

## **Supplementary Materials**

### **Supplementary methods: Description of the LDCT screening cohorts participating in this study**

#### **PanCan: Pan-Canadian Early Detection of Lung Cancer Study**

This study prospectively recruited 2537 subjects between September 2008, and December, 2010 [1]. Participants had at least a 2% 6-year risk of lung cancer as estimated by the PanCan model, a precursor to the validated PLCom2012 model. The eligible subjects were ever smokers, aged 50-75 years who had no previous history of invasive cancer. The subjects were recruited at eight centers across Canada (Vancouver, Calgary, Hamilton, Toronto, Ottawa, Quebec, Halifax, and St. John's). All study participants completed a detailed questionnaire, spirometry, collection of blood specimens for biomarker measurement and LDCT at baseline. Participants with abnormal LDCT results (the presence of any non-calcified non-perifissural pulmonary nodule or a nodule with non-solid area of at least 1mm diameter) were followed up every 3-24 months. A nodule was considered benign if it showed a benign calcification pattern (e.g., fully calcified, popcorn calcification) or if the size and density of was unchanged for longer than 2 years.

#### **UKLS: The UK Lung Cancer Pilot Screening Trial**

The UKLS is a randomized controlled trial of LDCT screening in individuals aged 50-75 years with at least 4.5% risk of developing lung cancer in 5 years based on the Liverpool Lung Project (LLPv2) risk model [2, 3]. The subjects with any comorbidity that unequivocally contraindicated either screening or treatment were excluded. There were 4055 individuals recruited in two geographical areas, covering six primary care trusts around Liverpool and Cambridge, from 2011 to 2013. The participants were randomized to receive either LDCT or usual care, 1994 of which underwent LDCT scans. Most participants had their lung function measured, provided blood samples, and completed detailed epidemiological, psychosocial and health economics questionnaires at baseline. The LDCT subjects received a baseline scan and nodules

were assessed by two local radiologists and placed in one of four pre-defined nodule categories: Category 4 (Solid: volume  $>500\text{mm}^3$  or diameter  $>10\text{mm}$ ; Part solid: solid component- volume  $>500\text{mm}^3$ ); Category 3 (Solid: volume 50 -  $500\text{mm}^3$  or diameter 5 - 9.9mm; Part solid: non-solid component diameter  $>5\text{mm}$ , solid component :15 -  $500\text{mm}^3$  or diameter 3 - 9.9mm; Non-solid: diameter  $\geq 5\text{mm}$ ), Category 2 (Solid: volume 15 -  $49\text{mm}^3$  or diameter 3 - 4.9mm; Part solid: solid component volume  $<15\text{mm}^3$  or diameter  $< 3\text{mm}$ ; Non solid diameter 3 - 4.9mm); Category 1 (Benign nodule or volume  $<15\text{mm}^3$  or diameter  $<3\text{mm}$  or features of an intrapulmonary lymph node). Consensus nodule category was assigned following central reading at the Royal Brompton Hospital, with a read for arbitration when necessary. Category 4 nodules were immediately referred for work-up and clinical management. Category 3 nodules identified in the baseline scan were re-analyzed in follow-up CT scans at three and twelve months; Category 2 nodules at twelve months only. Category 1 nodules were considered normal and had no further follow-up. Growth of nodules was based on their characteristics and volume doubling time (VDT); i.e. VDT  $<400$  days or new solid component of non-solid nodule was classified as growth in the UKLS trial and were referred to the trial participating MDT clinic for work-up and clinical management.

### **IELCAP-Toronto: International Early Lung Cancer Action Program – Toronto**

This single arm early detection of lung cancer research study at Princess Margaret Cancer Centre, Toronto, Canada joined the international Early Lung Cancer Action Program (I-ELCAP) in 2003 [4, 5]. Ever smokers of more than 10 pack-years, aged 50 years or older were eligible for LDCT screening. A total of 4782 individuals have been enrolled in the Greater Toronto Area, and 4161 of those have the clinic-epidemiological data needed to be part of INTEGRAL. Participants were administered a LDCT scan along with a standard study questionnaire at baseline. Blood samples were systematically collected at baseline since 2006. In this study, a LDCT results was considered to be normal if no nodules of the size of 5 mm or

larger was found. In case of a positive LDCT result, an indeterminate noncalcified solid or part-solid nodule 5 mm or larger, or a nonsolid nodule 8 mm or larger was detected. The abnormal scans were followed up at 1-2 years depending on nodule size and shape. The location, size, consistency, calcifications and spiculations were determined for each nodule.

