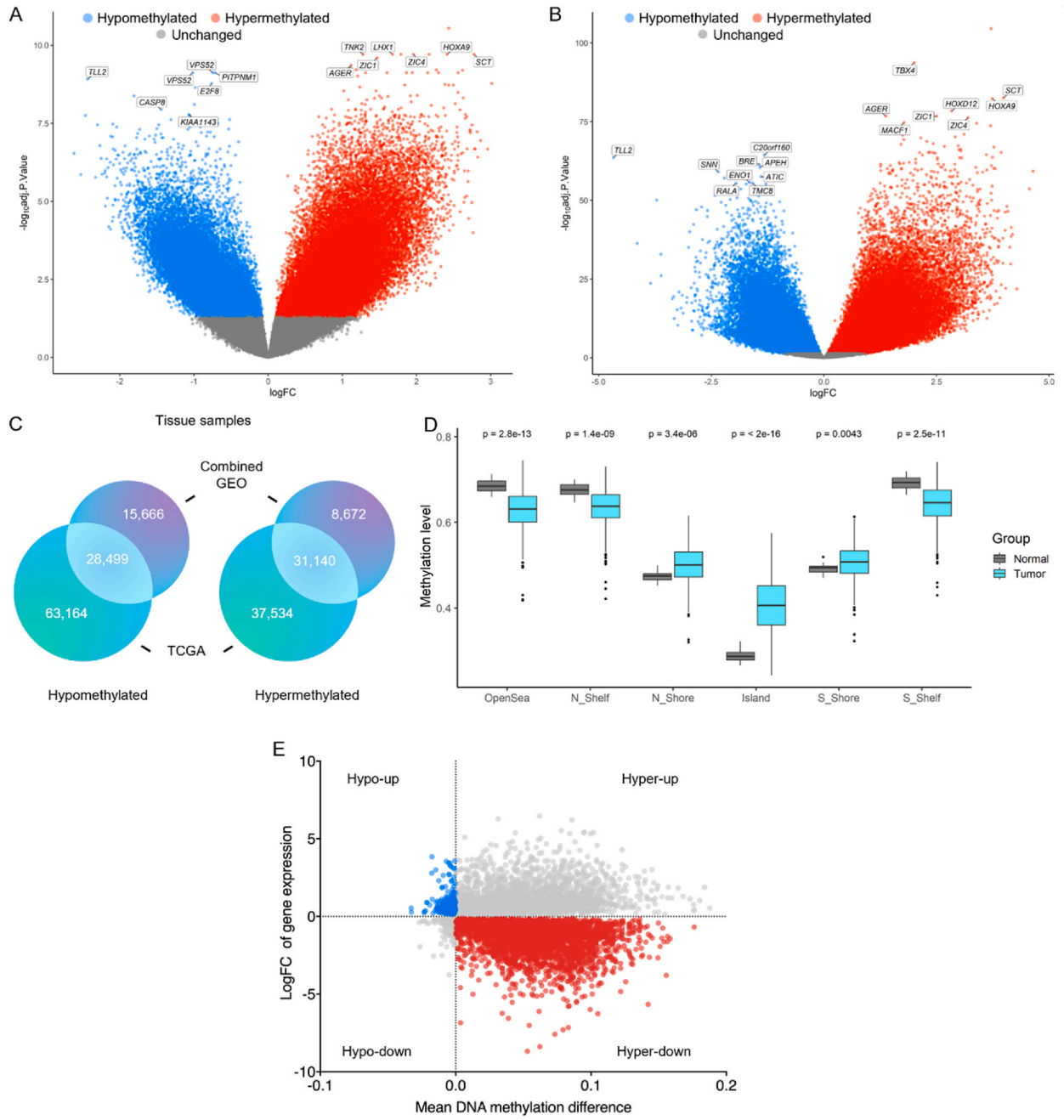

Fig. 2. DNA methylation (DNAm) profiling of tissue samples of lung adenocarcinoma. A, B, Volcano plots of the differentially methylated positions (DMPs) from the combined GEO dataset (A) and the TCGA dataset (B). **C,** Venn diagram depicting DMPs overlapping between the combined GEO and TCGA datasets by methylation change. **D,** The distribution of DMPs by CpG density within the combined GEO dataset. **E,** Scatter plot of mean methylation difference versus log fold change of gene expression in the TCGA dataset (blue: hypomethylated and upregulated; red: hypermethylated and down-regulated).

3.3. Epigenetic changes in the blood of LUAD patients

We next examined DNAm in blood samples from Thai patients and healthy elderly Thai cohort controls. This revealed a total of 46,680 DMPs. Among these, 27,073 (58 %) CpG sites exhibited hypomethylation, while 19,607 (42 %) CpG sites showed hypermethylation (Fig. 3A). Gains of methylation were again most commonly observed at regions of high CpG density (CpG islands), while losses were observed at regions of low CpG density (shelves and shores). GO analysis of the 46,680 DMPs revealed a significant enrichment of genes involved in the structural constituents of ribosomes and protein serine/threonine kinase activity (Fig. 3B). KEGG

Table 2
Top 10 significant CpG sites by dataset.

