|  |
| --- |
| **Supplementary Table 1**. Demographic characteristics among cases and controls |
| **Variable** | **Total (N)** | **Cases, n (%)**  | **Controls, n (%)**  | **a*p-*value** |
| **Age (years)** | 550 |  |  |  |
| Mean ± SD |  | 59.23 ± 10.28 | 58.49 ± 11.67 | 0.27 |
| Range |  | 28-86 | 27-89 |  |
| **Gender**  | 550 |  |  |  |
| Male  |  | 458 (83.27) | 459 (83.45) |  |
| Female  |  | 92 (16.73) | 91 (16.55) | 0.94 |
| **Smoking status** | 550 |  |  |  |
| Smokers |  | 444 (80.73) | 425 (77.27) |  |
| Non-smokers  |  | 106 (19.27) | 125 (22.73) | 0.18 |
| **Pack years** | 550 |  |  |  |
| Mean ± SD |  | 31.16 ± 28.28 | 20.76 ± 15.95 | ***p*<0.0001** |
| **Histological types** | 550 |  |  |  |
| SQCC |  | 216 (39.27) |  |  |
| ADCC |  | 200 (36.36) |  |  |
| SCLC |  | 127 (23.1) |  |  |
| Others |  | 6 (1.09) |  |  |
| Unknown |  | 1 (0.18) |  |  |
| **Overall survival** | 473 |  |  |  |
| Dead |  | 362(76.53) |  |  |
| Alive |  | 111(23.47) |  |  |
| **TNM staging** | 550 |  |  |  |
| I |  | 5 (0.91) |  |  |
| II |  | 19 (3.45) |  |  |
| III |  | 229 (41.64) |  |  |
| IV |  | 247 (44.91) |  |  |
| Unclassified  |  | 50 (9.09) |  |  |
| **Tumor Size** | 550 |  |  |  |
| Tx |  | 17 (3.09) |  |  |
| T1 |  | 27 (4.91) |  |  |
| T2 |  | 42 (7.64) |  |  |
| T3 |  | 115 (20.91) |  |  |
| T4 |  | 299 (54.36) |  |  |
| Unknown |  | 50 (9.09) |  |  |
| **Lymph Node** | 550 |  |  |  |
| Nx |  | 1 (0.18) |  |  |
| N0 |  | 65 (11.82) |  |  |
| N1 |  | 44 (8) |  |  |
| N2 |  | 233 (42.36) |  |  |
| N3 |  | 153 (27.82) |  |  |
| N4 |  | 4 (0.73) |  |  |
| Unknown |  | 50 (9.09) |  |  |
| **Metastasis** | 550 |  |  |  |
| Mx |  | 2 (0.36) |  |  |
| M0 |  | 254 (46.18) |  |  |
| M1 |  | 246 (44.73) |  |  |
| Unknown  |  | 48 (8.73) |  |  |
| **Performance status** |  |  |  |  |
| KPS (below 60) | 516 | 76(14.73) |  |  |
| KPS (70-80) |  | 285(55.23) |  | **<0.0001** |
| KPS (90-100) |  | 155(30.04) |  | **<0.0001** |
| ECOG (0-1) | 517 | 243(47.00) |  |  |
| ECOG (2) |  | 206(39.85) |  | 0.089 |
| ECOG (3-4) |  | 68(13.15) |  | **<0.0001** |
| **Chemotherapy regimen** |  |  |  |  |
| Pemetrexedcis/carboplatin | 382 | 119(31.15) |  |  |
| Irinotecancis/carboplatin |  | 76(19.89) |  |  |
| Docetaxelcis/carboplatin |  | 75(19.63) |  |  |
| Paclitaxel cis/carboplatin |  | 59(15.45) |  |  |
| Gemcitabine cis/carboplatin |  | 20(5.23) |  |  |
| Gefitinib |  | 12(3.15) |  |  |
| Etoposide |  | 11(2.88) |  |  |
| Ceretinib |  | 1(0.26) |  |  |
| Others  |  | 9(2.36) |  |  |
| Abbreviations: SD = Standard Deviation, n = total number of lung cancer cases or control subjects, a) *p*-values were derived from Pearson Chi – square test except age and pack-years; Student t-test was used for age and pack-years. All *p*-values are two – sided. *p*< 0.05 was considered statistically significant |

|  |
| --- |
| **Supplementary Table 2.** Relationship of different *EPHX1* polymorphisms on the basis of gender |
| **FEMALE** | **MALE** |
