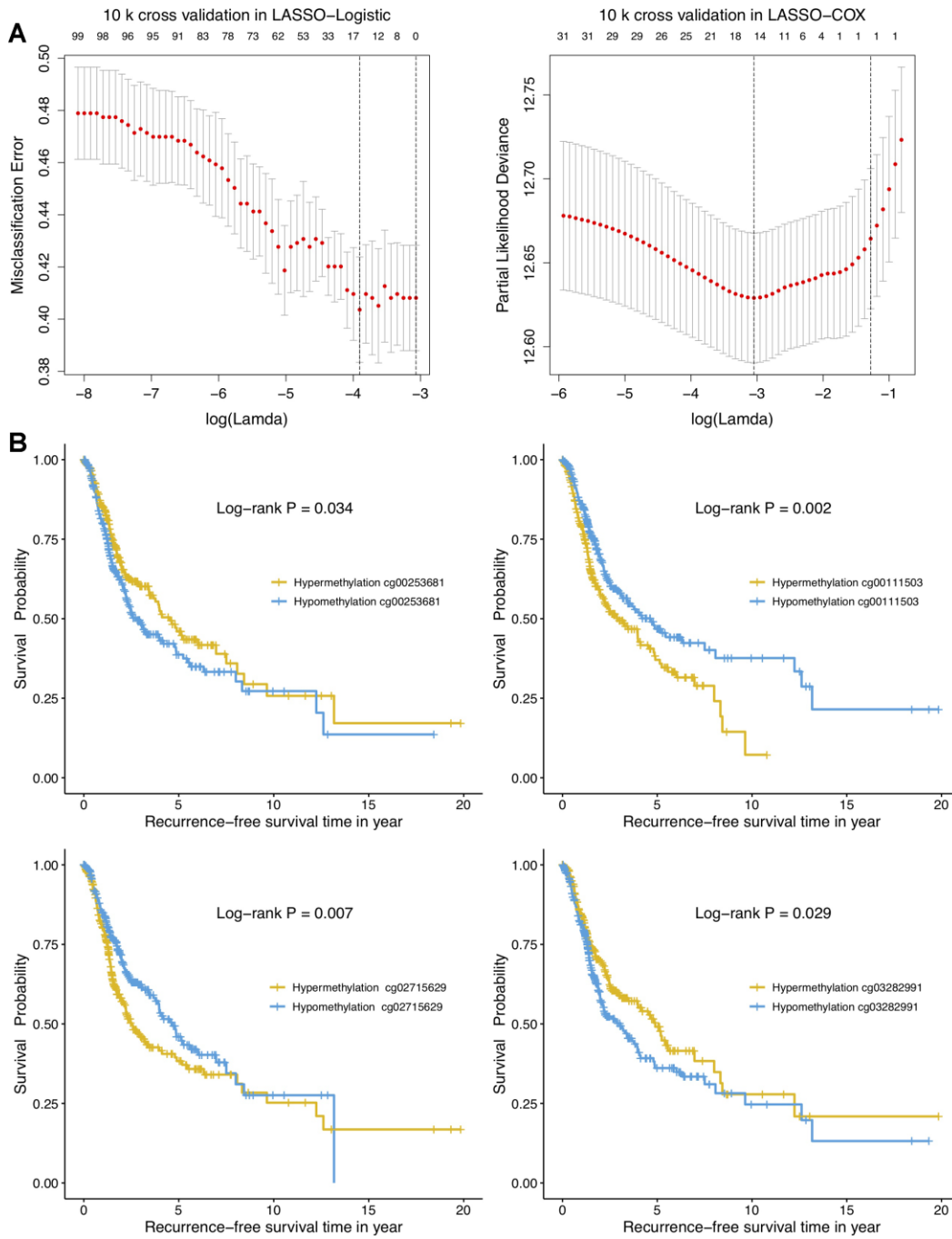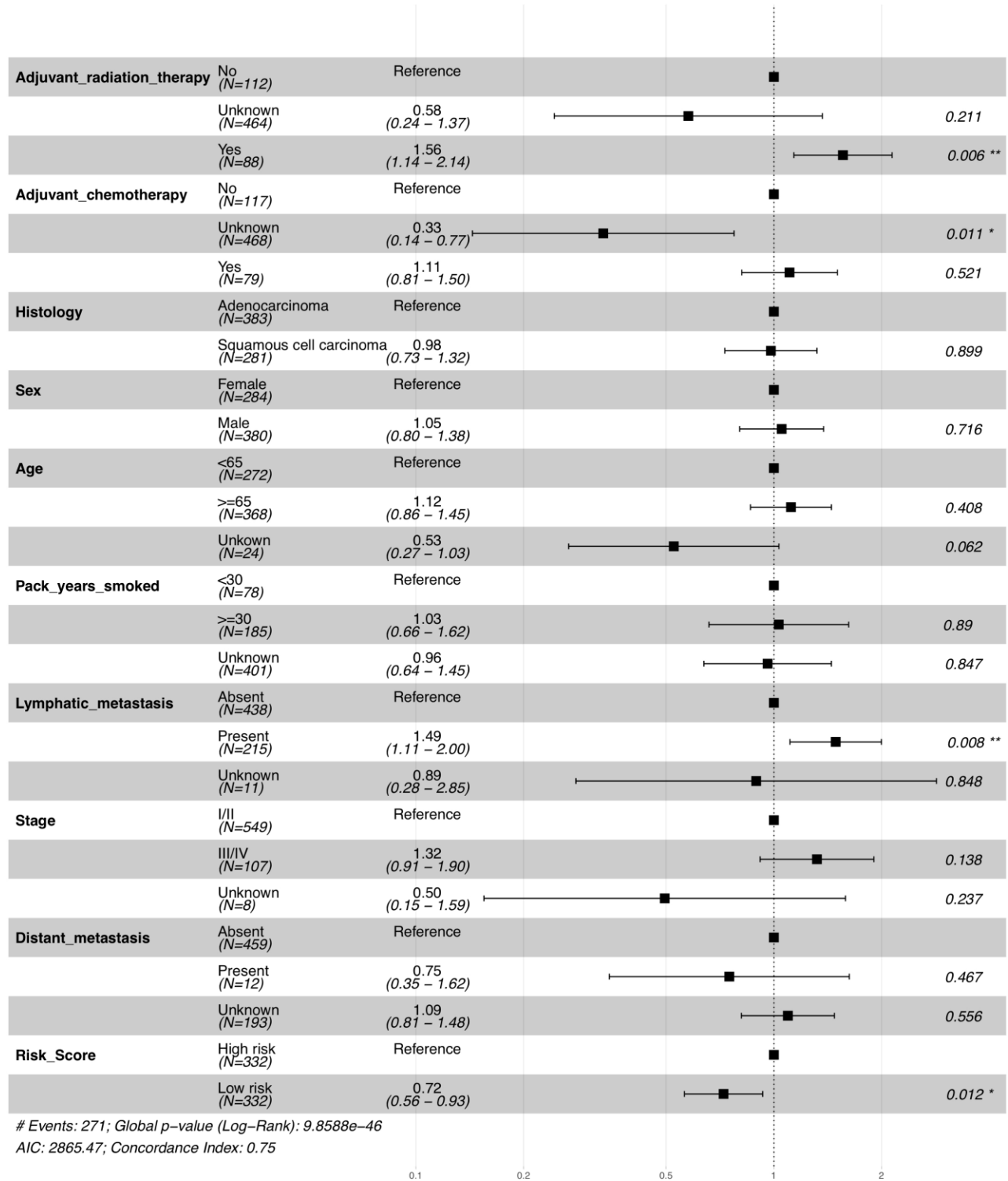**Supplementary Figure 1.** (**A**) Ten-fold cross validations performed for obtaining optimal parameter lambda (λ) in LASSO-Logistic (left) and LASSO-Cox analysis (right). The dotted vertical lines were plotted at values of log(λ) by minimum criteria and 1-Standard Error criteria in two LASSO models, respectively. The optimal values of λ were determined by minimum criteria where two dotted vertical lines were drawn in Figure 3A and Figure3D, and thus 14 and 9 CpG markers with nonzero coefficients were screened out. (**B**) Kaplan-Meier curves of four final selected CpGs present their correlation with recurrence and prognostic prediction of NSCLC patients in training cohort.