### **P-IELCAP: Pamplona International Early Lung Cancer Action Program**

P-IELCAP is an ongoing lung cancer screening program in Spain which began recruitment in September 2000 [6]. It is part of the I-ELCAP consortium. As of December 2019, the program had enrolled 3825 asymptomatic subjects older than 40 years of age, including former or current smokers ( $\geq 10$  pack-years), cancer-free 5 years prior to enrollment. The subjects completed LDCT, spirometry, blood collection and a health questionnaire at baseline and during follow up. Subjects with a positive baseline LDCT (noncalcified nodules  $\geq 5$  or 6 mm depending on the year of enrollment), had imaging follow up at 1-3 months. Subjects with negative scans or small pulmonary nodules at baseline had follow up LDCTs at annual intervals.

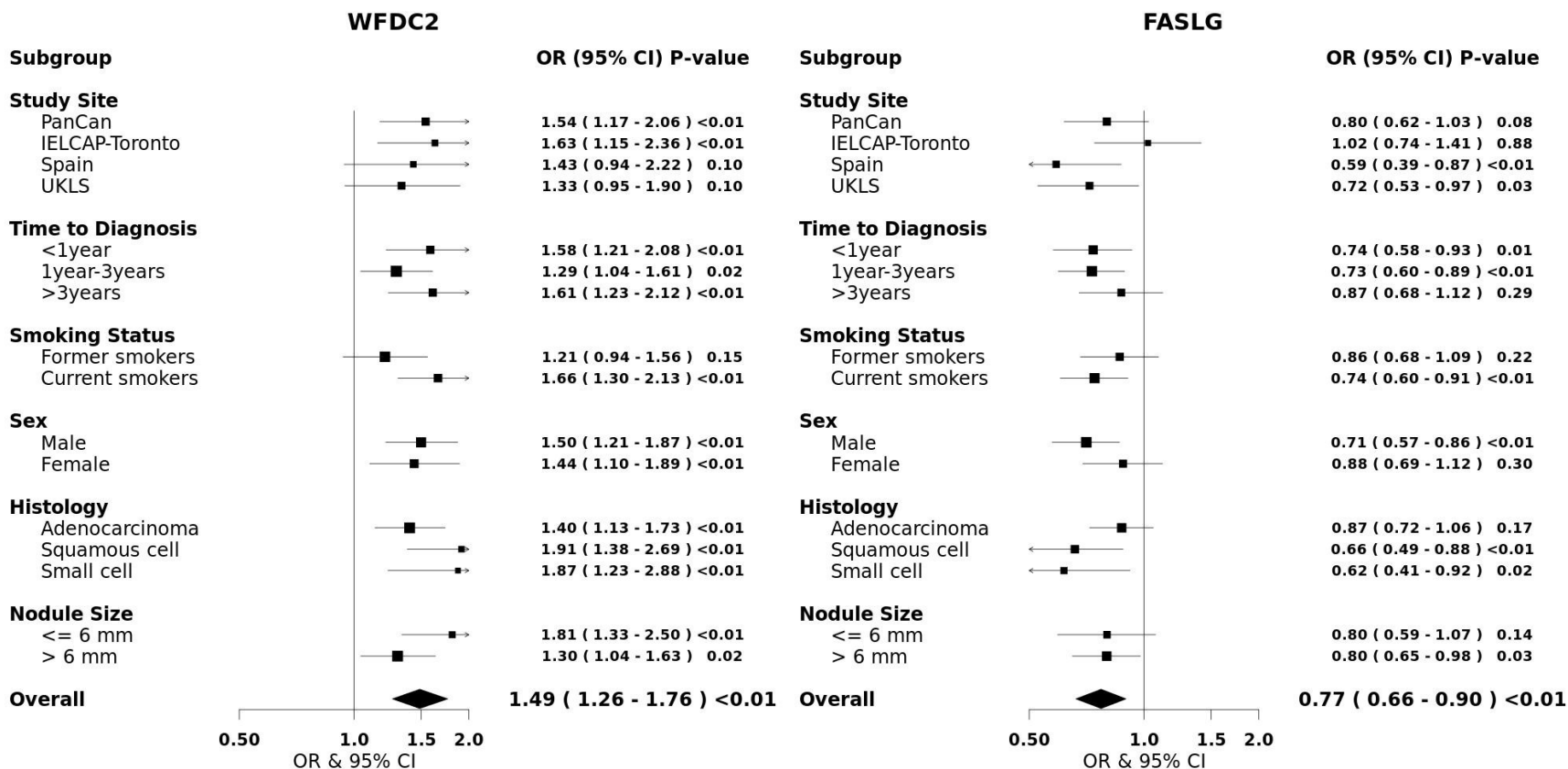
**Supplementary Table 2: Summary of the top 11 informative circulating protein markers**