| **Genotype****Tyr113His** | **Controls (91) N(%)** | **Cases (92) N(%)** | **AOR (95% CI) a** | ***p* b** | **Controls (459) N(%)** | **Cases (458) N(%)** | **AOR (95% CI) a** | ***p* b** |
| **Co-dominant model** |  |
| Tyr/Tyr | 27 (29.67) | 31 (33.70) | 1.0(Reference) |  | 188 (40.96) | 157 (34.28) | 1.0(Reference) |  |
| Tyr/His | 50 (54.95) | 45 (48.91) | 0.69 (0.34-1.39) | 0.30 | 196 (42.70) | 221 (48.25) | 1.36 (1.02-1.82) | **0.03** |
| His/His | 14 (15.38) | 16 (17.39) | 0.70 (0.25-1.89) | 0.48 | 75 (16.34) | 80 (17.47) | 1.29 (0.88-1.89) | 0.19 |
| **Dominant model** |  |
| Tyr/Tyr | 27 (29.67) | 31 (33.70) | 1.0(Reference) |  | 188 (40.96) | 157 (34.28) | 1.0(Reference) |  |
| Tyr/His+His/His | 64 (70.33) | 61 (66.30) | 0.70 (0.36-1.38) | 0.30 | 271 (59.04) | 301 (65.72) | 1.34 (1.02-1.75) | **0.03** |
| **Recessive model** |  |
| Tyr/Tyr+Tyr/His | 77 (84.62) | 76 (82.61) | 1.0(Reference) |  | 384 (83.66) | 378 (82.53) | 1.0(Reference) |  |
| His/His | 14 (15.38) | 16 (17.39) | 0.90 (0.37-2.19) | 0.81 | 75 (16.34) | 80 (17.47) | 1.07 (0.76-1.52) | 0.69 |
| **Genotype****His139Arg** | **Controls (91) N(%)** | **Cases (92) N(%)** | **AOR (95% CI) a** | ***p* b** | **Controls (459) N(%)** | **Cases (458) N(%)** | **AOR (95% CI) a** | ***p* b** |
| **Co-dominant model** |  |
| His/His | 58 (63.74) | 60 (65.22) | 1.0(Reference) |  | 322 (70.15) | 315 (68.78) | 1.0(Reference) |  |
| His/Arg | 31 (34.06) | 30 (32.61) | 0.97 (0.49-1.90) | 0.92 | 125 (27.24) | 135 (29.47) | 1.10 (0.82-1.47) | 0.52 |
| Arg/Arg | 2 (2.20) | 2 (2.17) | 1.26 (0.17-9.52) | 0.82 | 12 (2.61) | 8 (1.75) | 0.69 (0.28-1.72) | 0.43 |
| **Dominant model** |  |
| His/His | 58 (63.74) | 60 (65.22) | 1.0(Reference) |  | 322 (70.15) | 315 (68.78) | 1.0(Reference) |  |
| His/Arg + Arg/Arg | 33 (36.26) | 32 (34.78) | 1.003 (0.52-1.94) | 0.99 | 137 (29.85) | 143 (31.22) | 1.06 (0.80-1.41) | 0.68 |
| **Recessive model** |  |
| His/His + His/Arg | 89 (97.80) | 90 (97.83) | 1.0(Reference) |  | 447 (97.39) | 450 (98.25) | 1.0(Reference) |  |
| Arg/Arg | 2 (2.20) | 2 (2.17) | 1.46 (0.20-10.78) | 0.71 | 12 (2.61) | 8 (1.75) | 0.66 (0.27-1.63) | 0.37 |
| a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis after adjusting for age, gender. (b)-Two-sided χ2 test for either genotype distribution or allelic frequencies between the cases and controls. The number in bold indicates the significant values in the table. |

|  |
| --- |
| **Supplementary Table 3.** Combinatorial analysis of the risk associated with different polymorphic sites of *EPHX1* gene |
|  | **OVERALL** | **ADCC** |
| **Controls (215) N(%)** | **Cases (219) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (215) N(%)** | **Cases (74) N(%)** | **AOR****(95% CI)a** | ***p*b** |
| Tyr/Tyr+His/His | 130 (60.47) | 116 (52.97) | 1.00 (Reference) |  | 130 (60.47) | 43 (58.11) | 1.00 (Reference) |  |
| (Tyr/His+His/His)+(His/Arg+Arg/Arg) | 85 (39.53) | 103 (47.03) | 1.36 (0.92-1.99) | 0.12 | 85 (39.53) | 31 (41.89) | 0.98 (0.56-1.71) | 0.94 |
| **SQCC** | **SCLC** |