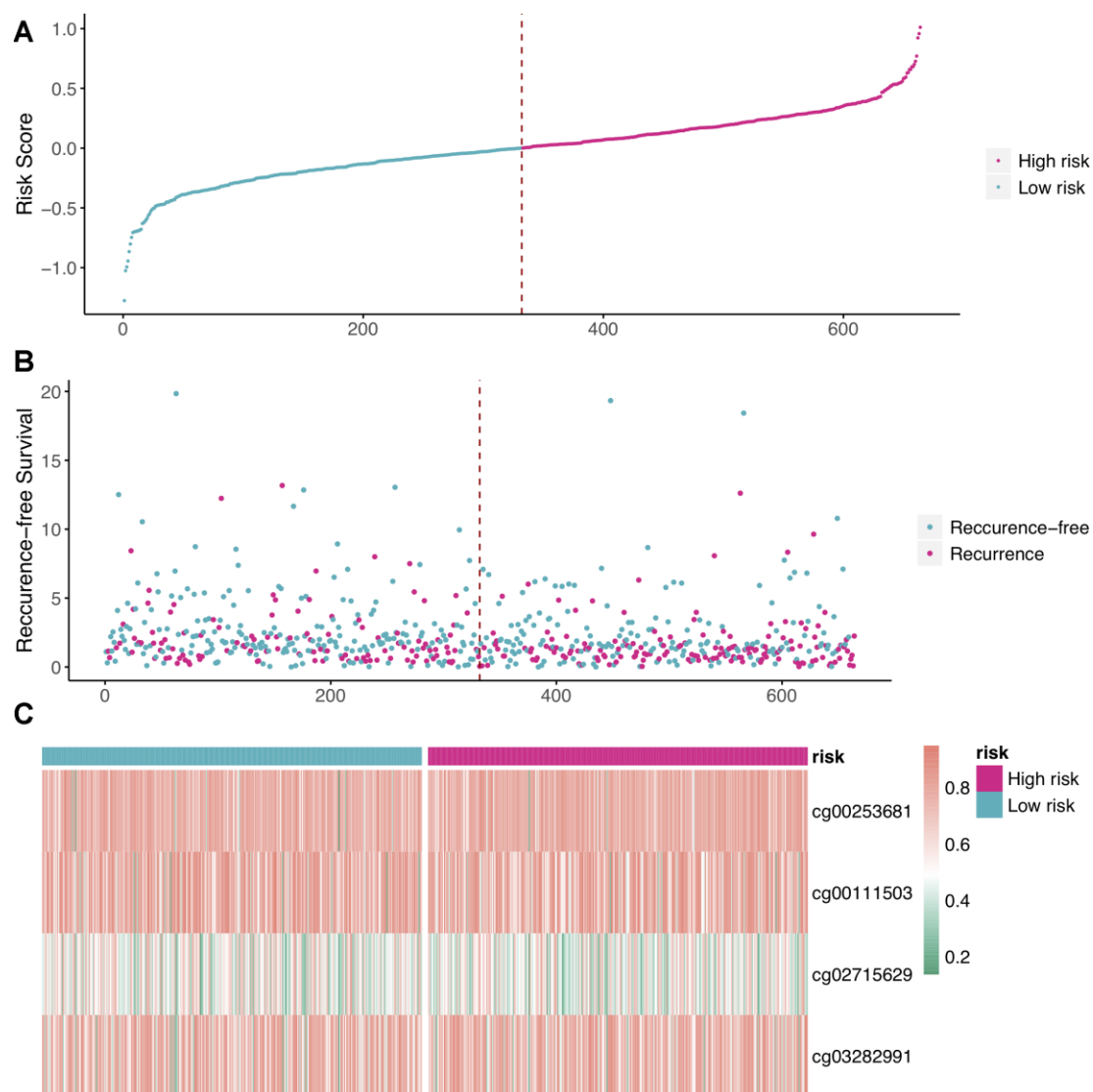
Supplementary Figure 2. Association chart of relapse risk factors. (**A**, **B**) Distribution of risk score and RFS of TCGA NSCLC patients in high- and low-risk groups. (**C**) The heatmap of DNAm profile of 4 final selected CpGs.

