Supplementary Table 1. List of genes, primers and sequences

Gene	PCR primers	Sequencing primers	Modified sequence/corresponding unmodified sequence analyzed
CDH1 (sense)	5'-TTTGATTTTAGGTTTTAGTGAGT	5'-TAGTAATTTTAGGTTAGAGG	5'-GTTAT**Y**G**YG**TTTATG**YG**AGGT**YG**GGTGGG**YG**GGT**YG**
CDH1 (antisense)	5'-BIOTIN-ACCACAACCAATCAACAA		5'-GTCAC**CG****CG**TCTATGC**G**AGGCC**G**GGTGGGC**G**GGCC**G**
CDKN2A (sense)	5'-TTGAGGGTGGGAAGATGGT	5'-GGAGGGAGAGGAA	5'-**YGYG**GGTTTTGAGT**YG**TT**YGYGYGYG**
CDKN2A (antisense)	5'-BIOTIN-CCCRAACCTCCAAAATCTC		5'-**CGCG**GGCCCTGAGC**CG**CC**CGCGCGCG**
GSTP1 (sense)	5'-GTGATTTAGTATTGG	5'-AGTTTTTGTTATTAGTGAGTA	5'-**YGYGYG**GTT**YGYG**TTTT**YG**
GSTP1 (antisense)	5'-BIOTIN-AACTCTAAACCCCATC		5'-**CGCGCG**GCCC**GCG**TCCCC**G**
LINE-1 (sense)	5'-BIOTIN-TAGGGAGTGTTAGATAGTGG	5'-AACTCCCTAACCCCTTAC	5'-**R**CCCTACTT**CR**ACT**CRCR**CA**CR**ATAC**R**
LINE-1 (antisense)	5'-AACTCCCTAACCCCTTAC		5'-**G**CCCTGCTT**CG**GCT**CGCG**CAC**G**GTGC**G**
MTHFR (sense)	5'-BIOTIN-TTTTAATTTTTGTTTGGAGGGTAGT	5'-TGAGAGTTTTAAAGATAGTT	5'-**YG**AGATGTTTTATTT**YG**
			5'-**CG**AGATGTTCCACCC**CG**
		5'-GGGTTTGGATTTTGAG	5'-**YG**GTATGAGAGATTT**YG**GGAGAAGATGAGG**YG**GY**G**
MTHFR (antisense)	5'-BIOTIN-AAAAAAACCACTTATCACCAAATTC		5'-**CG**GCATGAGAGACTC**CG**GGAGAAGATGAGGC**G**GC**G**
RASSF1A (sense)	5'-AGTTTGGATTTTGGGGGAGG	5'-GGGTTAGTTTTGTGGTTT	5'-**YG**TT**YG**GTT**YGYG**TTTGTTAG**YG**TTTAAAGTTAG**YG**
RASSF1A (antisense)	5'-BIOTIN-CAACTCAATAAACTCAAACTCCCC		5'-**CG**CCC**GG**CCC**GCG**CTTGCTAG**CG**CCCAAAGCCAGC**G**

Analyzed CpG sites are highlighted

Underline sequences are SNPs

Supplementary Table 2. Frequency of CpG hypermethylation in lung tumours

| Gene	Hypermethylated tumour samples*, (%)	Methylation level cut-off
CDH1	1.2	15.3
GSTP1	0.6	4.2
CDKN2A	91.5	9.2
RASSF1A	36.0	4
MTHFR	39.4	75.6

*Samples with methylation levels above the quantile representing the upper 95% of methylation in blood samples

Supplementary Table 3. Multivariate analysis of association between DNA methylation levels over CpG sites of *MTHFR* and smoking status

