



HHS Public Access

Author manuscript

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2019 December 01.

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2018 December ; 27(12): 1527–1535. doi:
10.1158/1055-9965.EPI-18-0454.

DNA Methylation of *LRRC3B*: A Biomarker for Survival of Early-Stage Non-Small Cell Lung Cancer Patients

Yichen Guo^{1,2}, Ruyang Zhang³, Sipeng Shen^{1,3}, Yongyue Wei³, Sebastian Moran Salama⁴, Thomas Fleischer⁵, Maria Moksnes Bjaanæs⁵, Anna Karlsson⁶, Maria Planck⁶, Li Su¹, Zhaozhong Zhu¹, Johan Staaf⁶, Åslaug Helland^{5,7}, Manel Esteller⁴, and David C. Christiani^{1,8,*}

¹Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, USA

²Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, USA

³Department of Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 211166, China

⁴Bellvitge Biomedical Research Institute and University of Barcelona and Institutio Catalana de Recerca i Estudis Avançats, Barcelona, Spain

⁵Department of Genetics, Institute for Cancer Research, Oslo University Hospital - The Norwegian Radium Hospital, Oslo, Norway

⁶Division of Oncology and Pathology, Department of Clinical Sciences Lund, Lund University, Lund, Sweden

⁷Institute of Clinical Medicine, University of Oslo, Norway

⁸Pulmonary and Critical Care Division, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, USA

Abstract

Background: Previous studies support a tumor-suppressor role for *LRRC3B* across various types of cancers. We aim to investigate the association between DNA methylation of *LRRC3B* and overall survival for early-stage non-small cell lung cancer patients.

Methods: This study included 1,230 early-stage non-small cell lung cancer patients. DNA was extracted from lung tumor tissues and DNA methylation was measured using Illumina Infinium HumanMethylation450 BeadChips. The association between DNA methylation and overall survival was first tested using Cox regression on a discovery cohort and then validated in an independent cohort. Next, the association between DNA methylation and gene expression was

*Send correspondence to Dr. David C. Christiani, Building I Room 1401, 665 Huntington Avenue, Boston, MA 02115, USA. dchris@hsph.harvard.edu.

Conflicts of interest statements/Financial Disclosure statement: None.

Ethical Adherence

The study was approved by institutional review boards at Harvard School of Public Health, Massachusetts General Hospital, Bellvitge Biomedical Research Institute, Oslo University Hospital, Lund University and Skåne University Hospital.

investigated in two independent cohorts. Finally, the association between gene expression and overall survival was investigated in three independent groups of patients.

Results: Three novel DNA methylation sites in *LRRC3B* were significantly associated with overall survival in two groups of patients. Patients with hypermethylation in the DNA methylation sites had significantly longer survival than the others in both the discovery cohort (HR = 0.62; $P = 2.02 \times 10^{-05}$) and validation cohort (HR = 0.55; $P = 4.44 \times 10^{-04}$). The three DNA methylation sites were significantly associated with *LRRC3B* expression, which was also associated with overall survival.

Conclusion: Using clinical data from a large population, we illustrated the association between DNA methylation of *LRRC3B* and overall survival of early-stage non-small cell lung cancer.

Impact: We provide evidence of plausibility for building biomarkers on DNA methylation of *LRRC3B* for overall survival of early-stage non-small cell lung cancer, thus filling a gap between previous in vitro studies and clinical applications.

Keywords

LRRC3B; biomarker; non-small cell lung cancer; DNA methylation

Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide, with an estimated 224,390 new cases and 158,080 deaths in the U.S. alone in 2016 (1). Non-small cell lung cancer (NSCLC) comprises ~80% of lung cancers diagnosed in the U.S. (2,3). Clinically relevant prognostic information for lung cancer is mainly obtained by staging. However, patients within a stage have differing survival outcomes, suggesting there are other factors affecting prognosis (4). Hundreds of studies have been published on discovering prognostic factors beyond staging, such as molecular prognostic markers, and these have provided some biological and clinical insights (4–6). For early-stage (stage I and II) NSCLC, such markers might help further identify patients with poor prognosis, who can be targeted for aggressive adjuvant treatments to improve survival outcomes.

Epigenetic alterations, such as DNA methylation, are regarded as innovative cancer biomarkers because of their stability, frequency, and reversibility (7–11). In recent years, several epigenetic markers of lung cancer have been identified by in vitro and animal studies (7–11). However, to demonstrate the clinical relevance of those potential markers, large multicenter studies based on clinical data remain necessary (7).

LRRC3B encodes leucine-rich repeat-containing 3B, which is an evolutionarily highly conserved leucine-rich repeat-containing protein (12). Aberrant DNA methylation of *LRRC3B* has been associated with several types of cancer. *LRRC3B* is a target of aberrant DNA methylation in gastric cancer and may function as a tumor suppressor (12). Further, *LRRC3B* methylation intensity is significantly higher in cancer tissues than in the corresponding non-cancerous tissues of colorectal cancer (13). Additionally, *LRRC3B* is methylated and/or deleted with high frequency in major epithelial cancers, including breast, cervical, lung, kidney, ovarian, colon, and prostate cancers (14). However, none of those

studies provided strong clinical evidence by showing that DNA methylation of *LRRC3B* could predict lung cancer survival.

In this study, we performed a two-stage multicenter association analysis between lung tumor DNA methylation in *LRRC3B* and overall survival of 613 early-stage (stage I and II) NSCLC patients from an independent Caucasian population, followed by additional independent replication based on another cohort of 617 early-stage lung cancer patients from The Cancer Genome Atlas (TCGA). We detected three specific DNA methylation probes in *LRRC3B* as biomarkers of overall survival and demonstrated clear associations between those methylation probes and *LRRC3B* expression.