CpG	logFC	Adjusted *P* value	Chromosome	CpG density	CpG location	Gene
GEO						
cg08566455	2.44	2.85E-11	2	Island		
cg00249511	2.77	1.97E-10	11	Island	5′UTR	*SCT*
cg04864807	1.79	1.97E-10	2	Island		
cg09628877	1.28	1.97E-10	3	Island	Gene body	*TNK2*
cg09887059	2.17	1.97E-10	12	Island		
cg14754787	1.68	1.97E-10	17	Island	Gene body	*LHX1*
cg16768018	1.96	1.97E-10	3	Island	Gene body	*ZIC4*
cg26521404	2.41	1.97E-10	7	Island	1stExon	*HOXA9*
cg12595013	1.47	2.56E-10	3	Island	1stExon	*ZIC1*
cg08328160	1.12	4.42E-10	6	Open sea	Gene Body	*AGER*
TCGA						
cg08566455	3.72	2.88E-105	2	Island		
cg14823851	2.01	2.05E-94	17	Island	TSS1500	*TBX4*
cg25774643	3.99	2.46E-83	11	Island	TSS200	*SCT*
cg26521404	3.74	4.40E-83	7	Island	1stExon	*HOXA9*
cg03964958	2.84	4.01E-79	2	Island	1stExon	*HOXD12*
cg12595013	2.51	1.74E-77	3	Island	1stExon	*ZIC1*
cg08328160	1.38	1.74E-77	6	Open sea	Gene Body	*AGER*
cg16768018	3.20	5.74E-77	3	Island	Gene Body	*ZIC4*
cg08943107	1.78	2.15E-75	1	Open sea	Gene body	*MACF1*
cg09887059	3.39	3.12E-75	12	Island		
Whole Blood (Thai)						
cg01957330	−0.68	5.73E-45	7	Island	TSS200	*SKAP2*
cg08604594	−0.62	5.73E-45	10	Open sea		
cg08786079	0.75	5.73E-45	17	Shelf	3′UTR	*BCL6B*
cg09830059	0.74	5.73E-45	3	Shelf	Gene body	*DNAJC19*
cg10124255	0.55	5.73E-45	7	Open sea		
cg25484319	−0.56	5.73E-45	14	Open sea	Gene body	*RDH11*
cg26306476	−0.61	5.73E-45	19	Shelf	Gene body	PRR22
cg10985479	−0.65	7.07E-45	4	Shore	TSS1500	*LEF1*
cg27626141	0.55	7.07E-45	8	Island	TSS200	*AZIN1*
cg02159896	−0.60	1.40E-44	18	Shore	5′UTR	*MPPE1*

LogFC = log fold change.

analysis demonstrated that the significant enrichment of pathways was related to cancer such as the ribosome, spliceosome, cell cycle, ubiquitin-mediated proteolysis and nucleocytoplasmic transport (Fig. 3C). The leading hits in blood were distinct from those identified through analysis of tissue samples (Table 2).

3.4. Signature DMPs within tumor tissue and blood samples from LUAD patients

To develop a common DMP signature between lung tissue and blood samples of LUAD patients, we conducted a comparative analysis. Furthermore, to enhance the robustness of our findings, we also included an analysis of DNAm from Thai patients *vs* an independent external cohort of blood samples taken from healthy Han Chinese individuals (Supp Fig 1). We restricted our DMP list to those that a) were identified in both blood DNAm dataset analyses and b) show concordant changes in the tissue and blood of LUAD patients (i.e. hypomethylation or hypermethylation in both tissues). This revealed a two DMP signature (Table 3) matching between lung tumors and blood samples from cancer patients: one locus mapping to the *PTGER2* gene that is hypermethylated in cancer cases (cg00717218; Fig. 3D); and another mapping to *STX8* that is hypomethylated (cg15837531; Fig. 3E).

3.5. Epigenetic signature performance in classifying LUAD patient and control blood samples

In order to assess the diagnostic accuracy of the two signature DMPs in LUAD, we performed receiver operating characteristic (ROC) curve analysis on both lung tissue and blood samples as testing datasets. The diagnostic performance of the signature in distinguishing LUAD samples from normal lung tissue was high. The AUC for the combined GEO dataset was 0.77 (95%CI: 0.67–0.88), and for the TCGA dataset it was 0.92 (95%CI: 0.89–0.95). Diagnostic performance in the two blood datasets (Thai participants and Han Chinese) were similarly high, with an AUC of 0.96 (95%CI: 0.91–1.00) and 0.92 (95%CI: 0.85–0.98), respectively (Fig. 4A). We then assessed the performance of the DMP signature in independent validation cohorts, using two 450K datasets: GSE56044 (N = 136) and GSE75008 (N = 80). The performance of the signature was similarly high, with AUC of 0.92 (95%CI: 0.88–0.97) and 0.90 (95%CI: 0.83–0.97) in the GSE56044 and GSE75008 datasets, respectively (Fig. 4B).