Smoking Status	CpG sites						
	CpG1 Mean (*P*)	CpG2 Mean (*P*)	CpG3 Mean (*P*)	CpG4 Mean (*P*)	CpG5 Mean (*P*)	CpG6 Mean (*P*)	All sites Mean (*P*)
Blood							
Never	34 -	46 -	49 -	60 -	65 -	53 -	51 -
0-20 py	18 (0.01)	27 (0.03)	40 (0.17)	51 (0.34)	56 (0.04)	44 (0.07)	39 (<0.01)
20-40 py	31 (0.89)	41 (0.73)	45 (0.51)	55 (0.51)	61 (0.36)	50 (0.61)	47 (0.02)
40-60 py	27 (0.14)	37 (0.32)	44 (0.20)	54 (0.21)	62 (0.44)	51 (0.72)	46 (<0.01)
60+ py	36 (0.96)	52 (0.65)	47 (0.98)	59 (0.99)	64 (0.95)	52 (0.99)	52 (0.99)
Never	33 -	46 -	49 -	60 -	65 -	53 -	51 -
Ever	28 (0.11)	40 (0.19)	44 (0.06)	55 (0.08)	61 (0.06)	50 (0.15)	46 (<0.01)
Tumours							
Never	31 -	35 -	50 -	51 -	75 -	66 -	52 -
0-20 py	55 (0.02)	73 (<0.01)	63 (0.24)	72 (0.07)	77 (0.98)	67 (0.99)	68 (<0.01)
20-40 py	70 (<0.01)	87 (<0.01)	69 (0.04)	85 (<0.01)	74 (0.99)	68 (0.96)	76 (<0.01)
40-60 py	49 (0.09)	65 (<0.01)	63 (0.23)	66 (0.21)	75 (0.99)	68 (0.99)	65 (<0.01)
60+ py	65 (0.01)	87 (<0.01)	78 (0.02)	79 (0.05)	84 (0.56)	79 (0.18)	79 (<0.01)
Never	30 -	34 -	50 -	50 -	76 -	66 -	51 -
Ever	61 (<0.01)	77 (<0.01)	66 (0.04)	76 (<0.01)	75 (0.97)	68 (0.79)	71 (<0.01)