|  | **Controls (215) N(%)** | **Cases (89) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (215) N(%)** | **Cases (53) N(%)** | **AOR****(95% CI) a** | ***p*b** |
| Tyr/Tyr+His/His | 130 (60.47) | 45 (50.56) | 1.00 (Reference) |  | 130 (60.47) | 26 (49.06) | 1.00 (Reference) |  |
| (Tyr/His+His/His)+(His/Arg+Arg/Arg) | 85 (39.53) | 44 (49.44) | 1.66 (0.99-2.79) | **0.05** | 85 (39.53) | 27 (50.94) | 1.85 (0.98-3.49) | **0.05** |
|  | **SMOKERS** | **NON-SMOKERS** |
| **Controls (163) N(%)** | **Cases (175) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (52) N(%)** | **Cases (44) N(%)** | **AOR****(95% CI)a** | ***p*b** |
| Tyr/Tyr+His/His | 106 (65.03) | 91 (52.00) | 1.00 (Reference) |  | 24 (46.15) | 25 (56.82) | 1.00 (Reference) |  |
| (Tyr/His+His/His)+(His/Arg+Arg/Arg) | 57 (34.97) | 84 (48.00) | 1.67 (1.07-2.60) | **0.02** | 28 (53.85) | 19 (43.18) | 0.62 (0.27-1.44) | 0.27 |
| **LIGHT SMOKERS** | **HEAVY SMOKERS** |
|  | **Controls (117) N(%)** | **Cases (91) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (46) N(%)** | **Cases (84) N(%)** | **AOR****(95% CI) a** | ***p*b** |
| Tyr/Tyr+His/His | 78 (66.67) | 49 (53.85) | 1.00 (Reference) |  | 28 (60.87) | 42 (50.00) | 1.00 (Reference) |  |
| (Tyr/His+His/His)+(His/Arg+Arg/Arg) | 39 (33.33) | 42 (46.15) | 1.62 (0.91-2.88) | 0.10 | 18 (39.13) | 42 (50.00) | 1.64 (0.78-3.45) | 0.19 |
|  | **Males** | **Females** |
| **Controls (179) N(%)** | **Cases (182) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (36) N(%)** | **Cases (37) N(%)** | **AOR****(95% CI)a** | ***p*b** |
| Tyr/Tyr+His/His | 115 (64.25) | 98 (53.85) | 1.00 (Reference) |  | 15 (41.67) | 18 (48.65) | 1.00 (Reference) |  |
| (Tyr/His+His/His)+(His/Arg+Arg/Arg) | 64 (35.75) | 84 (46.15) | 1.54 (1.01-2.35) | **0.04** | 21 (58.33) | 19 (51.35) | 0.68 (0.25-1.85) | 0.45 |
| a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis after adjusting for age, gender and smoking status. (b)-Two-sided χ2 test for either genotype distribution or allelic frequencies between the cases and controls. The number in bold indicates the significant values in the table. |

|  |
| --- |
| **Supplementary Table 4.** Combined effects of *Tyr113His* (rs2854455) and *His139Arg* (rs2234922) risk alleles and risk of lung cancer |
| **OVERALL** | **ADCC** |
| **Number of Risk alleles$** | **Control (550), N(%)** | **Cases (550), N(%)** | **AOR (95% CI)** | **p-value** | **Control (550), N(%)** | **Cases (200), N(%)** | **AOR (95% CI)** | **p-value** |
| 0 | 130 (23.64) | 116 (21.10) | 1.0 (Reference) |  | 130 (23.64) | 43 (21.50) | 1.0 (Reference) |  |
| 1 | 260 (47.27) | 251 (45.64) | 1.07 (0.79-1.45) | 0.66 | 260 (47.27) | 97 (48.50) | 1.10 (0.72-1.70) | 0.65 |
| 2 | 132 (24.00) | 157 (28.54) | 1.32 (0.94-1.86) | 0.11 | 132 (24.00) | 53 (26.50) | 1.16 (0.71-1.87) | 0.55 |
| 3 | 28 (5.09) | 26 (4.72) | 1.05 (0.58-1.90) | 0.88 | 28 (5.09) | 7 (3.50) | 0.57 (0.22-1.47) | 0.25 |
| **SQCC** | **SCLC** |