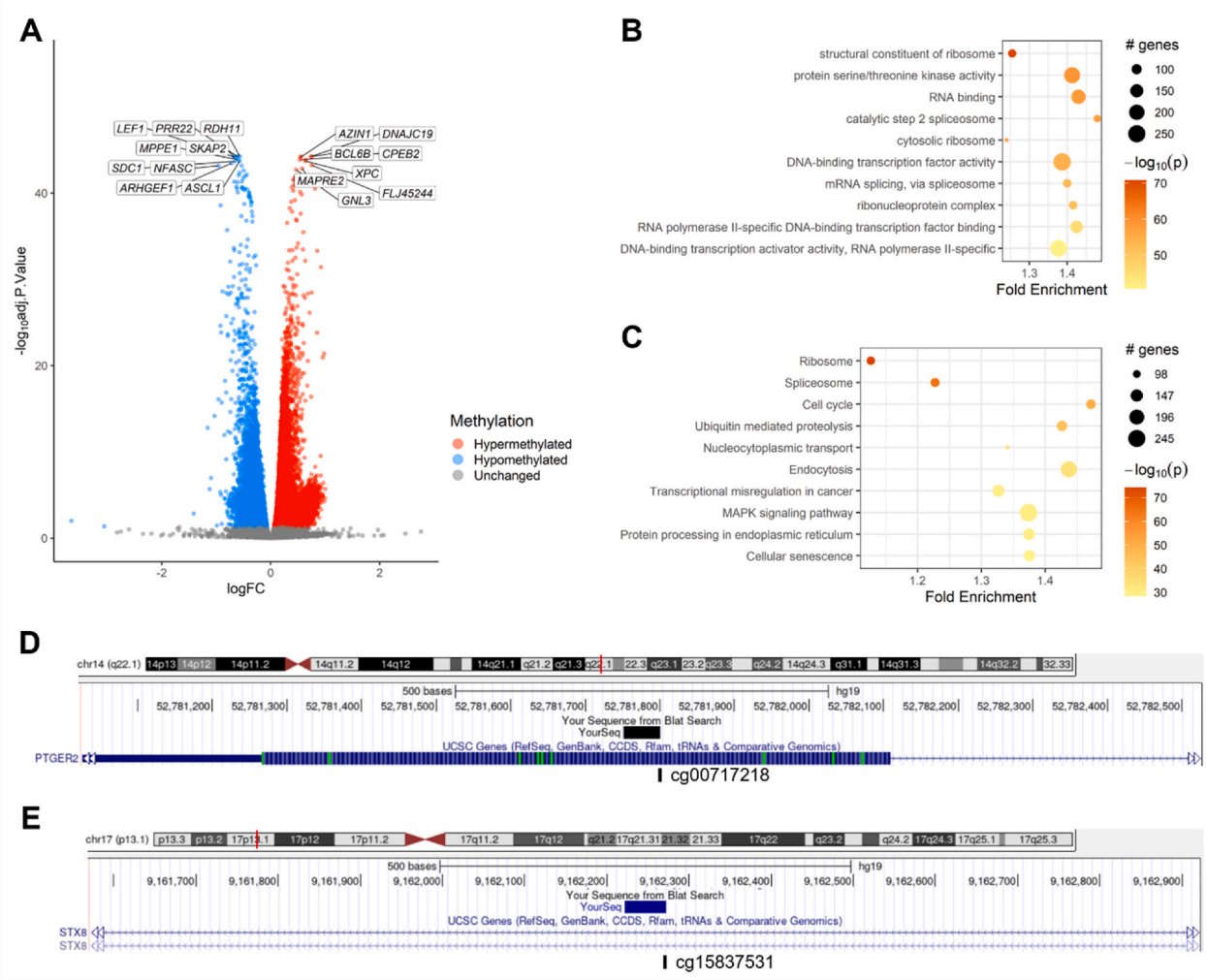

Fig. 3. DNAm profiling of blood samples from lung adenocarcinoma patients. A, the differentially methylated positions (DMPs). **B, C,** Pathway and gene oncology (GO) analysis of the 46,680 DMPs (B) and Kyoto Encyclopedia of Genes and Genomes pathway analysis (C). **D**, **E**, CpG sites mapping of cg00717218 (D) and cg15837531 (E).

Table 3
Two DMP signatures for tissue and blood.

CpG	Methylation	Adjusted *P* value				chr	CpG density	CpG location	Gene
		Combined GEO (GSE85845 and GSE66836)	TCGA	Whole Blood (Thai)	Whole Blood (Thai and Han Chinese)				
cg00717218	Hypermethylated	1.62E-05	1.23E-32	0.02	1.31E-02	14	Island	1stExon	*PTGER2*
cg15837531	Hypomethylated	4.27E-02	1.55E-04	1.65E-13	1.82E-04	17	Open sea	Gene body	*STX8*

chr = chromosome.

4. Discussion

The early detection of lung tumors is an urgent clinical need. Most patients are diagnosed with advanced disease that has already metastasized. Non-invasive testing, such as through the use of blood or sputum samples, offers a means to screen populations and to identify tumors while still localized, thereby offering improved patient outcomes. By taking a stringent approach, our study has been able to identify epigenetic changes present in tumors that are also measurable in blood and which offer high sensitivity and specificity