Supplementary Figure 1
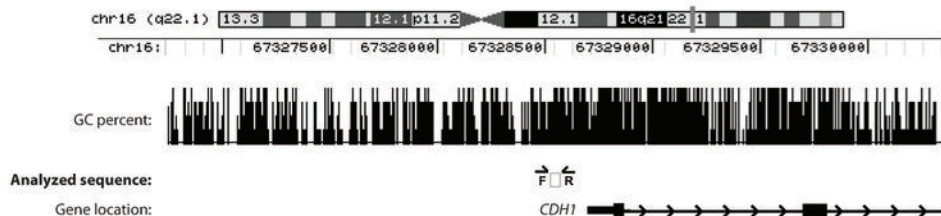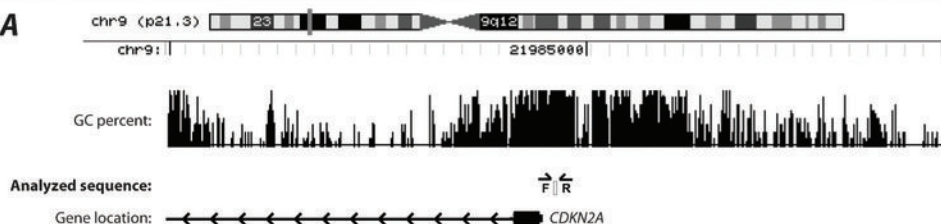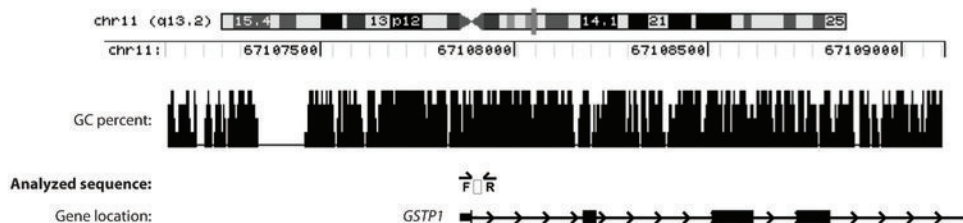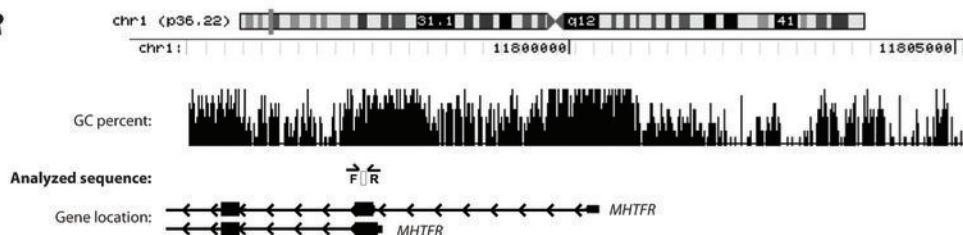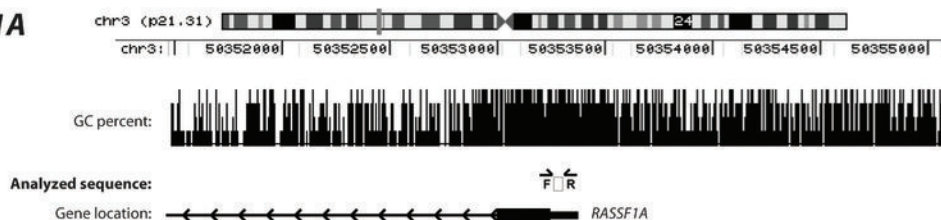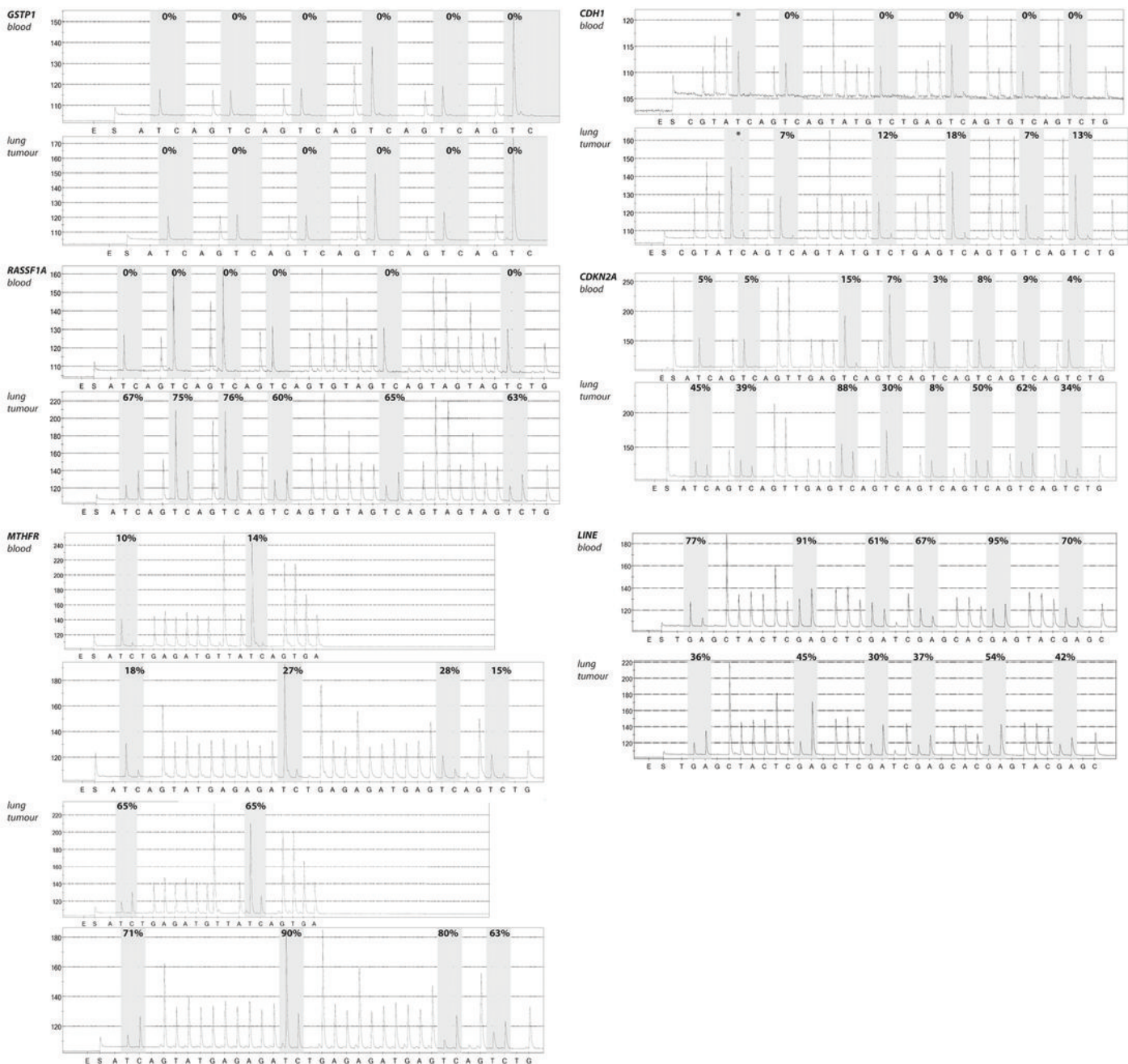