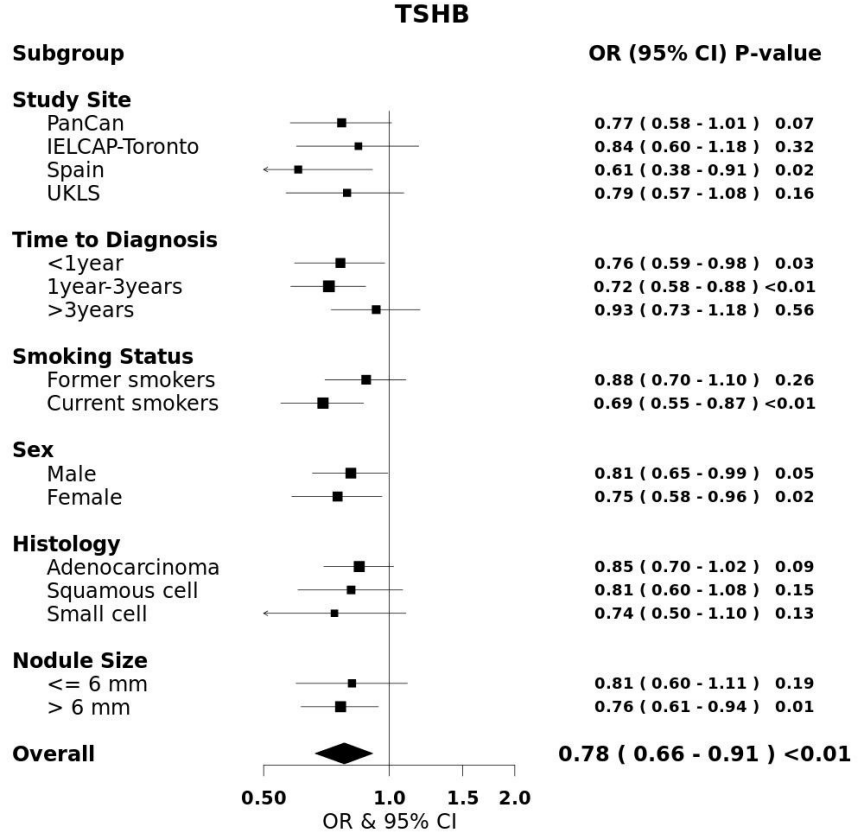
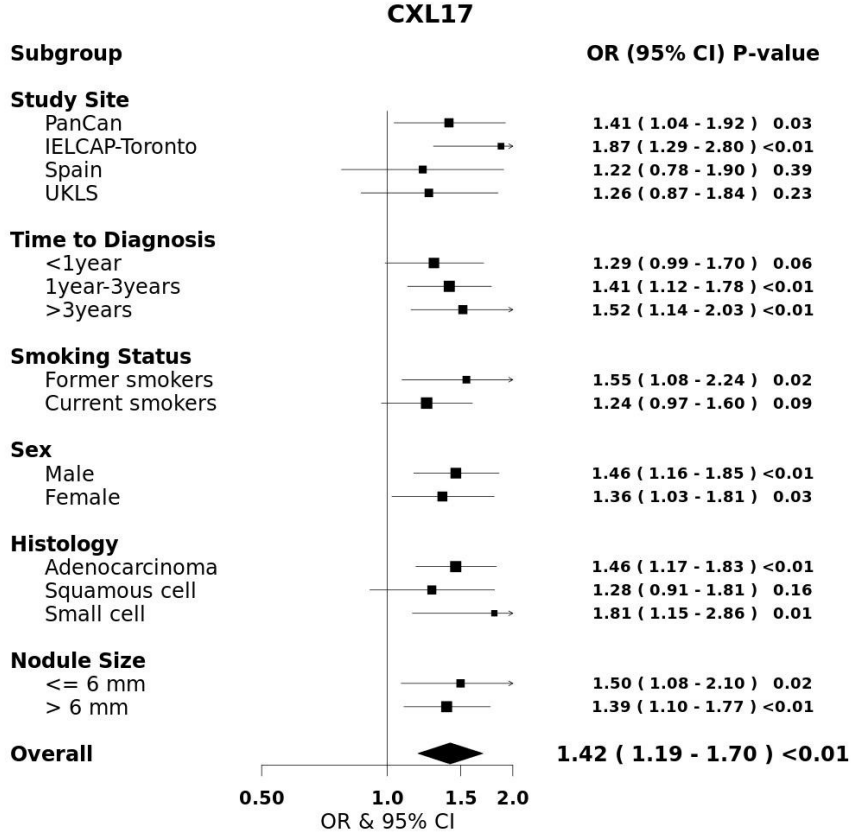
Gene name	Full name (Ensembl Gene ID)	Cytoband	Expression in selected tissues focuses on lung and immune system			
			Normal tissues: Human Protein Atlas (HPA) [7] and Proteomics DB (PDB) [8-10]	Protein expression	mRNA expression-HPA	Cancer tissues: Human Protein Atlas (HPA)[7]
					Protein	mRNA
WFDC2	WAP four-disulfide core domain 2 (ENSG00000101443)	20q13	<b>HPA:</b> Bronchus <sup>H</sup> <b>PDB:</b> Lung	Lung	Lung cancer	Lung cancer
FASLG	Fas ligand (ENSG00000117560)	1q24	<b>HPA:</b> Bone marrow <sup>M</sup> , Lymph node <sup>L</sup> <b>PDB:</b> Data not available	Lymph node, Lung	No expression	Lung cancer
CXCL17	C-X-C motif chemokine ligand 17 (ENSG00000189377)	19q13	<b>HPA:</b> Data not available <b>PDB:</b> Lung	Lung	Data not available	Lung cancer
TSHB	Thyroid stimulating hormone subunit beta (ENSG00000134200)	1q13	<b>HPA:</b> None <b>PDB:</b> None	Lung	No expression	Lung cancer
TRAIL-R2 (TNFRSF10B)	TNF receptor superfamily member 10b (ENSG00000120889)	8p21	<b>HPA:</b> Bronchus <sup>H</sup> , Lung <sup>H</sup> , Lymph node <sup>M</sup> <b>PDB:</b> Lung, Blood platelet	Lung, Bone marrow, Lymph node	Lung cancer	Lung cancer, Blood platelet
SCGB3A2	Secretoglobin family 3a member 2 (ENSG00000164265)	5q32	<b>HPA:</b> Data not available <b>PDB:</b> Lung	Lung	Data not available	Lung cancer
FCRL5	Fc receptor like 5 (ENSG00000143297)	1q23	<b>HPA:</b> Data not available <b>PDB:</b> Lung	Lymph node, Bone marrow, Lung	Data not available	Lung cancer