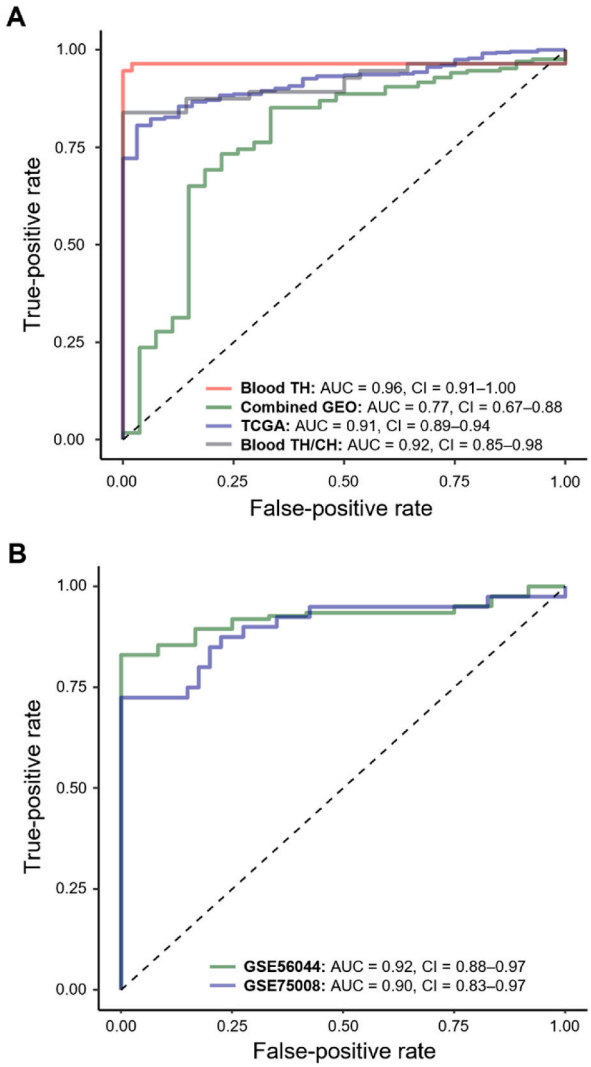

Fig. 4. Performance of the two DMP signatures assessed by ROC curve analysis. A, Training datasets of tissue samples (combined GEO, and TCGA) and two blood datasets (Thai and Thai/Han Chinese). **B,** Testing datasets (GSE75008 and GSE56044).

for the identification of the disease.

Our novel signature is comprised of two DMPs, mapping to the *PTGER2* and *STX8* genes. *PTGER2* is the more well-characterized of these, known to promote tumor development [40] and stimulate the growth of malignant cells [41], and it has been reported as hypermethylated in 58–70 % of NSCLC cases [42,43]. The *STX8* gene encodes syntaxin 8, a component of the SNARE complex involved in vesicular trafficking, which is normally expressed at low levels in lung tissue [44]. This gene demonstrates increased expression in NSCLC tissues [45], where it plays a pivotal role in regulating the DNA damage response and facilitating epithelial-to-mesenchymal transition [46]. These mechanisms contribute to the development of radiation resistance in lung cancer, emphasizing its clinical relevance and potential therapeutic implications [46].

Our approach has enabled us to account for epigenetic changes in the blood of patients that may not be specific to the disease. The presence of solid tumors within the body is associated with changes leukocyte composition in the blood that can be inferred through epigenetic analysis [23]. Indeed, many early studies examining blood samples from cancer patients reported differential methylation of genes related to immune function [47], including for lung cancer [48]. Such changes are unlikely to be specific to malignancies and able to serve as robust clinical biomarkers. Furthermore, many risk factors for cancer are associated with epigenetic changes in the blood. Of particular relevance, hypomethylation of *AHRR* and *F2RL3* has been widely reported in response to smoking [24]. While these loci may predict risk of lung cancer [25,49,50] and cancer-associated mortality [51], they are not suitable for diagnostic purposes. The impact of these risk factors is demonstrated by a recent blood-based case-control study in which 23 of 25 the identified DMPs were associated with smoking [52]. The DMPs identified through our study are more likely to be directly relevant to tumor biology. The benefit, and necessity, of this tumor-informed analysis of patient blood samples is demonstrated by fewer than 2 % of the DMPs observed in blood also showing corresponding changes within malignant lung tissue. Notably, DMPs identified in blood samples

are associated with pathways implicated in lung cancer, including the cell cycle, MAPK signaling, and cellular senescence. In a detailed characterization of the tumor immune infiltrate in patients with LUAD, Lin et al. [53] observed that genes related to cellular senescence could significantly influence the composition and distribution of the tumor immune infiltrate. The unregulated activation of the MAPK pathway plays a significant role in the oncogenic transformation of mammalian cells [54]. Senescence is a complex biological process with both cell-autonomous and paracrine effects, exerting a significant impact on the microenvironment [55,56]. However, the fold enrichment observed in the KEGG analysis was modest, which could be attributed to the relatively small sample size used in this study and interpatient heterogeneity. To fully assess the biological relevance of these pathways, further validation in larger, more diverse cohorts is necessary. Expanding the sample size in future studies will help confirm the significance of these pathways and provide a more robust understanding of their functional role in carcinogenesis.

The epigenetic signature identified in our study compares well to other blood-based studies of lung adenocarcinomas. Methylation of *PTGER4*, *RASSF1A* and *SHOX2* has been the most extensively examined, with these markers offering an AUC of 0.94 in blood for the diagnosis of all lung cancer types, with 89 % sensitivity for adenocarcinomas [57]. A similar model using only *PTGER4* and *SHOX2* offered an AUC of 0.95 in plasma [58]. Other commercially-available marker panels such as Lung EpiCheck have been similarly effective [59]. Other proposed markers include *FUT7* [60] and *PNPLA2* [61]. Recently, a panel of seven genes demonstrated promising detection rates for lung cancer compared to healthy controls, with a sensitivity of 86.7 %, specificity of 81.4 %, and an AUC of 0.891 [62]. We also cross-evaluated the methylation probes identified in Moon's [63] and Fasanelli's [25] studies using validation datasets. Our signature loci demonstrated superior performance compared to the probes reported in those studies (Supp Fig. 2). However, further work will be required to evaluate its performance as a diagnostic test, such as assessing the sensitivity in detection of early disease.

