**Table 1. Characteristics of eligible studies considered in the report**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Sample**  **Type** | **Age**a | **Stage**  **I%** | **Stage**  **(I+II)%** | **Gender**  **Ratio** | **Patients**  **(M/T)** | **Control**  **(M/T)** | **Methods** | **Aim** | **Multipletarget** | **Control design** | **Ad%** | **primer set** |
| Dong et al | tissue | 63 | 0.579545455 | 0.761363636 | 0.795454545 | 26/88 | 7/88 | MSP | Non-Diagnosis | Multi | hom | 0.397727273 | 1 |
| Feng et al | tissue | 64.3 | 0.428571429 | 0.775510204 | 0.530612245 | 28/49 | 1/49 | qMSP | Non-Diagnosis | Multi | hom | 0.588235294 | 1 |
| Hanabata et al | tissue | NA | 0.685714286 | 0.842857143 | 0.628571429 | 26/70 | 2/30 | MSP | Non-Diagnosis | Multi | hom | 0.685714286 | 1 |
| Hsu et al | plasma | 66.81 | NA | 0.650793651 | NA | 21/63 | 6/36 | MSP | Diagnosis | Multi | heter | NA | 2 |
| Hsu et al | tissue | 66.81 | NA | 0.650793651 | NA | 28/63 | 10/63 | MSP | Diagnosis | Multi | hom | 0.759259259 | 1 |
| Jin et al | tissue | 66.7 | NA | NA | 0.708333333 | 25/72 | 2/63 | qMSP | Non-Diagnosis | Multi | hom | 0.651515152 | 1 |
| Nikolaidis et al | tissue | 65.58 | NA | 0.770833333 | 0.520833333 | 11/48 | 0/48 | qMSP | Diagnosis | Multi | hom | 0.5 | 3 |
| Toyooka et al | tissue | NA | NA | NA | NA | 18/42 | 2/25 | MSP | Non-Diagnosis | Single | hom | 0.738095238 | 1 |
| Toyooka et al | tissue | 63 | NA | NA | 0.690661479 | 180/514 | 5/84 | MSP | Non-Diagnosis | Multi | hom | 0.606490872 | 1 |
| Tsou et al | tissue | NA | NA | NA | NA | 37/51 | 24/49 | qMSP | Diagnosis | Multi | both | 1 | 2 |
| Ulivi et al | serum | 70 | 0.557377049 | 0.721311475 | 0.803278689 | 14/61 | 0/15 | qMSP | Diagnosis | Multi | heter | 0.692307692 | 1 |
| Wang et al | tissue | NA | NA | NA | 0.607142857 | 15/28 | 3/12 | MSP | Diagnosis | Multi | hom | 0.681818182 | 2 |
| Zhai et al | serum | 62.39 | 0.095238095 | 0.142857143 | 0.761904762 | 23/42 | 0/40 | MSP | Non-Diagnosis | Multi | heter | 0.761904762 | 1 |
| Zhang et al | plasma | 59 | 0.320512821 | 0.743589744 | 0.743589744 | 37/110 | 2/50 | MSP | Non-Diagnosis | Multi | heter | NA |  |
| Zhang et al | tissue | 59 | 0.320512821 | 0.743589744 | 0.743589744 | 38/78 | 8/78 | MSP | Non-Diagnosis | Multi | hom | 0.454545455 |  |

amean or median age from articles; bwith two records since there are Tissue and serum data simultaneously in this article. M and T means methylation positive and total, respectively.

**Table 2 Subgroup analysis for the main potential confounding factors with fixed effects model**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Subgroup | Number of study | OR | lower | upper | I2 | P-value |
| Overall | 15 | 6.74 | 4.46 | 10.2 | 35.50% | 0.0847 |
| Age<65 | 6 | 8.91 | 5.59 | 14.2 | 39.00% |  |
| Age>=65 | 5 | 4.88 | 2.76 | 8.63 | 33.20% | 0.0801 |
| Stage I<50.45% | 4 | 16.15 | 8.04 | 31.05 | 47.00% |  |
| Stage I>=50.45% | 3 | 10.85 | 6.68 | 17.62 | 0.00% | 0.05 |
| Stage I+II<65% | 6 | 10.4 | 6.58 | 16.43 | 42.00% |  |
| Stage I+II>=65% | 5 | 5.83 | 3.37 | 10.11 | 37.00% | 0.11 |
| M2F<69% | 6 | 10.17 | 5.63 | 18.36 | 19.00% |  |
| M2F>=69% | 5 | 7.89 | 4.55 | 13.67 | 10.60% | 0.62 |
| MSP | 10 | 5.98 | 4.2 | 8.51 | 12.30% |  |
| qMSP | 5 | 6.34 | 3.31 | 12.13 | 65.00% | 0.44 |
| Diagnose | 6 | 3.48 | 2.18 | 5.58 | 0.00% |  |
| Non-diagnose | 9 | 9.26 | 6.13 | 13.97 | 9.00% | 0.002 |
| Multiple targets | 14 | 5.98 | 4.36 | 8.2 | 39.60% |  |
| Single target | 1 | 8.62 | 1.8 | 41.41 | - | 0.83 |
| Heterogeneous | 4 | 5.6 | 2.58 | 12.16 | 59.70% |  |
| Autogenous | 11 | 6.15 | 4.39 | 8.63 | 29.60% | 0.72 |
| Plasma | 4 | 5.6 | 2.58 | 12.16 | 59.70% |  |
| Tissue | 11 | 6.15 | 4.39 | 8.63 | 29.60% | 0.72 |
| Ad%<65% | 4 | 7.8 | 4.68 | 12.99 | 0.00% |  |
| Ad%>=65% | 9 | 7.63 | 5.05 | 11.54 | 52.00% | 0.95 |
| Asian | 9 | 6.79 | 4.7 | 9.81 | 31.00% |  |
| Caucasian | 6 | 8.11 | 4.88 | 13.5 | 55.00% | 0.58 |
| Primerset I | 11 | 7.82 | 1.91 | 12.69 | 35.00% |  |
| Primerset II | 1 | 2.75 | 1.20 | 6.32 | - |  |
| Primerset III | 1 | 15.25 | 3.44 | 67.53 | - |  |
| Primerset IV | 1 | 29.75 | 1.70 | 521.15 | - | 0.06 |