Methods

Populations

We harmonized datasets from five international study centers. The discovery phase consisted of four cohorts: USA, Spain, Norway, and Sweden. The validation phase was based on a TCGA cohort.

Cohort 1: USA. 152 newly diagnosed early-stage NSCLC patients were recruited at Massachusetts General Hospital (MGH) from 1992 to 2005. During curative surgery, tumor specimens were collected with complete resection and snap-frozen. Tumor DNA was extracted from 5 µm-thick histopathologic sections. Amount (tumor cellularity >70%) and quality of tumor cells for each specimen were evaluated by an MGH pathologist. In addition, specimens were histologically classified using World Health Organization criteria. The study protocol was approved by the institutional review boards at the Harvard School of Public Health and MGH. All patients provided written informed consent.

Center 2: Spain. Descriptions of this study population were previously reported (15). In brief, tumors specimens were collected by surgical resections from 226 early-stage NSCLC patients recruited during 1991–2009. Tumor DNA was extracted from fresh-frozen tumor specimens and further checked for integrity and quantity. The study was approved by the Bellvitge Biomedical Research Institute institutional review board. All patients provided written informed consent.

Center 3: Norway. Descriptions of this study population were previously reported (16). In brief, tumors specimens were collected by surgical resection from 144 early-stage NSCLC patients recruited at Oslo University Hospital-Rikshospitalet during 2006–2011. Tumor tissues obtained during surgery were snap-frozen in liquid nitrogen and stored at –80 °C until DNA isolation. This study was approved by the institutional review board and regional ethics committee (S-05307). All patients received oral and written information and signed a written consent.

Center 4: Sweden. Descriptions of this study population were previously reported (17). In brief, specimens were collected by surgical resection of 104 tumors of early-stage NSCLC patients recruited at Skåne University Hospital during 2004–2008. The study was approved

by the regional ethical review board (registration no. 2004/762 and 2008/702). All patients provided written informed consent.

Center 5: TCGA. Data from 617 early-stage NSCLC patients were included from The Cancer Genome Atlas (TCGA). Overall survival times and covariates were included. HumanMethylation450 DNA methylation image data (IDAT files) were downloaded on October 1, 2015.

Gene Expression Omnibus.—We also collected gene expression data of 437 early-stage NSCLC patients from the Gene Expression Omnibus (GEO) database. Those patients had complete clinical information, including overall survival time, smoking status, clinical stage, and histology.

DNA methylation profiles

DNA methylation was assessed by Illumina Infinium HumanMethylation450 BeadChips (Illumina Inc., San Diego, CA, USA). GenomeStudio Methylation Module V1.8 (Illumina Inc.) was used to transform raw image data into beta values (continuous numbers ranging from 0 to 1). Probes meeting the following criteria were removed: a) detection $P > 0.05$ in more than 5% of samples; b) coefficient of variance (CV) $< 5\%$; c) potentially contains or extends SNPs (for Infinium I probes, SNPs at the site of single base extension; for Infinium II probes, SNPs at CpG site; for both SNPs, SNPs located within 10 bp from site of single base extension) with MAF > 0.05 in 1000 Genomes Project 20110521 release for European population (18); d) cross-reactive probes (18) or cross-hybridizing probes (19); or e) probes passing QC in only one center. Samples with $>5\%$ undetectable probes were excluded. Methylation signals were further processed for quantile normalization (20), type I and II probe design bias correction (21), and batch effects adjustment (22).

Gene expression profiles

In the USA cohort, a whole-genome DASL HT assay (Illumina Corp., San Diego, CA) was used to measure gene expression values from a subset of NSCLC patients. Expression of all genes was normalized using dChip software before analysis.

In the Norway cohort, a subset of lung adenocarcinoma samples had both methylation and mRNA expression data. Gene expression was assessed through microarrays from Agilent Technologies (SurePrint G3 Human GE, 8×60 K). Gene expression data were \log_2 -transformed and normalized by Genespring GX analysis software v.12.1 (Agilent Technologies).

In the Sweden cohort, gene expression data were available on 117 tumor samples, measured by Illumina Human HT-12 V4 microarrays, and 97 samples had both methylation and expression data. Gene expression data were quantile-normalized and mean-centered for each probe across all samples. Probe sets without signal intensity above the median of negative control intensity signals in at least 80% of samples were excluded from analysis.

In the TCGA cohort, gene expression was measured by RNA sequencing (RNA-seq). Data preprocessing was done by the TCGA workgroup. Raw counts were normalized by

expectation maximization (RSEM). Level-3 (gene level) gene quantification data were downloaded from the TCGA data portal and were further checked for quality. Expression of all genes was extracted and quantile-normalized before analysis.

Survival analysis of single DNA methylation probe

After data preprocessing, there are 24 DNA methylation probes in *LRRC3B*. We performed survival analysis by applying univariate Cox proportional hazard models to test the association between each DNA methylation probe and overall survival time (OS) of 613 patients in the discovery cohorts (USA, Spain, Norway, and Sweden). To minimize the influence of potential outliers, DNA methylation level of each probe was dichotomized by median value. The Cox proportional hazard model was:

$$h(t)_{ij} = h_0(t)_{ij} \exp(\beta_j \times \text{methylation}_{ij})$$

Where i = index of patients; j = index of probes, and

$$\text{methylation}_{ij} = \begin{cases} 1 & \text{high methylation level (above median)} \\ 0 & \text{low methylation level (below median)} \end{cases}, \text{ for the } i^{\text{th}} \text{ patient}$$

Three out of 24 DNA methylation probes in *LRRC3B* were significantly associated with OS after adjusting for multiple testing ($P < 10^{-3}$). Hazard ratio (HR, between high methylation and low methylation) and 95% confidence interval (CI) were estimated for each probe. The same analyses were performed on 617 patients in the validation cohort to evaluate validity of the results. We further performed sensitivity analyses in both discovery and validation cohorts by building a multivariate Cox model adjusting for covariates, including age, gender, smoking status, cancer stage, and histology. The detailed model was:

$$h(t)_{ij} = h_0(t)_{ij} \exp(\beta_j \times \text{methylation}_{ij} + \alpha_1 \times \text{age}_i + \alpha_2 \times \text{gender}_i + \alpha_3 \times \text{smoking}_i + \alpha_4 \times \text{stage}_i + \alpha_5 \times \text{histology}_i)$$

Where i = index of patients; j = index of probes; and methylation_{ij} was defined the same as in (1).