Due to the absence of publicly-available and blood-based validation datasets, we have been unable to test our signature further to evaluate its relative performance. While the relatively low number of blood samples from patients is a further limitation of our study, it should be noted that the blood samples originated from individuals of Asian descent while the selection of signature loci was based on methylation and gene expression data from a tissue database representing a diverse ethnic mix. This may provide some confidence that the signature may hold true across diverse ethnic groups. It is also important to note that while these loci show large changes in methylation within tumors ($\Delta\beta > 0.2$), the differences measured in blood were significantly more subtle. Further work will be required to determine the sensitivity of the signature in cfDNA in comparison to whole blood samples.

Sputum samples may offer an alternative approach to non-invasive testing that avoids some of the issues inherent in the use of blood. *RASSF1A* and *SHOX2* methylation testing offers a sensitivity of 81 % and specificity of 97 % in bronchoalveolar lavage fluid, with an AUC of 0.89 [64], while the use of *SHOX2* as a single gene marker offers similar performance in bronchial aspirates (AUC = 0.94) [4] and blood (AUC = 0.91) [65]. Elsewhere, methylation-specific PCR-based testing of alternative three-gene panels has demonstrated superior performance in sputum in comparison to plasma [66]. In parallel to evaluating the two DMP signatures in blood samples from an independent patient cohort, it will be important to examine the signatures in alternative biofluids such as sputum or bronchial washings, to determine the most effective method for non-invasive screening of lung cancer. Since sputum and bronchial washings are more directly associated with the respiratory tract, they may provide a more sensitive and specific reflection of tumor-related methylation changes. Comparing the diagnostic performance of these sample types alongside blood-based biomarkers could lead to a more comprehensive understanding of the optimal non-invasive screening approach, potentially improving early detection and patient outcomes. Moreover, functional assays could further demonstrate clinical relevance by revealing the mechanistic function of the changes in carcinogenesis and disease progression.

5. Conclusions

Our two DMP signatures offer promising performance in both lung tissue and blood for the identification of LUAD. Furthermore, to harness the advantages of utilizing these signature loci, longitudinal studies on DNA methylation as a potentially crucial indicator of tumorigenesis will contribute to the identification of potential biomarkers for prognostic purposes.

Funding

This work was supported by the National Research Council of Thailand (N41A640123) to C.S. and Grant #CA209414 to D.C.

CRediT authorship contribution statement

Pitaksin Chitta: Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Timothy M. Barrow:** Writing – review & editing, Writing – original draft, Visualization. **Atchara Dawangpa:** Writing – review & editing, Writing – original draft, Investigation. **David C. Christiani:** Writing – review & editing, Writing – original draft. **Naravat Poungvarin:** Writing – review & editing, Writing – original draft. **Chanachai Sae-Lee:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Conceptualization.

Data availability

Tissue data are publicly available from The Cancer Genome Atlas (TCGA) via the URL: https://portal.gdc.cancer.gov/using the TCGAbiolinks R package and from Gene Expression Omnibus (GEO): GSE85845 and GSE66836. Validation data are also publicly

available from GEO: GSE56044 and GSE75008. Thai blood DNA methylation data has been deposited in Science Data Bank under the accession link https://doi.org/10.57760/sciencedb.15236.

Ethics approval

Ethical approval for this study (COA no. Si701/2021 and COA. no.Si213/2021) was granted by the Siriraj Institutional Review Board, Siriraj Hospital, Mahidol University, Bangkok, Thailand. All study protocols were performed according to the Declaration of Helsinki, and results were expressed by using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