Hazard ratio for disease recurrence of TCGA–LUNG

Adjuvant_radiation_therapy	No (N=112)	Reference	
	Unknown (N=464)	0.58 (0.24 – 1.37)	0.211
	Yes (N=88)	1.56 (1.14 – 2.14)	0.006 **
Adjuvant_chemotherapy	No (N=117)	Reference	
	Unknown (N=468)	0.33 (0.14 – 0.77)	0.011 *
	Yes (N=79)	1.11 (0.81 – 1.50)	0.521
Histology	Adenocarcinoma (N=383)	Reference	
	Squamous cell carcinoma (N=281)	0.98 (0.73 – 1.32)	0.899
Sex	Female (N=284)	Reference	
	Male (N=380)	1.05 (0.80 – 1.38)	0.716
Age	<65 (N=272)	Reference	
	>=65 (N=368)	1.12 (0.86 – 1.45)	0.408
	Unkown (N=24)	0.53 (0.27 – 1.03)	0.062
Pack_years_smoked	<30 (N=78)	Reference	
	>=30 (N=185)	1.03 (0.66 – 1.62)	0.89
	Unknown (N=401)	0.96 (0.64 – 1.45)	0.847
Lymphatic_metastasis	Absent (N=438)	Reference	
	Present (N=215)	1.49 (1.11 – 2.00)	0.008 **
	Unknown (N=11)	0.89 (0.28 – 2.85)	0.848
Stage	I/II (N=549)	Reference	
	III/IV (N=107)	1.32 (0.91 – 1.90)	0.138
	Unknown (N=8)	0.50 (0.15 – 1.59)	0.237
Distant_metastasis	Absent (N=459)	Reference	
	Present (N=12)	0.75 (0.35 – 1.62)	0.467
	Unknown (N=193)	1.09 (0.81 – 1.48)	0.556
Risk_Score	High risk (N=332)	Reference	
	Low risk (N=332)	0.72 (0.56 – 0.93)	0.012 *

Events: 271; Global p–value (Log–Rank): 9.8588e–46
AIC: 2865.47; Concordance Index: 0.75

Supplementary Figure 3. Multivariate Cox regression analysis for RFS of TCGA NSCLC patients with combinations of DNAm-based risk score and clinical factors.