Associations between methylation groups and overall survival

The methylation level of each selected probe was dichotomized by median. Then, Kaplan-Meier curves were plotted to compare OS of patients with high methylation levels (above median) with those with low methylation levels (below median) separately in discovery and validation cohorts.

We separated patients into different groups based on methylation levels of three probes. The association between overall survival and methylation group was tested by univariate as well as multivariate Cox models, as shown below. Kaplan-Meier curves were separately plotted for different groups in the discovery cohort, validation cohort, and total population. Log-rank tests were performed to test the differences between the curves.

Univariate Cox model:

$$h(t)_i = h_0(t)_i \exp(\beta \times group_i)$$

Multivariate Cox model:

$$h(t)_i = h_0(t)_i \exp(\beta \times group_i + \alpha_1 \times age_i + \alpha_2 \times gender_i + \alpha_3 \times smoking_i + \alpha_4 \times stage_i + \alpha_5 \times histology_i)$$

Also, among 805 patients who had treatment data, we performed a sensitivity analysis to adjust for confounding effects of adjuvant treatment using the following model:

$$h(t)_i = h_0(t)_i \exp(\beta \times group_i + \alpha_1 \times age_i + \alpha_2 \times gender_i + \alpha_3 \times smoking_i + \alpha_4 \times stage_i + \alpha_5 \times histology_i + \alpha_t \times adjunct_treatment_i)$$

Where i = index of patients and

$$adjunct_treatment_i = \begin{cases} 1 & \text{with chemotherapy and/or radiation before surgery, for } i^{th} \text{ patient.} \\ 0 & \text{otherwise} \end{cases}$$

Methylation risk scores

We performed Cox regressions which include all three selected methylation sites in it. For each person, a risk score was calculated by multiplying methylation levels and the correspondent Cox regression coefficients. People with high risk scores (above median of discovery cohort) and low risk scores were assigned to two groups. Kaplan-Meier curves were separately plotted for those groups in the discovery cohort, validation cohort, and total population. Log-rank tests were performed to test the differences between the curves.

Analysis of association between gene expression and DNA methylation

In the discovery cohort, 216 early-stage NSCLC patients had complete gene expression data and DNA methylation data. Gene expression data in different centers were measured by different platforms. Data were combined after being \log_2 -transformed and standardized (centralized by mean and scaled by standard deviation) in each center. Linear regression showed that there was no significant residual batch effect by center. Associations between expression of *LRRC3B* and DNA methylation level of the three selected probes were tested by linear regression.

In the validation cohort, 427 early-stage NSCLC patients had complete mRNA sequencing data and clinical information. Of those, 344 also had complete DNA methylation data. Expression of *LRRC3B* was \log_2 -transformed and standardized before analysis. Associations between *LRRC3B* expression and DNA methylation level of the three selected probes were tested by linear regression.

Survival analysis of gene expression

In summary, 216 patients in the discovery cohort and 427 patients in the validation cohort had *LRRC3B* expression data and clinical information, including OS, smoking status, clinical stage, and histology. We also collected gene expression data for 437 early-stage NSCLC patients from the Gene Expression Omnibus (GEO) database. Those patients had complete clinical information, including OS, smoking status, clinical stage, and histology. We dichotomized *LRRC3B* expression by median in each cohort. Associations between expression and OS were tested using a multivariate Cox model, adjusted for age, gender, smoking status, clinical stage, and histology.

Analysis of differential methylation and expression across tissues

DNA methylation data of 69 pairs of tumor and adjacent non-tumor tissues and gene expression data of 53 pairs of tumor and adjacent non-tumor tissues were collected from TCGA. Student's t-tests were used to test the difference in methylation levels in 24 probes of *LRRC3B* between tumor and adjacent non-tumor tissues. Gene expression data were \log_2 -transformed and student's t-tests were then used to test the difference in *LRRC3B* expression between tumor and adjacent non-tumor tissues.

All statistical analyses were performed using R version 3.3.0 (The R Foundation).

Results

DNA methylation of *LRRC3B* is associated with early-stage NSCLC survival

The study design is shown in Fig. 1, and Supplementary Table S1 lists demographic and clinical pathology data for all populations. We first evaluated the association between each of 24 DNA methylation probes in *LRRC3B* and OS of 613 patients in the discovery cohort using a univariate Cox model. Three DNA methylation probes (cg13046257, cg17623116, and cg19600115) located at the 5' untranslated region (5' UTR) were significantly associated with OS after adjusting for multiple testing ($P < 10^{-3}$) (Supplementary Table S2). We then tested the validity of associations in 617 patients in the validation cohort; associations remained significant with consistent effect size (Supplementary Table S3).