| **Number of Risk alleles$** | **Control (550), N(%)** | **Cases (216), N(%)** | **AOR (95% CI)** | **p-value** | **Control (550), N(%)** | **Cases (127), N(%)** | **AOR (95% CI)** | **p-value** |
| 0 | 130 (23.64) | 45 (20.83) | 1.0 (Reference) |  | 130 (23.64) | 26 (20.47) | 1.0 (Reference) |  |
| 1 | 260 (47.27) | 94 (43.52) | 1.05 (0.69-1.60) | 0.82 | 260 (47.27) | 57 (44.88) | 1.09 (0.64-1.89) | 0.74 |
| 2 | 132 (24.00) | 66 (30.56) | 1.50 (0.94-2.38) | 0.08 | 132 (24.00) | 36 (28.35) | 1.50 (0.84-2.67) | 0.17 |
| 3 | 28 (5.09) | 11 (5.09) | 1.36 (0.61-3.04) | 0.46 | 28 (5.09) | 8 (6.30) | 1.88 (0.72-4.93) | 0.20 |
| **SMOKERS** | **NON-SMOKERS** |
| **Number of Risk alleles$** | **Control (425), N(%)** | **Cases (444), N(%)** | **AOR (95% CI)** | **p-value** | **Control (125), N(%)** | **Cases (106), N(%)** | **AOR (95% CI)** | **p-value** |
| 0 | 106 (24.94) | 91 (20.50) | 1.0 (Reference) |  | 24 (19.20) | 25 (23.59) | 1.0 (Reference) |  |
| 1 | 203 (47.76) | 203 (45.72) | 1.15 (0.81-1.62) | 0.42 | 57 (45.60) | 48 (45.28) | 0.76 (0.37-1.56) | 0.45 |
| 2 | 100 (23.53) | 128 (28.83) | 1.45 (0.98-2.13) | 0.06 | 32 (25.60) | 29 (27.36) | 0.92 (0.43-1.97) | 0.83 |
| 3 | 16 (3.76) | 22 (4.95) | 1.55 (0.76-3.15) | 0.23 | 12 (9.60) | 4 (3.77) | 0.32 (0.09-1.15) | 0.08 |
| **LIGHT-SMOKERS** | **HEAVY-SMOKERS** |
| **Number of Risk alleles$** | **Control (300), N(%)** | **Cases (248), N(%)** | **AOR (95% CI)** | **p-value** | **Control (125), N(%)** | **Cases (196), N(%)** | **AOR (95% CI)** | **p-value** |
| 0 | 78 (26.00) | 49 (19.76) | 1.0 (Reference) |  | 28 (22.40) | 42 (21.43) | 1.0 (Reference) |  |
| 1 | 141 (47.00) | 112 (45.16) | 1.25 (0.80-1.95) | 0.31 | 62 (49.60) | 91 (46.43) | 0.97 (0.54-1.74) | 0.92 |
| 2 | 71 (23.67) | 72 (29.03) | 1.47 (0.90-2.42) | 0.12 | 29 (23.20) | 56 (28.57) | 1.36 (0.70-2.63) | 0.36 |
| 3 | 10 (3.33) | 15 (6.05) | 2.39 (0.99-5.81) | 0.05 | 6 (4.80) | 7 (3.57) | 0.74 (0.21-2.56) | 0.63 |
| **MALES** | **FEMALES** |
| **Number of Risk alleles$** | **Control (459), N(%)** | **Cases (458), N(%)** | **AOR (95% CI)** | **p-value** | **Control (91), N(%)** | **Cases (92), N(%)** | **AOR (95% CI)** | **p-value** |
| 0 | 115 (25.05) | 98 (21.40) | 1.0 (Reference) |  | 15 (16.48) | 18 (19.57) | 1.0 (Reference) |  |
| 1 | 215 (46.84) | 209 (45.63) | 1.14 (0.82-1.59) | 0.42 | 45 (49.45) | 42 (45.65) | 0.72 (0.30-1.69) | 0.44 |
| 2 | 107 (23.31) | 130 (28.38) | 1.44 (0.99-2.09) | 0.05 | 25 (27.47) | 27 (29.35) | 0.79 (0.30-2.05) | 0.63 |
| 3 | 22 (4.79) | 21 (4.59) | 1.14 (0.59-2.21) | 0.69 | 6 (6.59) | 5 (5.43) | 0.63 (0.15-2.73) | 0.54 |
| $ 0 risk allele: Tyr/Tyr + His/His; 1 risk allele: Tyr/Tyr + His/Arg, Tyr/His + His/His; 2 risk allele: Tyr/Tyr + Arg/Arg, His/His + His/His, Tyr/His + His/Arg; 3 risk allele: Tyr/His + Arg/Arg, His/His + His/Arg |

|  |
| --- |
| **Supplementary Table 5.** Genotypic and allelic distribution of the *EPHX1* genetic variant and its association with risk of lung cancer as Surrogate of Enzyme Activity |
|  | **OVERALL** | **ADCC** |