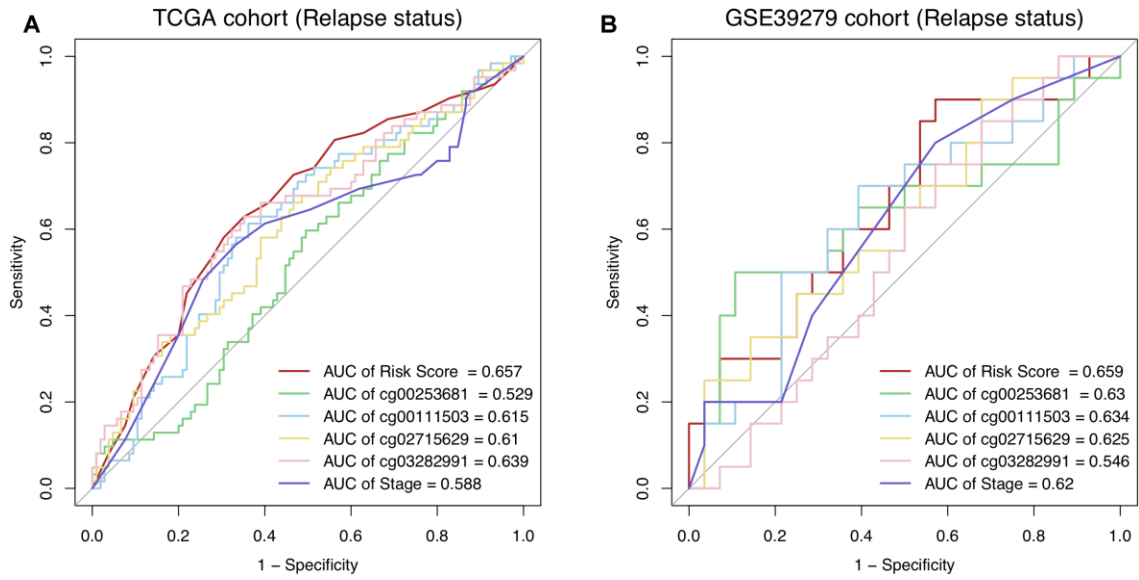
Supplementary Figure 4. Receiver operating characteristic (ROC) curves of the combined risk score, clinical stage and 4 separate CpGs demonstrate their performance of discrimination for relapse status in training (**A**) and validation (**B**) cohorts.
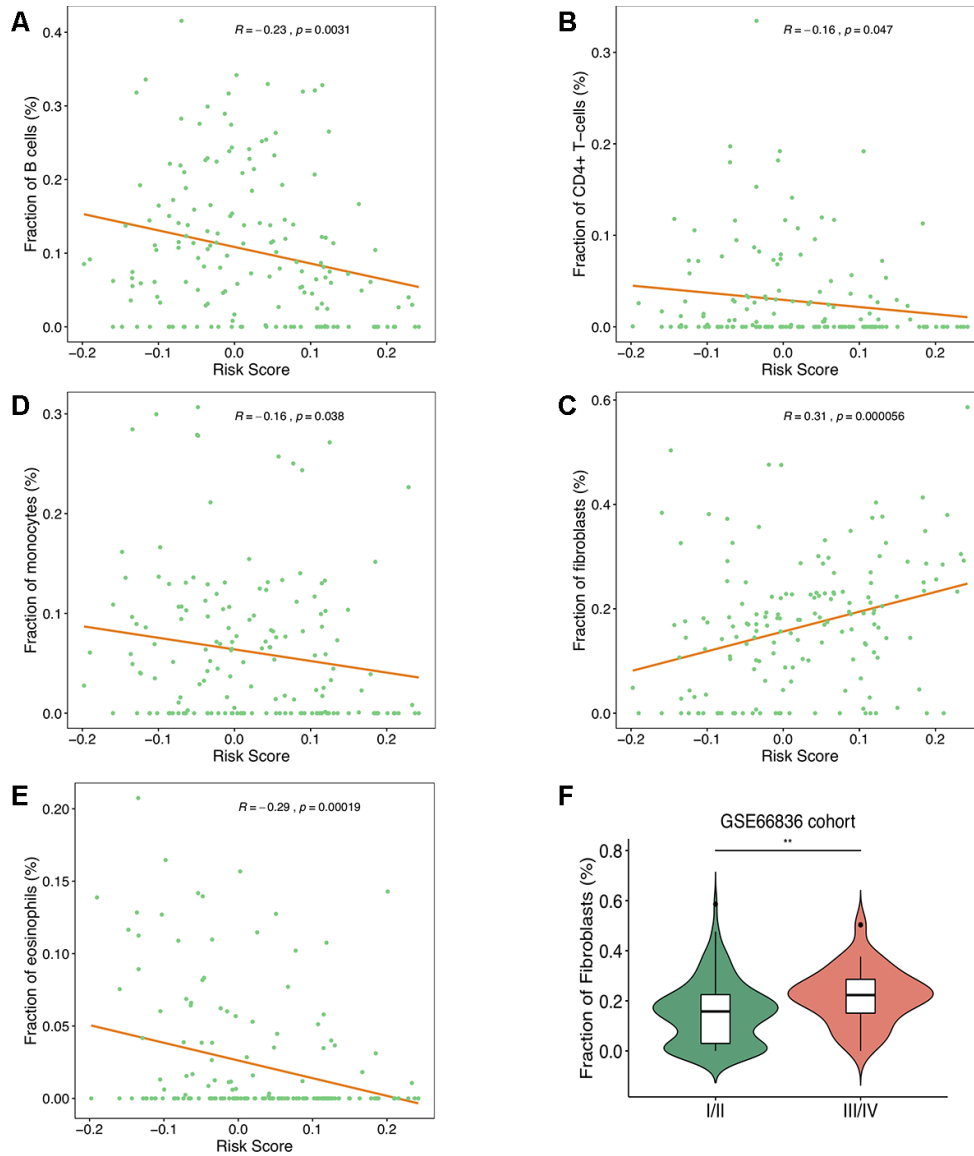
Supplementary Figure 5. Estimated compositions of B-cells (**A**), CD4+ T-cells (**B**), fibroblasts (**C**), monocytes (**D**) and eosinophils (**E**) were significantly