Methylation levels for these three probes were consistently distributed in both discovery and validation cohort samples (Fig. 2A). In addition, associations remained significant after adjusting for covariates, including age, gender, smoking status, cancer stage, and histology in the discovery cohort (cg13046257: HR = 0.76, 95% CI: 0.60–0.95; cg17623116: HR = 0.73, 95% CI: 0.59–0.91; cg19600115: HR = 0.72, 95% CI: 0.57–0.90), as well as in the validation cohort (cg13046257: HR = 0.53, 95% CI: 0.38–0.76; cg17623116: HR = 0.61, 95% CI: 0.43–0.85; cg19600115: HR = 0.54, 95% CI: 0.38–0.76) (Fig. 2B and Supplementary Table S4).

DNA methylation of *LRRC3B* is associated with early-stage NSCLC overall survival

For each of three selected probes, we plotted Kaplan-Meier curves for patients with high methylation levels (above median) at that probe against those with low methylation levels in discovery and validation cohorts respectively. Compared with those with low methylation

levels, patients with high methylation levels had significantly longer OS in both discovery (cg13046257: HR = 0.66, 95% CI: 0.53–0.82; cg17623116: HR = 0.68, 95% CI: 0.55–0.85; cg19600115: HR = 0.66, 95% CI: 0.53–0.82) and validation cohorts (cg13046257: HR = 0.51, 95% CI: 0.36–0.72; cg17623116: HR = 0.62, 95% CI: 0.45–0.87; cg19600115: HR = 0.57, 95% CI: 0.41–0.80) for all three probes (Fig.3 A-C).

By considering methylation profiles of all three probes together, we separated patients into different groups. Patients with at least two hypermethylated sites had significantly longer OS than the others in the discovery cohort (HR = 0.62; 95% CI: 0.50–0.77; $P = 2.02 \times 10^{-05}$), validation cohort (HR = 0.55; 95% CI: 0.39–0.77; $P = 4.44 \times 10^{-04}$), and total population (HR = 0.61; 95% CI: 0.52–0.74; $P = 2.80 \times 10^{-07}$) (Fig.4 A-C and Supplementary Table S5). Patients with three hypermethylated sites also had significantly longer OS than the others in the discovery cohort (HR = 0.57; 95% CI: 0.43–0.75; $P = 5.80 \times 10^{-05}$), validation cohort (HR = 0.45; 95% CI: 0.29–0.68; $P = 1.74 \times 10^{-04}$), and total population (HR = 0.55; 95% CI: 0.43–0.69; $P = 2.30 \times 10^{-07}$) (Fig.4 A-C and Supplementary Table S5). After adjusting for age, gender, smoking status, clinical stage, and histology, the association remained statistically significant (Supplementary Table S5). Among 805 patients who had treatment information, a sensitivity analysis to adjust for confounding effects of adjuvant treatment (chemotherapy and/or radiation) showed that the association remained consistent and significant (Supplementary Table S20). For each person, we also built a risk score based on methylation levels (Methods). People with high risk scores (above the median of discovery cohort) have significant shorter survival time than those with low risk scores, in both discovery and validation cohort. Notably, the risk scores of people in validation cohort were calculated based on Cox regression coefficients estimated in discovery cohort (Fig.4 A-C).

DNA methylation of *LRRC3B* is associated with gene expression

DNA methylation plays an important role in regulating gene expression (23). We tested the association between DNA methylation at the three probes and *LRRC3B* expression, as well as the association between gene expression and OS.

Linear regression showed that all three DNA methylation probes (cg13046257, cg17623116, and cg19600115) were positively associated with *LRRC3B* expression in both discovery and validation cohorts (Fig.5 A-C and Supplementary Table S6). In each cohort, regression coefficients for the three probes were similar, indicating potential coherent regulating effects of DNA methylation in that region. Note that expression profiles were measured by different platforms in the discovery and validation cohorts, which resulted in different regression coefficients across the two cohorts. The P-value of probe cg13046257 in the discovery cohort was borderline, which was likely due to the small sample size. Further, expression data from different centers in the discovery cohort were measured by different platforms, which likely induced variation in expression profiles.

LRRC3B expression is associated with overall survival

We further tested the association between *LRRC3B* expression and OS among 216 patients in the discovery cohort, 427 patients in the validation cohort, and 437 patients in the GEO dataset. After adjusting for age, gender, smoking status, clinical stage, and histology,

patients with high expression levels had longer OS than those with low expression levels (discovery: HR = 0.85, 95% CI: 0.57–1.27; validation: HR = 0.63, 95% CI: 0.41–0.98; GEO: HR = 0.59, 95% CI: 0.40–0.87) (Supplementary Table S7). The association in the discovery cohort was not significant, which was likely due to small sample size and platform variations of expression profiles. The associations in the two larger cohorts (validation and GEO) were significant, and the effect sizes were similar.

***LRRC3B* is differently methylated and expressed in tumor tissues against non-tumor tissues**

We also analyzed differences in *LRRC3B* methylation and expression levels between 53 pairs of tumor and adjacent non-tumor tissues. Probes cg13046257 and cg17623116 were hypomethylated in tumor compared to adjacent non-tumor tissues (cg13046257: fold change = 0.92, FDR = 7.06×10^{-08} ; cg17623116: fold change = 0.81, FDR = 1.14×10^{-14}) (Supplementary Table S8). In addition, *LRRC3B* had significantly lower expression in tumor tissues (fold change = 0.26; $P = 1.11 \times 10^{-04}$) (Supplementary Table S9).

Discussion

We investigated putative implications of DNA methylation biomarkers in the tumor suppressor gene *LRRC3B* on 1,230 early-stage NSCLC patients. Patients could be distinguished into different risk groups based on their methylation levels at three loci in *LRRC3B*. The validity of these methylation biomarkers was demonstrated by: 1) associations between DNA methylation and overall survival were replicated in two large populations; 2) coherence with results of potential downstream gene expression pathways in multiple independent large populations; and 3) consistency with previously published studies across different types of cancer.