Supplementary Figure 1. Schematic representation of the chromosomal loci of the genes analyzed. The CG density is indicated by vertical bars. Exons are indicated by black box and arrows indicate the orientation of each gene. Location of the analyzed sequence for each gene is represented by empty box and is obtained from the UCSC genome browser. Sets of primers (F - forward; R – reverse) for each gene analyzed are indicated by arrows. The LINE-1 sequences analyzed are repetitive elements in the genome and therefore are not shown here (accession number M80343).
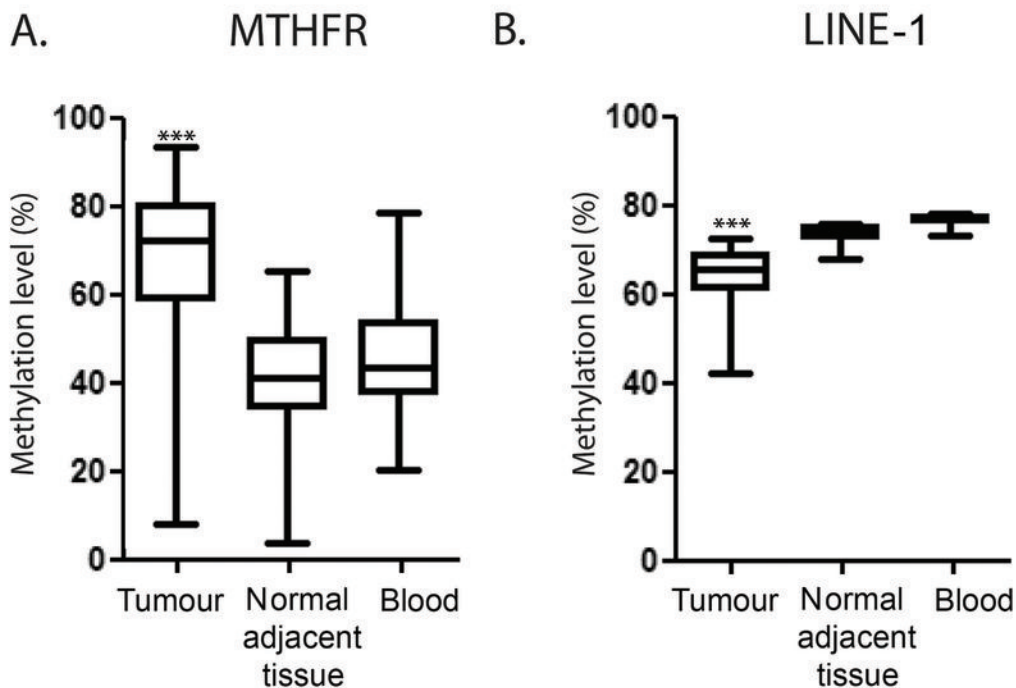
Supplementary Figure 2

Supplementary Figure 2. Examples of pyrograms obtained from the analysis of DNA extracted from lung tumour (upper panel) and blood sample (lower panel) of the same patient. Representative programs of GSTP1, RASSF1A, CDH1, CDKN2A, MTHFR, and LINE-1 are shown.

