Table 1. Sample size distribution from each population in the study.

|  |  |  |  |
| --- | --- | --- | --- |
|  | EUR | EAS | AA |
| Strata | CONTROL | CASE | Total | CONTROL | CASE | Total | CONTROL | CASE | Total |
| *Ever-smokers*  |
| Overall | 16165 | 22018 | 38183 | 1032 | 1495 | 2527 | 2309 | 1804 | 4113 |
| ADE | 16165 | 7838 | 24003 | 1032 | 586 | 1618 | 2309 | 734 | 3043 |
| SQC | 16165 | 5619 | 21784 | 1032 | 514 | 1546 | 2309 | 436 | 2745 |
| SCLC | 16165 | 1919 | 18084 | 1032 | 88 | 1120 | 2309 | 111 | 2420 |
| *Never-smokers* |
| Overall | 6396 | 2207 | 8603 | 4335 | 5561 | 9896 | 1405 | 170 | 1575 |
| ADE | 6396 | 1268 | 7664 | 4335 | 4019 | 8354 | 1405 | 105 | 1510 |
| SQC | 6396 | 189 | 6585 | 4335 | 771 | 5106 | 1405 | 12 | 1417 |
| SCLC | 6396 | 60 | 6456 | 4335 | 4 | 4339 | 1405 | 2 | 1407 |

EUR, European population; EAS, East Asian population; AA, African population. Overall, overall lung cancer; ADE, lung adenocarcinoma; SQC, squamous lung cancer; SCLC, small-cell lung cancer.

Table 2. Variants associated with lung cancer in only ever- or never-smokers.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Strata  | SNP | Position  | Gene | EAFEUR|EAS|AA | Weighted score | OR \_PEUR|EAS|AA | Joint effect size (p-value) | Q |
|  | **Ever-smokers** |
| LUNG | rs62303696\* | 4p12 | GABRA4 | 0.074|0.275|0.028  | 0.94 | 1.17 (2.71x10-7)|1.22 (4.81x10-3)|1.33 (6.08x10-2) | 1.18 (1.22x10-9) | 0.62 |
| LUNG | rs55781567 | 15q25.1 | CHRNA5 | 0.414|0.039|0.299 | 0.99 | 1.31 (5.67X10-69)|0.99 (9.65X10-1)|1.32 (8.51X10-8) | 1.31 (1.66X10-74) | 0.65 |
| SQUAM  | rs17879961# | 22q12.1 | CHEK2 | 0.002|0.000|0.000 | 0.89 | 0.25 (2.93X10-11)| NA|NA | 0.25 (2.93X10-11) | NA |
| SCLC | rs58778970\* | 12q24.33 | Intergenic  | 0.134|0.007|0.190 | 0.92 | 1.33 (1.50x10-7)|0.77 (8.05x10-1)|1.53 (2.40x10-2) | 1.34 (1.58x10-8) | 0.67 |
|  | **Never-smokers** |
| ADE | rs4756620\* | 11p12 | LRRC4C | 0.998|0.977|0.810 | 0.91 | 0.76 (5.62x10-1)|0.57 (1.37x10-8)|0.64 (1.28x10-2) | 0.59 (6.51x10-10) |  0.74 |
| SQC | rs5767055# | 2q34 | IKZF2 | 0.962|0.909|0.917 |  | 1.44 (1.94x10-1)|0.56 (1.51x10-11)|0.71 (6.49x10-1) | 0.61 (1.11x10-9) |  |
| SQC | rs1383429\* | 4q21.21 | LINC01088 | 0.909|0.878|0.492 | 0.97 | 0.73 (8.74x10-2)|0.64 (5.57x10-9)|1.56 (3.13x10-1) | 0.67 (6.44x10-9) | 0.12 |
| SQC | rs968516\* | 9q34.3 | LCNL1 | 0.947|0.966|0.923 | 0.86 | 0.62 (4.10x10-2)|0.36 (8.07x10-10)|0.92 (9.47x10-1) | 0.34 (8.19x10-10) | 0.12 |
|  | **Never-smoking women** |
| Overall  | rs12265047 | 10q25.2 | VTI1A | 0.949|0.701|0.626 | 0.93 | 0.63 (4.64x10-5)|0.77 (4.53x10-13)|0.63 (3.29x10-3) | 0.75 (1.10x10-17) | 0.68 |
| ADE | rs7962469 | 12q13.13 | ACVR1B | 0.684|0.674|0.443 | 0.90 | 1.12 (5.61x10-2)|1.18 (1.63x10-6)|1.74 (3.14x10-3) | 1.18 (3.73x10-8) | 0.03 |

Association results for variants associated with lung cancer in only ever- or never-smokers are summarized in the table. EAF, effective allele frequency. Q indicates the heterogeneity p value. EUR: European population; EAS: East Asian population; AA: African American population. Weighted score indicated the imputation quality score weighted by sample size from the studies. #, known variants identified from previous study but shown to confer risk effect in only ever- or never-smoking group. \*, novel variants identified in this study.

Table 3. Heritability and correlation analysis in EUR and EAS population.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Heritability\_Ever (sd) | Heritability\_Never (sd) | Correlation (P) |
|  |  | **EUR** |  |
| Overall lung cancer | 0.0980 (0.0226) | 0.0262 (0.0232) | 0.6722 (0.0966) |
| Lung adenocarcinoma | 0.1163 (0.0238) | 0.0534 (0.0291) | 0.7919 (0.0111) |
| Lung squamous carcinoma | 0.0874 (0.0222) | 0.0857 (0.0349) | 0.2560 (0.2235) |
|  |  | **EAS** |  |
| Overall lung cancer | 0.2062 (*0.1680*) | 0.1615 (0.0492) | 1.1376 (0.0576) |
| Lung adenocarcinoma | 0.6816 (*0.2466*) | 0.0868 (0.0575) | 0.4682 (0.2306) |
| Lung squamous carcinoma | 1.5381 (*0.2978*) | 0.8710 (0.1090) | 1.0366 (9.7513x10-14) |

LD score regression using LDSC program is not a bounded estimator, so it can produce estimates outside of [-1,1] due to sampling variation. The standard error for estimation of heritability in EAS population is much higher (Italic in the table) in ever-smokers compared with the standard error of the estimates from other strata due to the small sample size in EAS ever-smokers (N=2,527). The estimates were not reliable thus no valid or reliable estimate for correlation between ever- and never-smoking lung cancer in Asian population.