C.S. acknowledges support from the Postdoctoral Fellowship Program, Siriraj Hospital, Mahidol University, from Academic development scholarship. The authors are indebted to the Siriraj Clinical Molecular Pathology Laboratory team. This work was supported by the National Research Council of Thailand (N41A640123) to C.S. and Grant #CA209414 to D.C.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2025.e42581.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 71 (3) (2021) 209–249.
[2] R.L. Siegel, K.D. Miller, N.S. Wagle, A. Jemal, Cancer statistics, 2023, CA Cancer J. Clin. 73 (1) (2023) 17–48.
[3] K.C. Thandra, A. Barsouk, K. Saginala, J.S. Aluru, A. Barsouk, Epidemiology of lung cancer, Contemp. Oncol. 25 (1) (2021) 45–52.
[4] D. Dietrich, C. Kneip, O. Raji, T. Liloglou, A. Seegebarth, T. Schlegel, et al., Performance evaluation of the DNA methylation biomarker SHOX2 for the aid in diagnosis of lung cancer based on the analysis of bronchial aspirates, Int. J. Oncol. 40 (3) (2012) 825–832.
[5] S.R. Payne, From discovery to the clinic: the novel DNA methylation biomarker (m)SEPT9 for the detection of colorectal cancer in blood, Epigenomics 2 (4) (2010) 575–585.
[6] D.E. Piccioni, A.S. Achrol, L.A. Kiedrowski, K.C. Banks, N. Boucher, G. Barkhoudarian, et al., Analysis of cell-free circulating tumor DNA in 419 patients with glioblastoma and other primary brain tumors, CNS Oncol 8 (2) (2019) CNS34.
[7] A.L. Fymat, Genetics, epigenetics and cancer, Cancer Ther Oncol Int J 4 (2017) 555634.
[8] Y. Cheng, C. He, M. Wang, X. Ma, F. Mo, S. Yang, et al., Targeting epigenetic regulators for cancer therapy: mechanisms and advances in clinical trials, Signal Transduct. Targeted Ther. 4 (2019) 62.
[9] C. Sae-Lee, S. Corsi, T.M. Barrow, G.G.C. Kuhnle, V. Bollati, J.C. Mathers, et al., Dietary intervention modifies DNA methylation age assessed by the epigenetic clock, Mol. Nutr. Food Res. 62 (23) (2018) e1800092.
[10] L.B. Alexandrov, S. Nik-Zainal, D.C. Wedge, S.A. Aparicio, S. Behjati, A.V. Biankin, et al., Signatures of mutational processes in human cancer, Nature 500 (7463) (2013) 415–421.
[11] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, et al., Initial sequencing and analysis of the human genome, Nature 409 (6822) (2001) 860–921.
[12] K. Lokk, T. Vooder, R. Kolde, K. Valk, U. Vosa, R. Roosipuu, et al., Methylation markers of early-stage non-small cell lung cancer, PLoS One 7 (6) (2012) e39813.
[13] K. Quek, J. Li, M. Estecio, J. Zhang, J. Fujimoto, E. Roarty, et al., DNA methylation intratumor heterogeneity in localized lung adenocarcinomas, Oncotarget 8 (13) (2017) 21994–22002.
[14] C.J. Marsit, D.H. Kim, M. Liu, P.W. Hinds, J.K. Wiencke, H.H. Nelson, et al., Hypermethylation of RASSF1A and BLU tumor suppressor genes in non-small cell lung cancer: implications for tobacco smoking during adolescence, Int. J. Cancer 114 (2) (2005) 219–223.
[15] S.A. Belinsky, K.J. Nikula, W.A. Palmisano, R. Michels, G. Saccomanno, E. Gabrielson, et al., Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis, Proc. Natl. Acad. Sci. U. S. A. 95 (20) (1998) 11891–11896.
[16] H.H. Nelson, C.J. Marsit, B.C. Christensen, E.A. Houseman, M. Kontic, J.L. Wiemels, et al., Key epigenetic changes associated with lung cancer development: results from dense methylation array profiling, Epigenetics 7 (6) (2012) 559–566.
[17] A. Karlsson, M. Jonsson, M. Lauss, H. Brunnstrom, P. Jonsson, A. Borg, et al., Genome-wide DNA methylation analysis of lung carcinoma reveals one neuroendocrine and four adenocarcinoma epitypes associated with patient outcome, Clin. Cancer Res. 20 (23) (2014) 6127–6140.
[18] M.M. Bjaanaes, T. Fleischer, A.R. Halvorsen, A. Daunay, F. Busato, S. Solberg, et al., Genome-wide DNA methylation analyses in lung adenocarcinomas: association with EGFR, KRAS and TP53 mutation status, gene expression and prognosis, Mol. Oncol. 10 (2) (2016) 330–343.
[19] J. Krushkal, T. Silvers, W.C. Reinhold, D. Sonkin, S. Vural, J. Connelly, et al., Epigenome-wide DNA methylation analysis of small cell lung cancer cell lines suggests potential chemotherapy targets, Clin. Epigenet. 12 (1) (2020) 93.
[20] A. Moran, T. Fernandez-Marcelo, J. Carro, C. De Juan, I. Pascua, J. Head, et al., Methylation profiling in non-small cell lung cancer: clinical implications, Int. J. Oncol. 40 (3) (2012) 739–746.
[21] A.E. Teschendorff, A. Jones, H. Fiegl, A. Sargent, J.J. Zhuang, H.C. Kitchener, et al., Epigenetic variability in cells of normal cytology is associated with the risk of future morphological transformation, Genome Med. 4 (3) (2012) 24.
[22] Y.X. Shi, Y. Wang, X. Li, W. Zhang, H.H. Zhou, J.Y. Yin, et al., Genome-wide DNA methylation profiling reveals novel epigenetic signatures in squamous cell lung cancer, BMC Genom. 18 (1) (2017) 901.