Supplementary Figure 3
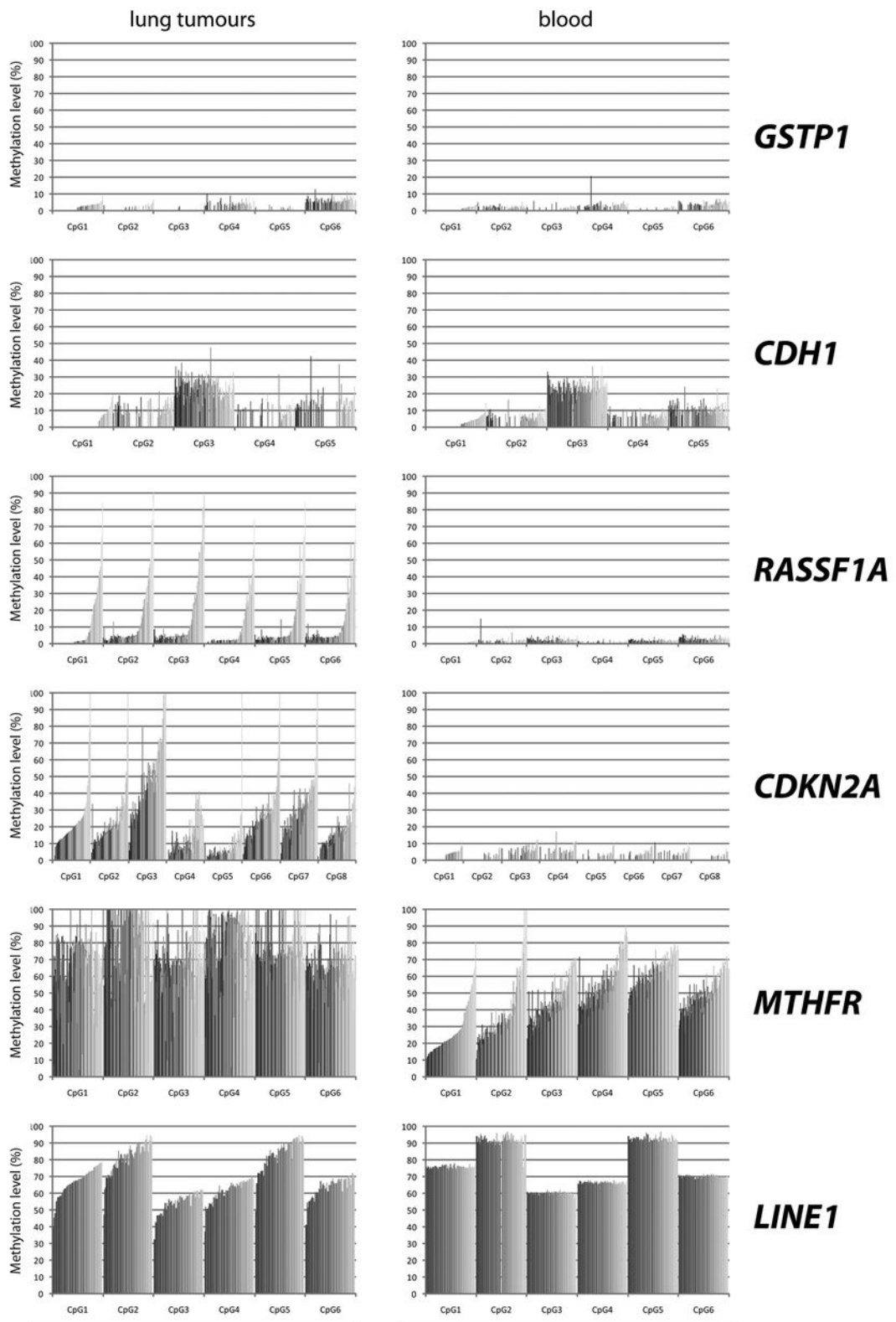
Supplementary Figure 3. Summary of the analysis of DNA methylation in 5 genes and LINE-1 in lung tumours and blood controls.

Supplementary Figure 4

Supplementary Figure 4. Differential methylation of MTHFR and LINE-1 in lung tumours and normal appearing adjacent tissues. (A) Level of methylation obtained for MTHFR in lung tumours (n=209), normal appearing adjacent tissues (n=51) and blood (n=336) are summarized. (B) Levels of LINE-1 methylation in lung tumours (n=55), normal appearing adjacent tissues (n=51) and blood (n=64). The statistical significance for differential methylation in tissues analyzed was calculated using Newman-Keuls' test. The groups exhibiting statistically different methylation levels at $P < 0.01$ are indicated (***).