LRRC3B encodes leucine-rich repeat-containing 3B, which is an evolutionarily highly conserved leucine-rich repeat-containing protein (12). It covers about 88 kb at the human chromosome 3p24.1 locus, with two predicted CpG islands in the promoter region (chr3:26664104–26664796 and chr3:26665950–26666164). Previous studies showed *LRRC3B* is a potential tumor suppressor in gastric and colorectal cancers (12,13). One study showed *LRRC3B* expression is reduced in 90.9% of gastric cancer cell lines and 88.5% of gastric tumor tissues (12), which is consistent with our finding that *LRRC3B* expression is repressed in 75.47% (40/53) of lung cancer tumor tissues.

LRRC3B hypermethylation was reported at six CpG sites around the first CpG island (chr3:26664104–26664796) in the promoter region in 78 gastric tumor tissues, comparing to non-tumor tissues (11), which is consistent with our results (Supplementary Table S8). However, we found that *LRRC3B* was hypomethylated in a non-CpG island region near the gene body (chr3: 26700158–26751141), which covered two of the three probes we identified (Supplementary Table S8). Methylation levels of three sites found in our study all show strong positive relationships with *LRRC3B* expression, consistent with previous results demonstrating that there is no inverse relationship between methylation of non-CpGs and gene expression (23,24).

Solid evidence from functional and *in vitro* experiments supports a tumor-suppressor role for *LRRC3B* across various types of cancers (12–14). Studies have also demonstrated the potential of DNA methylation or expression biomarkers as prognostic biomarkers for some cancer types (12–14). However, most studies focused on *in vitro* experiments or small-scale population data. Our study was based on clinical data from large populations, thus building a bridge between previous experimental studies and clinical applications. Considering the convenience, feasibility, and economy of future applications of the biomarkers, we selected only three methylation probes in a restrictive way and tested robustness of their associations with overall survival in two independent populations. To further validate reliability of the biomarkers, we tested 1) the association between DNA methylation of the three probes and *LRRC3B* expression; and 2) the association between gene expression and OS in at least two independent populations. The results of those analyses are robust and coherent with each other. We also performed sensitivity analysis to provide more information, which can be found in Supplementary Information (Supplementary Table S10-S20 and Supplementary Figure S1 A-C). However, we recognize some limitations of this study. Our study is based on observational data, and DNA methylation and gene expression were measured once at the same time point. The casual mechanism between DNA methylation and gene expression cannot be evaluated in our study.

In summary, we provide evidence of the potential development of a biomarker panel based on DNA methylation of *LRRC3B* for the overall survival of early-stage NSCLC, thus filling the gap between *in vitro* studies and clinical application. Future studies may focus on the additional implications of *LRRC3B* epigenetics, including early-diagnosis biomarkers and drug target discovery.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

We thank all participants in this study. This work was supported by the National Institutes of Health (National Cancer Institute) grants (CA209414, CA092824, CA090578, CA074386 and ES00002 to D.C.C) and the Raymond P. Lavietes Family Fund (to D.C.C.).

Financial Support

This work was supported by the National Institutes of Health and the National Cancer Institute (CA209414, CA092824, CA090578, CA074386 and ES00002 to D.C.C.) and the Raymond P. Lavietes Family Fund (to D.C.C.).

References

1. Cancer Facts & Figures. American Cancer Society Web site. 2017.
2. Govindan R, Page N, Morgensztern D, Read W, Tierney R, Vlahiotis A, et al. Changing epidemiology of small-cell lung cancer in the United States over the last 30 years: analysis of the surveillance, epidemiologic, and end results database. *Journal of clinical oncology* 2006;24(28): 4539–44. [PubMed: 17008692]
3. Khuder SA. Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. *Lung cancer* 2001;31(2–3):139–48. [PubMed: 11165392]