KRT19	Keratin 19 (ENSG00000171345)	17q21	<b>HPA:</b> Bronchus <sup>H</sup> , Lung <sup>H</sup> <b>PDB:</b> Lung, Lymph node, Cytotoxic T-lymphocyte, Blood platelet	Lung	Lung cancer	Lung cancer
KPNA1	Karyopherin subunit alpha 1 (ENSG00000114030)	3q21	<b>HPA:</b> Lung <sup>L</sup> , Lymph node <sup>L</sup> <b>PDB:</b> Natural killer cell, B- lymphocyte, Cytotoxic T- lymphocyte, Monocyte, Lymph node, Lung, Helper T-lymphocyte, Cardia, Blood platelet	Lymph node, Lung	Lung cancer	Lung cancer
MMP12	Matrix metallopeptidase 12 (ENSG00000262406)	11q22	<b>HPA:</b> Data not available <b>PDB:</b> Lymph node, Lung, Bone marrow, Blood platelet	Lymph node, Lung	Data not available	Lung cancer
CEACAM5	Cea cell adhesion molecule 5 (ENSG00000105388)	19q13	<b>HPA:</b> Lung <sup>L</sup> , Bone marrow <sup>L</sup> <b>PDB:</b> Lung	Lung	Lung cancer	Lung cancer