| **Controls (550) N(%)** | **Cases (550) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (550) N(%)** | **Cases (200) N(%)** | **AOR****(95% CI)a** | ***p*b** |
| Very low (VL) | 64 (11.64) | 73 (13.27) | 1.00 (Reference) |  | 64 (11.64) | 27 (13.50) | 1.00 (Reference) |  |
| Low (L) | 211 (38.36) | 209 (38.00) | 0.87 (0.59-1.28) | 0.48 | 211 (38.36) | 73 (36.50) | 0.80 (0.47-1.36) | 0.41 |
| Intermediate (I) | 187 (34.00) | 193 (35.09) | 0.90 (0.61-1.34) | 0.61 | 187 (34.00) | 67 (33.50) | 0.84 (0.49-1.45) | 0.54 |
| High (H) | 88 (16.00) | 75 (13.64) | 0.73 (0.46-1.16) | 0.19 | 88 (16.00) | 33 (16.50) | 0.87 (0.46-1.62) | 0.66 |
| Combined (L+I+H) | 486 (88.36) | 477 (86.73) | 0.86 (0.60-1.23) | 0.41 | 486 (88.36) | 173 (86.50) | 0.82 (0.50-1.34) | 0.43 |
| **SQCC** | **SCLC** |
|  | **Controls (550) N(%)** | **Cases (216) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (550) N(%)** | **Cases (127) N(%)** | **AOR****(95% CI) a** | ***p*b** |
| Very low (VL) | 64 (11.64) | 30 (13.89) | 1.00 (Reference) |  | 64 (11.64) | 15 (11.81) | 1.00 (Reference) |  |
| Low (L) | 211 (38.36) | 84 (38.89) | 0.88 (0.52-1.48) | 0.63 | 211 (38.36) | 50 (39.37) | 0.88 (0.45-1.72) | 0.71 |
| Intermediate (I) | 187 (34.00) | 78 (36.11) | 0.89 (0.53-1.50) | 0.66 | 187 (34.00) | 45 (35.43) | 0.97 (0.50-1.87) | 0.92 |
| High (H) | 88 (16.00) | 24 (11.11) | 0.58 (0.30-1.10) | 0.10 | 88 (16.00) | 17 (13.39) | 0.77 (0.35-1.69) | 0.51 |
| Combined (L+I+H) | 486 (88.36) | 186 (86.11) | 0.82 (0.51-1.33) | 0.43 | 486 (88.36) | 112 (88.19) | 0.89 (0.48-1.65) | 0.72 |
|  | **SMOKERS** | **NON-SMOKERS** |
| **Controls (425) N(%)** | **Cases (444) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (125) N(%)** | **Cases (106) N(%)** | **AOR****(95% CI)a** | ***p*b** |
| Very low (VL) | 50 (11.76) | 61 (13.74) | 1.00 (Reference) |  | 14 (11.20) | 12 (11.32) | 1.00 (Reference) |  |
| Low (L) | 158 (37.18) | 175 (39.41) | 0.93 (0.60-1.44) | 0.74 | 53 (42.40) | 34 (32.08) | 0.67 (0.27-1.69) | 0.40 |
| Intermediate (I) | 147 (34.59) | 153 (34.46) | 0.88 (0.56-1.37) | 0.57 | 40 (32.00) | 40 (37.73) | 1.20 (0.48-2.98) | 0.69 |
| High (H) | 70 (16.47) | 55 (12.39) | 0.66 (0.39-1.12) | 0.12 | 18 (14.40) | 20 (18.87) | 1.36 (0.49-3.77) | 0.56 |
| Combined (L+I+H) | 375 (88.24) | 383 (86.26) | 0.86 (0.57-1.29) | 0.47 | 111 (88.80) | 94 (88.68) | 0.95 (0.41-2.20) | 0.90 |
| **LIGHT SMOKERS** | **HEAVY SMOKERS** |
|  | **Controls (300) N(%)** | **Cases (248) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (125) N(%)** | **Cases (196) N(%)** | **AOR****(95% CI) a** | ***p*b** |
| Very low (VL) | 35 (11.67) | 40 (16.13) | 1.00 (Reference) |  | 15 (12.00) | 21 (10.71) | 1.00 (Reference) |  |
| Low (L) | 108 (36.00) | 94 (37.90) | 0.83 (0.48-1.45) | 0.52 | 50 (40.00) | 81 (41.33) | 1.12 (0.52-2.38) | 0.77 |
| Intermediate (I) | 107 (35.67) | 76 (30.65) | 0.65 (0.37-1.14) | 0.13 | 40 (32.00) | 77 (39.29) | 1.35 (0.62-2.90) | 0.45 |
| High (H) | 50 (16.66) | 38 (15.32) | 0.71 (0.38-1.36) | 0.30 | 20 (16.00) | 17 (8.67) | 0.56 (0.22-1.45) | 0.23 |