4. Zhu C, Shih W, Ling C, Tsao M. Immunohistochemical markers of prognosis in non-small cell lung cancer: a review and proposal for a multiphase approach to marker evaluation. *Journal of clinical pathology* 2006;59(8):790–800. [PubMed: 16873561]
5. Kanters S, Lammers J, Voest E. Molecular and biological factors in the prognosis of non-small cell lung cancer. *European Respiratory Journal* 1995;8(8):1389–97. [PubMed: 7489807]
6. Botling J, Edlund K, Lohr M, Hellwig B, Holmberg L, Lambe M, et al. Biomarker discovery in non-small cell lung cancer: integrating gene expression profiling, meta-analysis and tissue microarray validation. *Clinical cancer research* 2012;clincanres. 11392012.
7. Costa-Pinheiro P, Montezuma D, Henrique R, Jerónimo C. Diagnostic and prognostic epigenetic biomarkers in cancer. *Epigenomics* 2015;7(6):1003–15. [PubMed: 26479312]
8. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. *New England Journal of Medicine* 2003;349(21):2042–54. [PubMed: 14627790]
9. Baylin SB, Ohm JE. Epigenetic gene silencing in cancer—a mechanism for early oncogenic pathway addiction? *Nature Reviews Cancer* 2006;6(2):107. [PubMed: 16491070]
10. Laird PW. Early detection: the power and the promise of DNA methylation markers. *Nature Reviews Cancer* 2003;3(4):253. [PubMed: 12671664]
11. Højfeldt JW, Agger K, Helin K. Histone lysine demethylases as targets for anticancer therapy. *Nature reviews Drug discovery* 2013;12(12):917. [PubMed: 24232376]
12. Kim M, Kim J-H, Jang H-R, Kim H-M, Lee C-W, Noh S-M, et al. LRRC3B, encoding a leucine-rich repeat-containing protein, is a putative tumor suppressor gene in gastric cancer. *Cancer research* 2008;68(17):7147–55. [PubMed: 18757430]
13. Tian XQ, Zhang Y, Sun D, Zhao S, Xiong H, Fang J. Epigenetic silencing of LRRC3B in colorectal cancer. *Scandinavian journal of gastroenterology* 2009;44(1):79–84. [PubMed: 18815942]
14. Kondratov A, Stoliar L, Kvasha S, Gordiyuk V, Zgonnyk Y, Gerashchenko A, et al. Methylation pattern of the putative tumor-suppressor gene LRRC3B promoter in clear cell renal cell carcinomas. *Molecular medicine reports* 2012;5(2):509–12. [PubMed: 22101383]
15. Sandoval J, Mendez Gonzalez J, Nadal E, Chen G, Carmona FJ, Sayols S, et al. A prognostic DNA methylation signature for stage I non-small-cell lung cancer. *Journal of clinical oncology* 2013;31:4140–7. [PubMed: 24081945]
16. Bjaanaes MM, Fleischer T, Halvorsen AR, Daunay A, Busato F, Solberg S, et al. Genome-wide DNA methylation analyses in lung adenocarcinomas: association with EGFR, KRAS and TP53 mutation status, gene expression and prognosis. *Molecular oncology* 2016;10(2):330–43. [PubMed: 26601720]
17. Karlsson A, Jönsson M, Lauss M, Brunnström H, Jönsson P, Borg Å, et al. Genome-wide DNA methylation analysis of lung carcinoma reveals one neuroendocrine and four adenocarcinoma epitypes associated with patient outcome. *Clinical Cancer Research* 2014.
18. Chen Y-a, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 2013;8(2):203–9. [PubMed: 23314698]
19. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genomics data* 2016;9:22–4. [PubMed: 27330998]
20. Pidsley R, Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC genomics* 2013;14(1):293. [PubMed: 23631413]
21. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 2012;29(2):189–96. [PubMed: 23175756]
22. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;8(1):118–27. [PubMed: 16632515]
23. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics* 2012;13(7):484.

24. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature genetics* 2005;37(8):853. [PubMed: 16007088]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

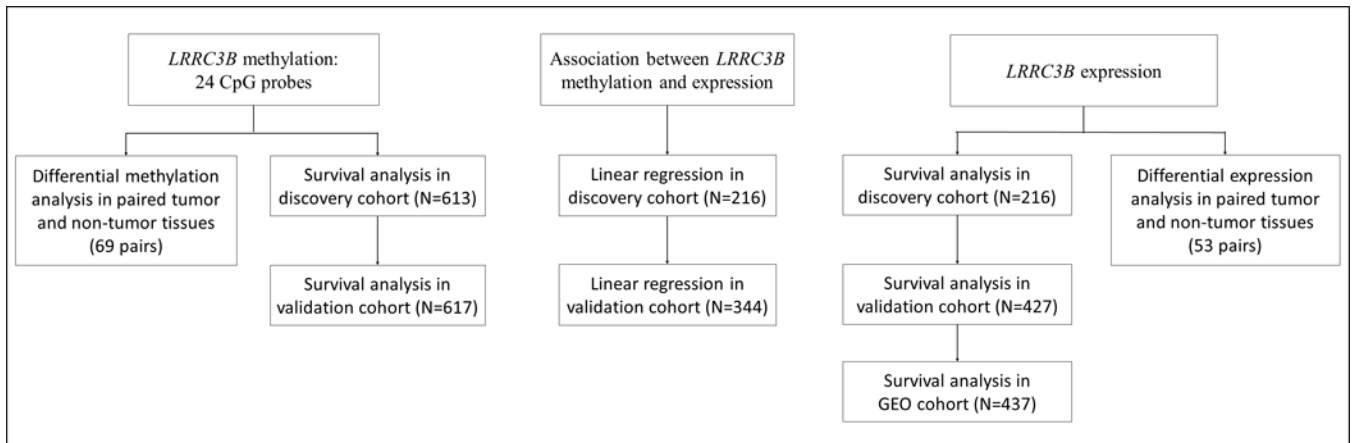


Figure 1: Study roadmap

This study included three parts. First, DNA methylation sites in *LRR3B* were selected and survival models were built based on a discovery cohort containing 613 samples, which were then validated in a validation cohort of 617 samples. Differential methylation analysis was performed in 69 pairs of tumor and adjacent non-tumor tissues. Second, the association between DNA methylation and *LRR3B* expression was investigated in 216 samples and validated in 344 samples. Third, the association between *LRR3B* expression and overall survival was investigated in three groups of patients: a discovery cohort of 216 patients; a validation cohort of 427 patients, and a Gene Expression Omnibus (GEO) cohort of 437 patients. Differential expression analysis was performed in 53 pairs of tumor and adjacent non-tumor tissues.