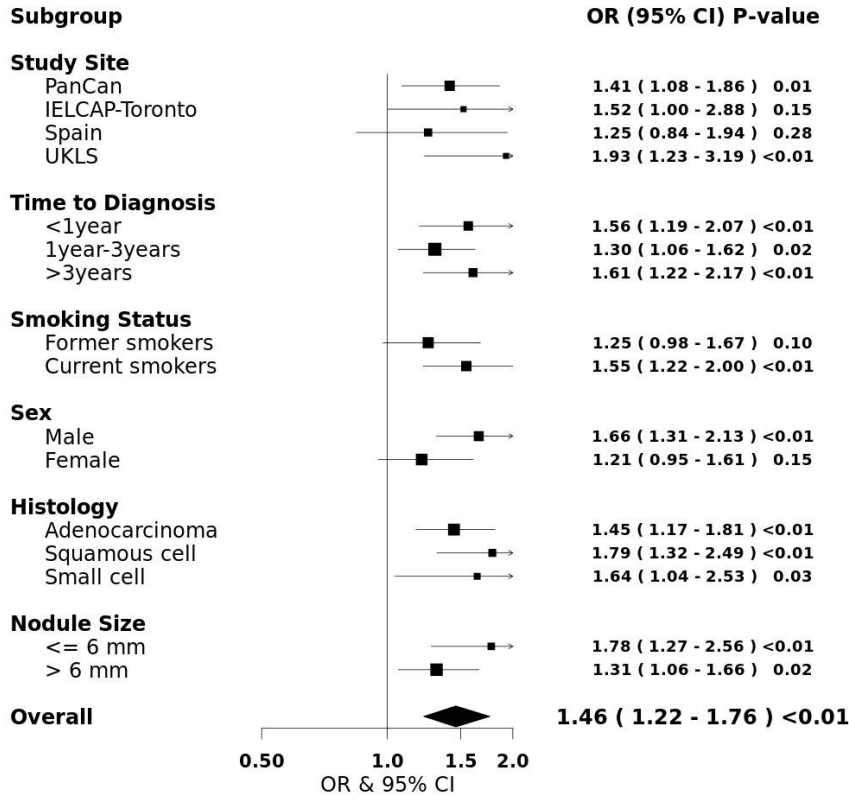
The Human Protein Atlas (HPA); H: High; M: Medium; L: Low  
ProteomicsDB (PDB)

**Supplementary Figure 1: Forest plots of the top circulating protein markers illustrating the association with nodule malignancy stratified by study site, time to diagnosis, smoking status, sex, histology and nodule size, after adjustment for Brock score. The Odds Ratios (OR) and the 95% Confidence Intervals (CI) are presented.**