| Combined (L+I+H) | 265 (88.33) | 208 (83.87) | 0.74 (0.45-1.23) | 0.24 | 110 (88.00) | 175 (89.29) | 1.11 (0.54-2.25) | 0.78 |
| **Males** | **Females** |
|  | **Controls (459) N(%)** | **Cases (458) N(%)** | **AOR****(95% CI) a** | ***p*b** | **Controls (91) N(%)** | **Cases (92) N(%)** | **AOR****(95% CI)a** | ***p*b** |
| Very low (VL) | 55 (11.98) | 61 (13.32) | 1.00 (Reference) |  | 9 (9.89) | 12 (13.04) | 1.00 (Reference) |  |
| Low (L) | 172 (37.47) | 175 (38.21) | 0.91 (0.60-1.39) | 0.67 | 39 (42.86) | 34 (36.96) | 0.79 (0.25-2.44) | 0.68 |
| Intermediate (I) | 157 (34.20) | 161 (35.15) | 0.92 (0.60-1.41) | 0.71 | 30 (32.97) | 32 (34.78) | 1.10 (0.36-3.41) | 0.86 |
| High (H) | 75 (16.34) | 61 (13.32) | 0.71 (0.43-1.17) | 0.17 | 13 (14.28) | 14 (15.22) | 1.44 (0.37-5.56) | 0.60 |
| Combined (L+I+H) | 404 (88.02) | 397 (86.68) | 0.88 (0.60-1.30) | 0.53 | 82 (90.11) | 80 (86.96) | 0.94 (0.33-2.66) | 0.90 |
| a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis after adjusting for age, gender and smoking status. (b)-Two-sided χ2 test for either genotype distribution or allelic frequencies between the cases and controls. The number in bold indicates the significant values in the table. |

|  |
| --- |
| **Supplementary Table 6.** Haplotype frequencies and lung cancer risk associated with *EPHX1* haplotypes |
| **Haplotype** | **Case frequency (%)** | **Control frequency (%)** | **OR (95% CI)$** | **p-value#** |
| *EPHX1 Tyr113His* and *EPHX1 His139Arg* |
| -His113-His139 | 407.38 (0.37) | 377.39 (0.34) | 1.13 (0.95-1.34) | 0.18 |
| -His113-Arg139 | 50.62 (0.05) | 46.61 (0.04) | 1.09 (0.73-1.64) | 0.68 |
| -Tyr113-His139 | 507.62 (0.46) | 538.61 (0.49) | 0.89 (0.75-1.06) | 0.18 |
| -Tyr113-Arg139 | 134.38 (0.12) | 137.39 (0.13) | 0.97 (0.76-1.26) | 0.84 |
| **$**ORs, 95% CIs and their corresponding p-values were calculated by logistic regression analysis.**#**Two- sided χ2 test for either genotype distribution or allelic frequencies between the cases and controls.OR: Odds ratio. |

|  |
| --- |
| **Supplementary Table 7.** Multifactor dimensionality reduction analysis showing interactions of *EPHX1* variants with lung cancer risk |
| **Number of risk factors** | **Best model** | **Cross-validation consistency** | **Prediction error** | **p-value** |
| 1 | *EPHX1 Tyr113His*  | 10/10 | 0.48 | 0.72 |
| 2 | *EPHX1 Tyr113His, EPHX1 His139Arg* | 10/10 | 0.42 | 0.86 |

|  |
| --- |
| **Supplementary Table 8.** Relationship of different genotypes with the clinical-pathological parameters |
|  | **CLINICAL STAGE** | **PRIMARY TUMOR EXTENSION** |
|  | **III** | **IV** |  | **T3** | **T4** |  |  |
| **Genotype****Tyr113His** | **N=229 (%)** | **N=247 (%)** | **AOR (95% CI) a** | ***p* b** | **N=115 (%)** | **N=299 (%)** | **AOR (95% CI) a** | ***p* b** |
| **Co-dominant model** |  |  |  |  |  |
| Tyr/Tyr | 75 (32.75) | 85 (34.41) | 1.0(Reference) |  | 38 (33.04) | 103 (34.45) | 1.0(Reference) |  |