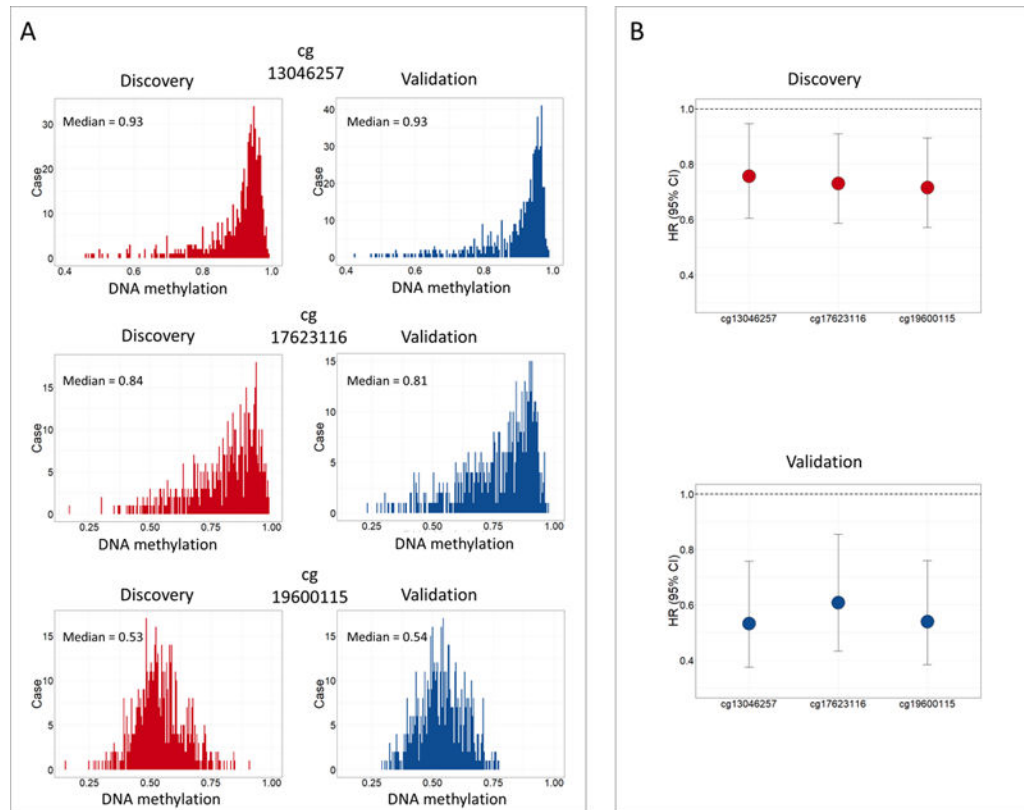


Figure 2: DNA methylation of three selected sites in *LRR3B*

(A) Plots showing consistent distribution of DNA methylation levels for each methylation site in the discovery and validation cohorts. (B) Hazard ratio (HR) with 95% confidence interval (95% CI) of the three methylation sites in Cox regression analysis of discovery and validation cohorts.

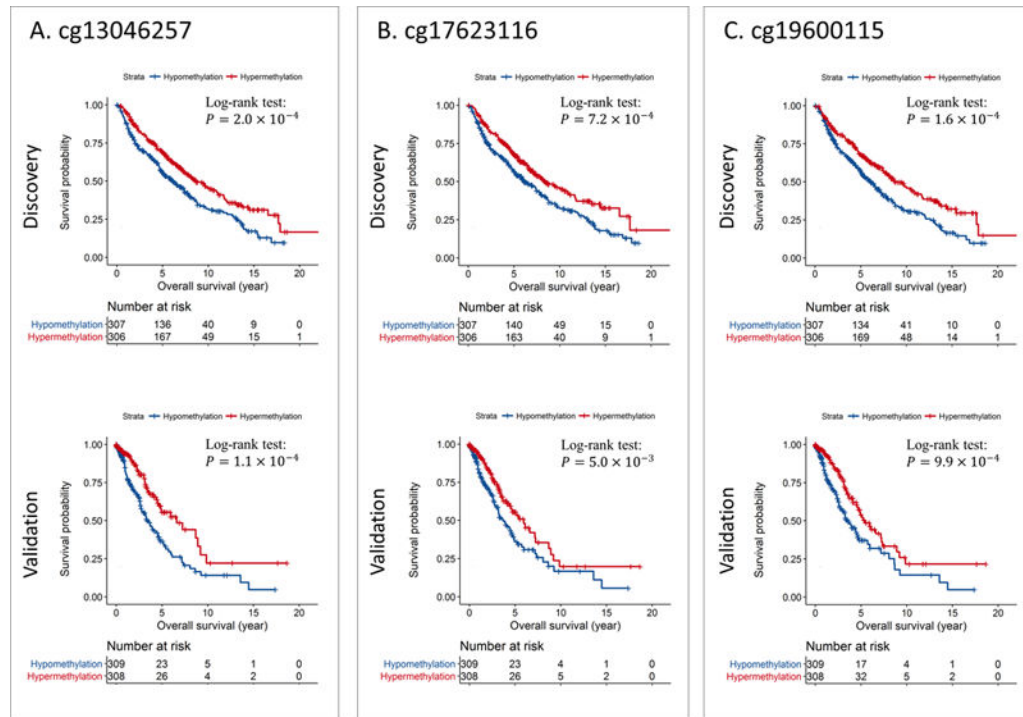


Figure 3: Survival analysis for three selected DNA methylation sites in *LRR3B*
 Kaplan-Meier curves of patients categorized into hypomethylation and hypermethylation groups based on median methylation value for site (A) cg13046257, (B) cg17623116, and (C) cg19600115 in discovery and validation cohorts. Log-rank tests were performed to test differences between survival curves of each group.