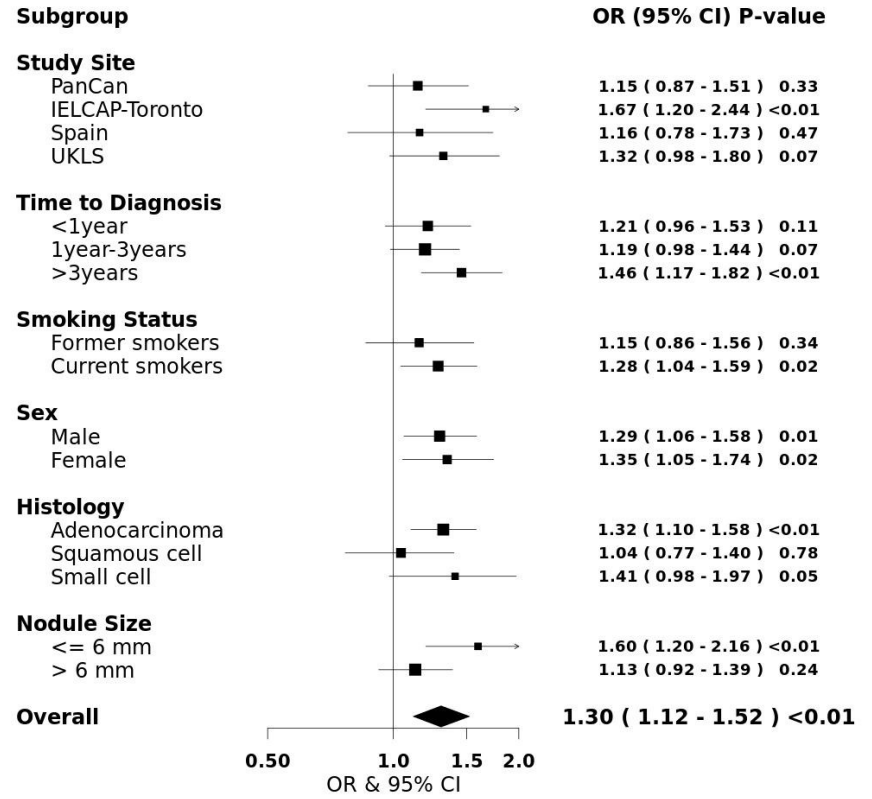




### TRAIL-R2

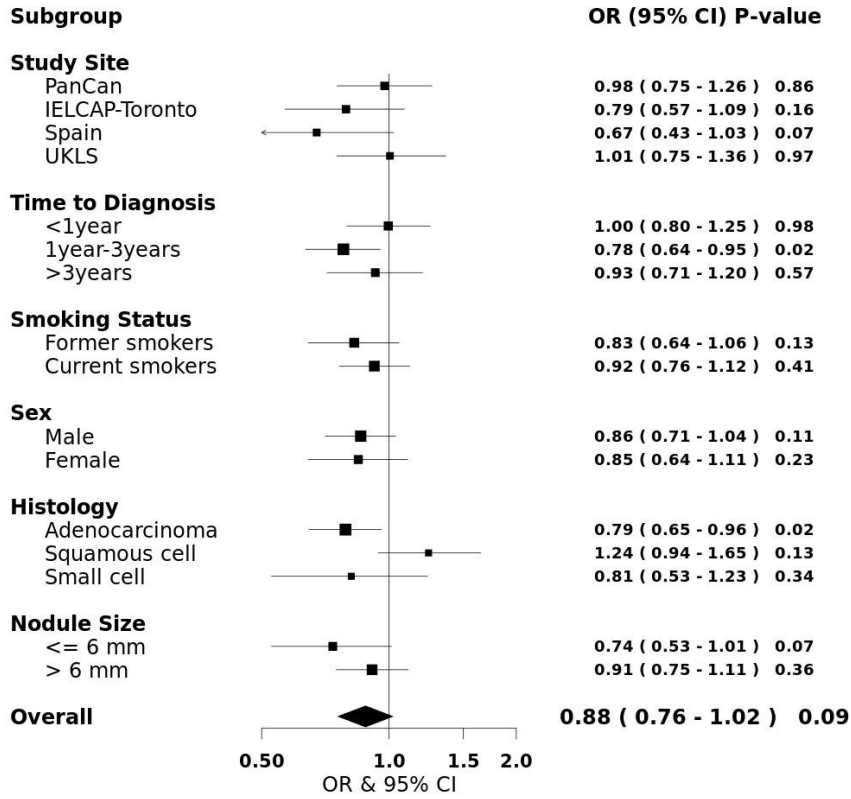