| Tyr/His | 117 (51.09) | 116 (46.96) | 0.93 (0.61-1.43) | 0.76 | 58 (50.43) | 145 (48.49) | 0.90 (0.56-1.47) | 0.69 |
| His/His | 37 (16.16) | 46 (18.62) | 1.22 (0.70-2.13) | 0.48 | 19 (16.52) | 51 (17.06) | 0.81 (0.41-1.59) | 0.54 |
| **Dominant model** |  |  |  |  |
| Tyr/Tyr | 75 (32.75) | 85 (34.41) | 1.0(Reference) |  | 38 (33.04) | 103 (34.45) | 1.0(Reference) |  |
| Tyr/His+His/His | 154 (67.25) | 162 (65.59) | 1.01 (0.67-1.50) | 0.97 | 77 (66.96) | 196 (65.55) | 0.92 (0.58-1.45) | 0.72 |
| **Recessive model** |  |  |  |  |
| Tyr/Tyr+Tyr/His | 192 (83.84) | 201 (81.38) | 1.0(Reference) |  | 96 (83.48) | 248 (82.94) | 1.0(Reference) |  |
| His/His | 37 (16.16) | 46 (18.62) | 1.23 (0.75-2.01) | 0.41 | 19 (16.52) | 51 (17.06) | 0.98 (0.55-1.76) | 0.95 |
| **Genotype****His139Arg** | **N=229 (%)** | **N=247 (%)** | **AOR** **(95% CI)a** | ***p*b** | **N=115 (%)** | **N=299 (%)** | **AOR** **(95% CI)a** | ***p*b** |
| **Co-dominant model** |  |
| His/His | 162 (70.74) | 165 (66.80) | 1.0(Reference) |  | 84 (73.04) | 201 (67.22) | 1.0(Reference) |  |
| His/Arg | 62 (27.08) | 77 (31.18) | 1.18 (0.78-1.79) | 0.43 | 30 (26.09) | 89 (29.77) | 1.27 (0.78-2.07) | 0.34 |
| Arg/Arg | 5 (2.18) | 5 (2.02) | 0.81 (0.21-3.02) | 0.75 | 1 (0.87) | 9 (3.01) | 3.69 (0.46-29.76) | 0.22 |
| **Dominant model** |  |
| His/His | 162 (70.74) | 165 (66.80) | 1.0(Reference) |  | 84 (73.04) | 201 (67.22) | 1.0(Reference) |  |
| His/Arg + Arg/Arg | 67 (29.26) | 82 (33.20) | 1.15 (0.77-1.73) | 0.50 | 31 (26.96) | 98 (32.78) | 1.35 (0.83-2.18) | 0.22 |
| **Recessive model** |  |
| His/His + His/Arg | 224 (97.82) | 242 (97.98) | 1.0(Reference) |  | 114 (99.13) | 290 (96.99) | 1.0(Reference) |  |
| Arg/Arg | 5 (2.18) | 5 (2.02) | 0.73 (0.19-2.82) | 0.65 | 1 (0.87) | 9 (3.01) | 3.34 (0.42-26.87) | 0.26 |
|  | **LYMPH NODE INVASION** | **METASTASIS** |
|  | **Nx+N0+N1** | **N2+N3+N4** |  |  | **No (M0)** | **Yes (M1)** |  |  |
| **Genotype****Tyr113His** | **N=110 (%)** | **N=390 (%)** | **AOR (95% CI) a** | ***p* b** | **N=256 (%)** | **N=246 (%)** | **AOR (95% CI) a** | ***p* b** |
| **Co-dominant model** |  |  |  |  |
| Tyr/Tyr | 46 (41.82) | 126 (32.31) | 1.0(Reference) |  | 90 (35.16) | 82 (33.33) | 1.0(Reference) |  |
| Tyr/His | 50 (45.45) | 191 (48.97) | 1.40 (0.88-2.22) | 0.15 | 125 (48.83) | 118 (47.97) | 1.14 (0.76-1.72) | 0.52 |
| His/His | 14 (12.73) | 73 (18.72) | 1.94 (0.99-3.80) | **0.04** | 41 (16.01) | 46 (18.70) | 1.41 (0.82-2.43) | 0.21 |
| **Dominant model** |  |  |  |  |
| Tyr/Tyr | 46 (41.82) | 126 (32.31) | 1.0(Reference) |  | 90 (35.16) | 82 (33.33) | 1.0(Reference) |  |
| Tyr/His+His/His | 64 (58.18) | 264 (67.69) | 1.51 (0.98-2.34) | 0.06 | 166 (64.84) | 164 (66.67) | 1.22 (0.83-1.79) | 0.32 |
| **Recessive model** |  |  |  |  |
| Tyr/Tyr+Tyr/His | 96 (87.27) | 317 (81.28) | 1.0(Reference) |  | 215 (83.99) | 200 (81.30) | 1.0(Reference) |  |
| His/His | 14 (12.73) | 73 (18.72) | 1.57 (0.85-2.91) | 0.15 | 41 (16.01) | 46 (18.70) | 1.27 (0.78-2.05) | 0.33 |