**Table 3. Differential FHIT methylation, odds ratio between adenocarcinoma, squamous cell carcinoma and their counterparts from TCGA dataset**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Type** | **CpG site** | **Position** | **McaM** | **McoM** | **∆β** | **P-valuea** | **P-valueb** | **ORb** | **95%CIb** |
| **LUAD** | cg22215728 | 61236652 | 0.12 | 0.16 | **0.04** | 0.00081 | 0.0097 | 4.33 | 1.93-12.55 |
| cg15931943 | 61236909 | 0.10 | 0.12 | **0.02** | 0.00718 | 0.18873 | 1.80 | 0.91-4.62 |
| cg02854288 | 61236911 | 0.11 | 0.13 | **0.02** | 0.00168 | 0.04887 | 2.81 | 1.29-7.97 |
| cg19049316 | 61237063 | 0.03 | 0.03 | **0.00** | 0.22515 | 0.14417 | 1.73 | 0.93-3.58 |
| cg26322434 | 61237156 | 0.03 | 0.04 | **0.01** | 0.29934 | 0.17031 | 1.63 | 0.89-3.28 |
| cg24796403 | 61237172 | 0.04 | 0.04 | **0.00** | 0.66265 | 0.40238 | 1.35 | 0.73-2.89 |
| cg16986494 | 61237206 | 0.04 | 0.05 | **0.01** | 0.01932 | 0.04887 | 3.12 | 1.33-9.44 |
| cg12030002 | 61237226 | 0.04 | 0.05 | **0.01** | 0.2052 | 0.24634 | 1.49 | 0.82-2.9 |
|  |  |  |  |  |  |  |  |  |  |
| **LUSC** | cg22215728 | 61236652 | 0.10 | 0.14 | **0.04** | 1.00E-05 | 0.00066 | 3.92 | 2.07-8.44 |
| cg15931943 | 61236909 | 0.09 | 0.10 | **0.01** | 1.00E-05 | 0.01163 | 2.51 | 1.35-5.13 |
| cg02854288 | 61236911 | 0.08 | 0.10 | **0.02** | 1.00E-05 | 0.00485 | 2.76 | 1.5-5.58 |
| cg19049316 | 61237063 | 0.03 | 0.03 | **0.00** | 0.52339 | 0.31038 | 0.66 | 0.26-1.16 |
| cg26322434 | 61237156 | 0.03 | 0.02 | **0.01** | 0.51658 | 0.29719 | 0.77 | 0.45-1.21 |
| cg24796403 | 61237172 | 0.04 | 0.03 | **0.01** | 0.52339 | 0.29486 | 0.31 | 0.03-1.13 |
| cg16986494 | 61237206 | 0.04 | 0.03 | **0.01** | 0.02458 | 0.54451 | 0.84 | 0.26-1.37 |
| cg12030002 | 61237226 | 0.04 | 0.03 | **0.01** | 0.11177 | 0.12536 | 0.10 | 0.005-0.99 |

McaM and McoM represent the mean of case methylation (Beta) and mean of control methylation (Beta). Methylation levels are calculated with formula: Beta = (M/M + U).

Position represents the chromosome position of each CpG site according to GRCh37/hg19.

P-valuesa are calculated from Wilcoxon rank sum test after false discovery rate(FDR adjustment).

P-valueb and ORb and 95%CIb are from logistic regression analysis with P-valueb were also after false discovery rate(FDR adjustment).

**Figure 1**

**Figure 2**

**Figure 3 CpG sites on the Illumina Infinium HumanMethylation450 Beadchip array across FHIT gene region and Gene expression scatterplot with paired data from TCGA dataset**

Methylation and gene expression status for FHIT gene [TCGA lung cancer dataset]. A-B each represents the different methylation status of lung cancer subtypes versus normal lung tissues in different datasets. For A-B, the x-axis shows the different CpG sites in FHIT genes and the y-axis shows the beta value of each CpG site to represent the methylation level of each CpG site. C-D each represents the gene expression status of paired samples. The x-ais of the two figures shows the different types and y-axis shows the gene expression level using RPKM as measurement.