[23] D.C. Koestler, C.J. Marsit, B.C. Christensen, W. Accomando, S.M. Langevin, E.A. Houseman, et al., Peripheral blood immune cell methylation profiles are associated with nonhematopoietic cancers, Cancer Epidemiol. Biomarkers Prev. 21 (8) (2012) 1293–1302.
[24] N.S. Shenker, S. Polidoro, K. van Veldhoven, C. Sacerdote, F. Ricceri, M.A. Birrell, et al., Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking, Hum. Mol. Genet. 22 (5) (2013) 843–851.
[25] F. Fasanelli, L. Baglietto, E. Ponzi, F. Guida, G. Campanella, M. Johansson, et al., Hypomethylation of smoking-related genes is associated with future lung cancer in four prospective cohorts, Nat. Commun. 6 (2015) 10192.
[26] H. Yan, Q. Guan, J. He, Y. Lin, J. Zhang, H. Li, et al., Individualized analysis reveals CpG sites with methylation aberrations in almost all lung adenocarcinoma tissues, J. Transl. Med. 15 (1) (2017) 26.
[27] M.M. Bjaanæs, T. Fleischer, A.R. Halvorsen, A. Daunay, F. Busato, S. Solberg, et al., Genome-wide DNA methylation analyses in lung adenocarcinomas: association with EGFR, KRAS and TP53 mutation status, gene expression and prognosis, Mol. Oncol. 10 (2) (2016) 330–343.
[28] C. Xiao, S. Yi, D. Huang, Genome-wide identification of age-related CpG sites for age estimation from blood DNA of Han Chinese individuals, Electrophoresis 42 (14–15) (2021) 1488–1496.
[29] T. Goldmann, B. Schmitt, J. Muller, M. Kroger, S. Scheufele, S. Marwitz, et al., DNA methylation profiles of bronchoscopic biopsies for the diagnosis of lung cancer, Clin. Epigenet. 13 (1) (2021) 38.
[30] A. Colaprico, T.C. Silva, C. Olsen, L. Garofano, C. Cava, D. Garolini, et al., TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data, Nucleic Acids Res. 44 (8) (2016) e71.
[31] S. Davis, P.S. Meltzer, GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor, Bioinformatics 23 (14) (2007) 1846–1847.
[32] F. Faul, E. Erdfelder, A.-G. Lang, A. Buchner, G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, Behav. Res. Methods 39 (2) (2007) 175–191.
[33] J.P. Fortin, T.J. Triche Jr., Hansen KD. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi, Bioinformatics 33 (4) (2017) 558–560.
[34] R. Pidsley, E. Zotenko, T.J. Peters, M.G. Lawrence, G.P. Risbridger, P. Molloy, et al., Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling, Genome Biol. 17 (1) (2016) 208.
[35] J.T. Leek, W.E. Johnson, H.S. Parker, A.E. Jaffe, J.D. Storey, The sva package for removing batch effects and other unwanted variation in high-throughput experiments, Bioinformatics (Oxford, England) 28 (6) (2012) 882–883.
[36] E.A. Houseman, W.P. Accomando, D.C. Koestler, B.C. Christensen, C.J. Marsit, H.H. Nelson, et al., DNA methylation arrays as surrogate measures of cell mixture distribution, BMC Bioinf. 13 (2012) 86.
[37] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47.
[38] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics 26 (1) (2010) 139–140.
[39] E. Ulgen, O. Ozisik, O.U. Sezerman, pathfindR: an R package for comprehensive identification of enriched pathways in omics data through active subnetworks, Front. Genet. 10 (2019) 858.
[40] R.L. Keith, M.W. Geraci, S.P. Nana-Sinkam, R.M. Breyer, T.M. Hudish, A.M. Meyer, et al., Prostaglandin E2 receptor subtype 2 (EP2) null mice are protected against murine lung tumorigenesis, Anticancer Res. 26 (4B) (2006) 2857–2861.
[41] T. Yamaki, K. Endoh, M. Miyahara, I. Nagamine, N. Thi Thu Huong, H. Sakurai, et al., Prostaglandin E2 activates Src signaling in lung adenocarcinoma cell via EP3, Cancer Lett. 214 (1) (2004) 115–120.
[42] L. Tian, M. Suzuki, T. Nakajima, R. Kubo, Y. Sekine, K. Shibuya, et al., Clinical significance of aberrant methylation of prostaglandin E receptor 2 (PTGER2) in nonsmall cell lung cancer: association with prognosis, PTGER2 expression, and epidermal growth factor receptor mutation, Cancer 113 (6) (2008) 1396–1403.
[43] M. Yoshino, M. Suzuki, L. Tian, Y. Moriya, H. Hoshino, T. Okamoto, et al., Promoter hypermethylation of the p16 and Wif-1 genes as an independent prognostic marker in stage IA non-small cell lung cancers, Int. J. Oncol. 35 (5) (2009) 1201–1209.
[44] V. Thoreau, T. Berges, I. Callebaut, Z. Guillier-Gencik, L. Gressin, A. Bernheim, et al., Molecular cloning, expression analysis, and chromosomal localization of human syntaxin 8 (STX8), Biochem. Biophys. Res. Commun. 257 (2) (1999) 577–583.
[45] J. Sun, T. Tian, L. Chen, Y. Liu, J. Kong, Research of the expression and significance of syntaxin 8 in non-small cell lung cancer tissues, Tianjin Med. J. (2017) 735–738.