| **Genotype****His139Arg** | **N=110 (%)** | **N=390 (%)** | **AOR** **(95% CI)a** | ***p*b** | **N=256 (%)** | **N=246 (%)** | **AOR** **(95% CI)a** | ***p*b** |
| **Co-dominant model** |  |
| His/His | 72 (65.45) | 268 (68.72) | 1.0(Reference) |  | 177 (69.14) | 165 (67.07) | 1.0(Reference) |  |
| His/Arg | 35 (31.82) | 115 (29.49) | 0.89 (0.56-1.41) | 0.63 | 73 (28.52) | 77 (31.30) | 1.10 (0.73-1.64) | 0.65 |
| Arg/Arg | 3 (2.73) | 7 (1.79) | 0.60 (0.15-2.38) | 0.46 | 6 (2.34) | 4 (1.63) | 0.58 (0.15-2.23) | 0.43 |
| **Dominant model** |  |
| His/His | 72 (65.45) | 268 (68.72) | 1.0(Reference) |  | 177 (69.14) | 165 (67.07) | 1.0(Reference) |  |
| His/Arg + Arg/Arg | 38 (34.55) | 122 (31.28) | 0.87 (0.55-1.36) | 0.53 | 79 (30.86) | 81 (32.93) | 1.05 (0.71-1.55) | 0.80 |
| **Recessive model** |  |
| His/His + His/Arg | 107 (97.27) | 383 (98.21) | 1.0(Reference) |  | 250 (97.66) | 242 (98.37) | 1.0(Reference) |  |
| Arg/Arg | 3 (2.73) | 7 (1.79) | 0.63 (0.16-2.50) | 0.51 | 6 (2.34) | 4 (1.63) | 0.53 (0.13-2.09) | 0.36 |
| a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis after adjusting for age, gender. (b)-Two-sided χ2 test for either genotype distribution or allelic frequencies between the cases and controls. The number in bold indicates the significant values in the table. |

|  |
| --- |
| **Supplementary Table 9.** Relationship of genotypic distribution with clinic-pathological parameter on the basis of response to chemotherapy |
|  | **Response of chemotherapy** | **AOR (95% CI) a** | ***p* b** |
| **Genotype****Tyr113His** | **CR+PR****(118), n (%)** | **SD+PD****(148), n (%)** |
| **Co-dominant model** |
| Tyr/Tyr | 40 (33.90) | 58 (39.19) | 1.0(Reference) |  |
| Tyr/His | 51 (43.22) | 61 (41.22) | 1.15 (0.66-2.00) | 0.63 |
| His/His | 27 (22.88) | 29 (19.59) | 1.32 (0.67-2.58) | 0.42 |
| **Dominant model** |
| Tyr/Tyr | 40 (33.90) | 58 (39.19) | 1.0(Reference) |  |
| Tyr/His+His/His | 78 (66.10) | 90 (60.81) | 1.21 (0.73-2.02) | 0.46 |
| **Recessive model** |
| Tyr/Tyr+Tyr/His | 91 (77.12) | 119 (80.41) | 1.0(Reference) |  |
| His/His | 27 (22.88) | 29 (19.59) | 1.20 (0.66-2.18) | 0.54 |
| **Genotype****His139Arg** | **CR+PR****(118), n (%)** | **SD+PD****(148), n (%)** | **AOR (95% CI) a** | ***p* b** |
| **Co-dominant model** |
| His/His | 84 (71.19) | 101 (68.24) | 1.0(Reference) |  |
| His/Arg | 32 (27.12) | 44 (29.73) | 0.91 (0.52-1.58) | 0.74 |
| Arg/Arg | 2 (1.69) | 3 (2.03) | 0.90 (0.14-5.71) | 0.91 |
| **Dominant model** |
| His/His | 84 (71.19) | 101 (68.24) | 1.0(Reference) |  |
| His/Arg + Arg/Arg | 34 (28.81) | 47 (31.76) | 0.89 (0.52-1.53) | 0.67 |
| **Recessive model** |
| His/His + His/Arg | 116 (98.31) | 145 (97.97) | 1.0(Reference) |  |
| Arg/Arg | 2 (1.69) | 3 (2.03) | 0.86 (0.14-5.31) | 0.87 |
| a)-Adjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis after adjusting for age, gender. (b)-Two-sided χ2 test for either genotype distribution or allelic frequencies between the cases and controls. The number in bold indicates the significant values in the table |