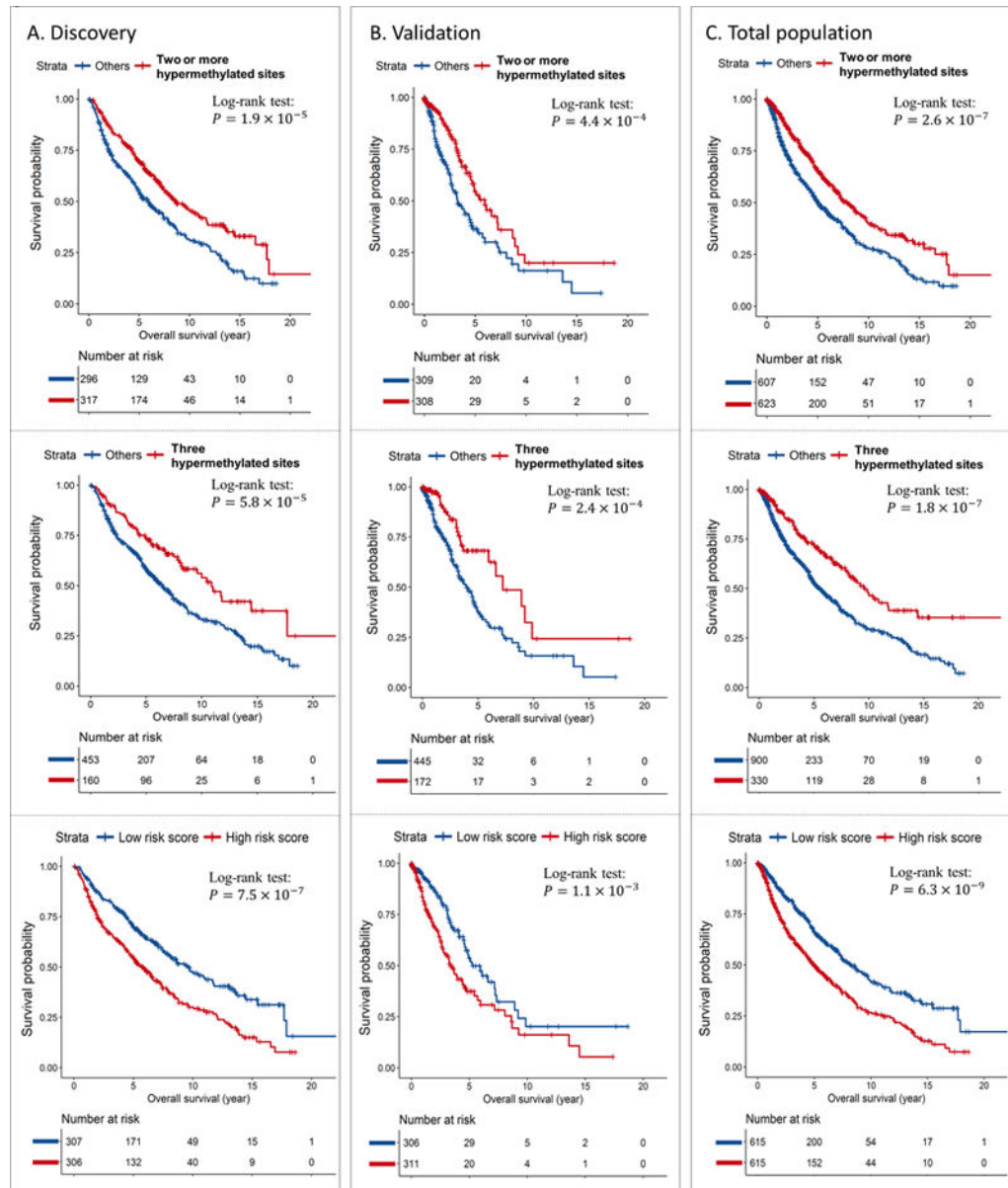


Figure 4: Survival analysis for combining effects of DNA methylation sites in *LRR3B*
 Kaplan-Meier curves of patients categorized into different groups based on methylation levels of all three selected sites or methylation risk scores in the (A) discovery cohort, (B) validation cohort, and (C) total population. Log-rank tests were performed to test differences between survival curves of each group.

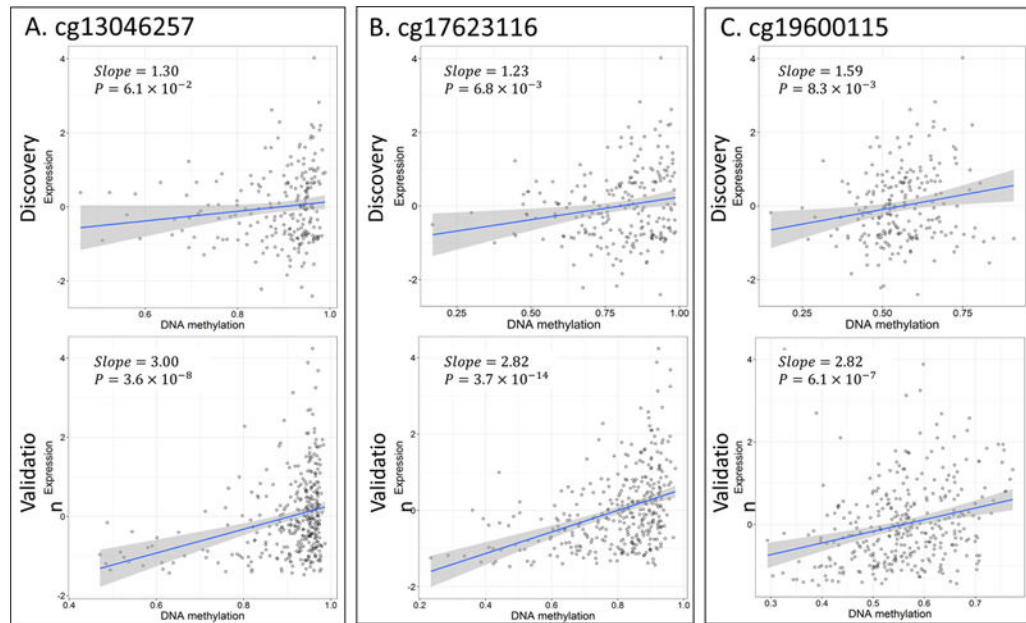


Figure 5: Associations between DNA methylation and *LRR3B* expression

Scatter plots of standardized *LRR3B* expression and DNA methylation levels for site (A) cg13046257, (B) cg17623116, and (C) cg19600115 in discovery and validation cohorts.

Linear models were fitted and tested with Wald tests.