[46] C. Thumser-Henner, S. Oeck, S. Kalmbach, J. Forster, F. Kindl, A. Sak, et al., Syntaxin 18 regulates the DNA damage response and epithelial-to-mesenchymal transition to promote radiation resistance of lung cancer, Cell Death Dis. 13 (6) (2022) 529.
[47] A.E. Teschendorff, U. Menon, A. Gentry-Maharaj, S.J. Ramus, S.A. Gayther, S. Apostolidou, et al., An epigenetic signature in peripheral blood predicts active ovarian cancer, PLoS One 4 (12) (2009) e8274.
[48] L. Wang, J.A. Aakre, R. Jiang, R.S. Marks, Y. Wu, J. Chen, et al., Methylation markers for small cell lung cancer in peripheral blood leukocyte DNA, J. Thorac. Oncol. 5 (6) (2010) 778–785.
[49] L. Baglietto, E. Ponzi, P. Haycock, A. Hodge, M. Bianca Assumma, C.H. Jung, et al., DNA methylation changes measured in pre-diagnostic peripheral blood samples are associated with smoking and lung cancer risk, Int. J. Cancer 140 (1) (2017) 50–61.
[50] Y. Zhang, M. Elgizouli, B. Schottker, B. Holleczek, A. Nieters, H. Brenner, Smoking-associated DNA methylation markers predict lung cancer incidence, Clin. Epigenet. 8 (2016) 127.
[51] Y. Tsuboi, H. Yamada, E. Munetsuna, R. Fujii, M. Yamazaki, Y. Ando, et al., Increased risk of cancer mortality by smoking-induced aryl hydrocarbon receptor repressor DNA hypomethylation in Japanese population: a long-term cohort study, Cancer Epidemiol 78 (2022) 102162.
[52] T.M. Sandanger, T.H. Nost, F. Guida, C. Rylander, G. Campanella, D.C. Muller, et al., DNA methylation and associated gene expression in blood prior to lung cancer diagnosis in the Norwegian Women and Cancer cohort, Sci. Rep. 8 (1) (2018) 16714.
[53] W. Lin, X. Wang, Z. Wang, F. Shao, Y. Yang, Z. Cao, et al., Comprehensive analysis uncovers prognostic and immunogenic characteristics of cellular senescence for lung adenocarcinoma, Front. Cell Dev. Biol. 9 (2021) 780461.
[54] S.J. Mansour, W.T. Matten, A.S. Hermann, J.M. Candia, S. Rong, K. Fukasawa, et al., Transformation of mammalian cells by constitutively active MAP kinase kinase, Science 265 (5174) (1994) 966–970.
[55] A. Lasry, Y. Ben-Neriah, Senescence-associated inflammatory responses: aging and cancer perspectives, Trends Immunol. 36 (4) (2015) 217–228.
[56] A. Hernandez-Segura, J. Nehme, M. Demaria, Hallmarks of cellular senescence, Trends Cell Biol. 28 (6) (2018) 436–453.
[57] B. Wei, F. Wu, W. Xing, H. Sun, C. Yan, C. Zhao, et al., A panel of DNA methylation biomarkers for detection and improving diagnostic efficiency of lung cancer, Sci. Rep. 11 (1) (2021) 16782.
[58] G. Weiss, A. Schlegel, D. Kottwitz, T. Konig, R. Tetzner, Validation of the SHOX2/PTGER4 DNA methylation marker panel for plasma-based discrimination between patients with malignant and nonmalignant lung disease, J. Thorac. Oncol. 12 (1) (2017) 77–84.
[59] M. Gaga, J. Chorostowska-Wynimko, I. Horvath, M.C. Tammemagi, D. Shitrit, V.H. Eisenberg, et al., Validation of Lung EpiCheck, a novel methylation-based blood assay, for the detection of lung cancer in European and Chinese high-risk individuals, Eur. Respir. J. 57 (1) (2021).
[60] Y. Fang, Y. Qu, L. Ji, H. Sun, J. Li, Y. Zhao, et al., Novel blood-based FUT7 DNA methylation is associated with lung cancer: especially for lung squamous cell carcinoma, Clin. Epigenet. 14 (1) (2022) 167.
[61] R. Qiao, M. Li, R. Zhong, Y. Wei, J. Wang, Z. Zhang, et al., The association between PNPLA2 methylation in peripheral blood and early-stage lung cancer in a case-control study, Cancer Manag. Res. 13 (2021) 7919–7927.
[62] C. Du, L. Tan, X. Xiao, B. Xin, H. Xiong, Y. Zhang, et al., Detection of the DNA methylation of seven genes contribute to the early diagnosis of lung cancer, J. Cancer Res. Clin. Oncol. 150 (2) (2024) 77.
[63] D.H. Moon, S.O. Kwon, W.J. Kim, Y. Hong, Identification of serial DNA methylation changes in the blood samples of patients with lung cancer, Tuberc. Respir. Dis. 82 (2) (2019) 126–132.

[64] C. Zhang, W. Yu, L. Wang, M. Zhao, Q. Guo, S. Lv, et al., DNA methylation analysis of the SHOX2 and RASSF1A panel in bronchoalveolar lavage fluid for lung cancer diagnosis, J. Cancer 8 (17) (2017) 3585–3591.
[65] C. Kneip, B. Schmidt, A. Seegebarth, S. Weickmann, M. Fleischhacker, V. Liebenberg, et al., SHOX2 DNA methylation is a biomarker for the diagnosis of lung cancer in plasma, J. Thorac. Oncol. 6 (10) (2011) 1632–1638.
[66] A. Hulbert, I. Jusue-Torres, A. Stark, C. Chen, K. Rodgers, B. Lee, et al., Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum, Clin. Cancer Res. 23 (8) (2017) 1998–2005.