correlated with DNAm-based risk score in GSE66836 cohort. (**F**) Fraction of fibroblasts in NSCLC samples at late stage was significantly higher than those at early stage.
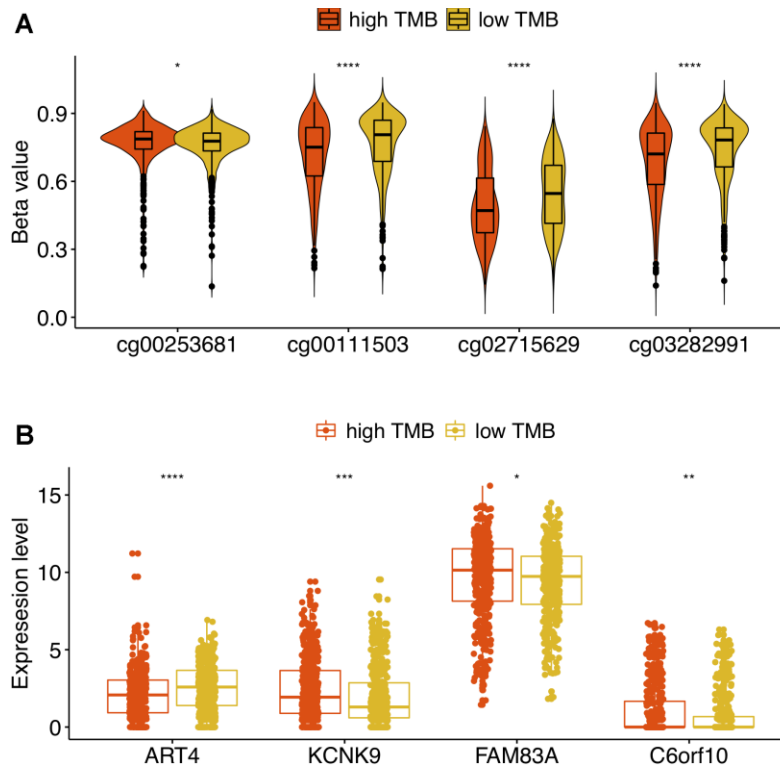
Supplementary Figure 6. The correlation of TMB with four CpGs methylation status and expression of four nearby genes. (**A**) Methylation levels of 4 identified DMPs in high TMB compared with low TMB group. (**B**) Differential expression of 4 reference genes in high and low TMB.