### SCGB3A2

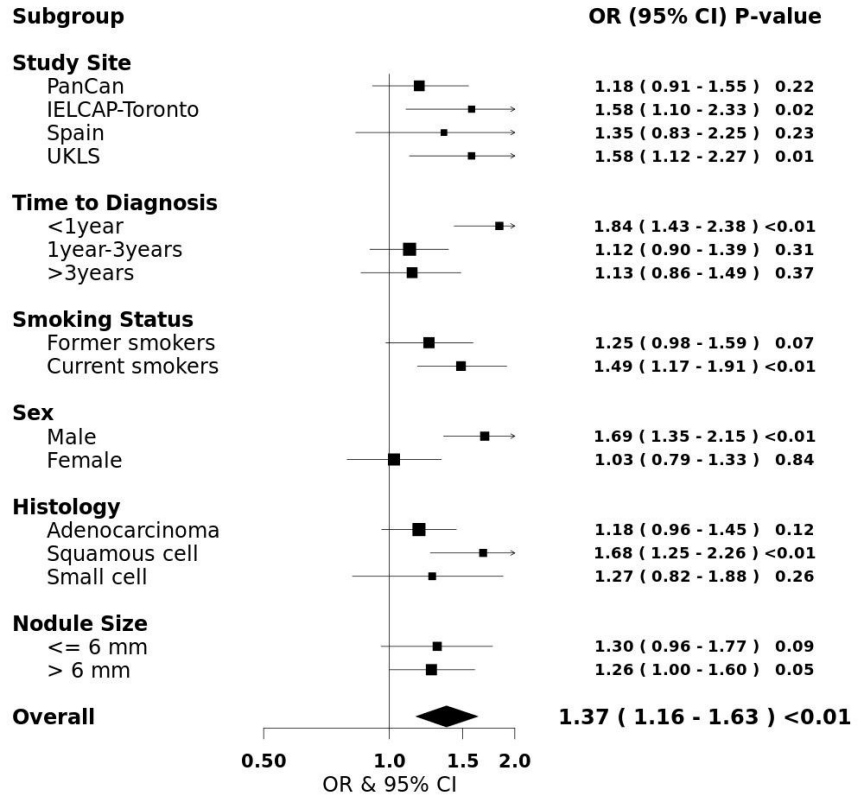


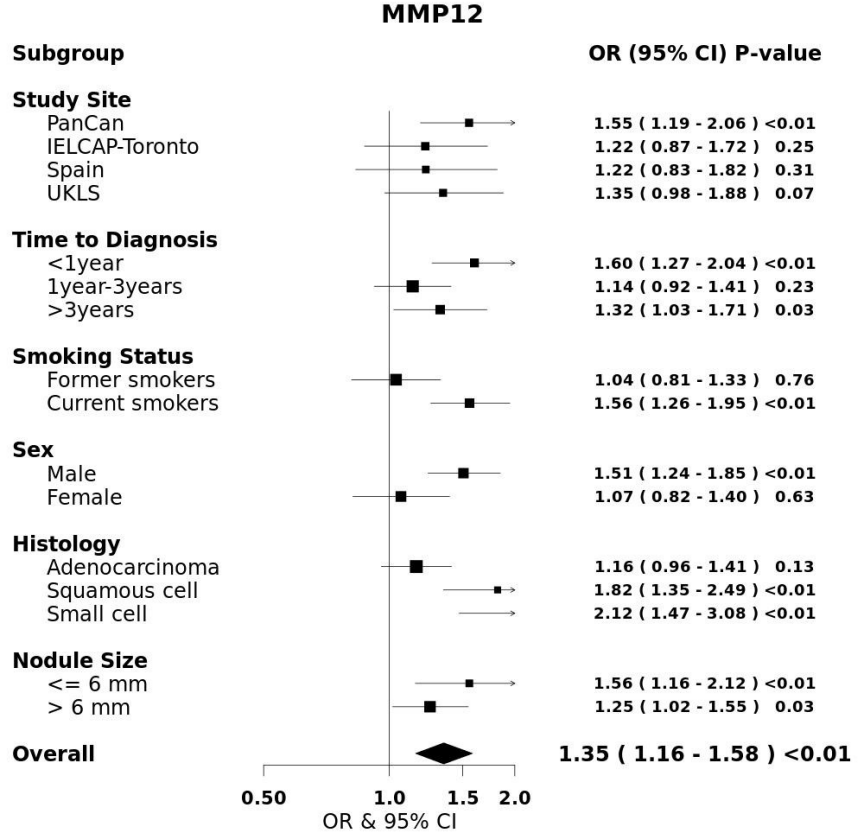
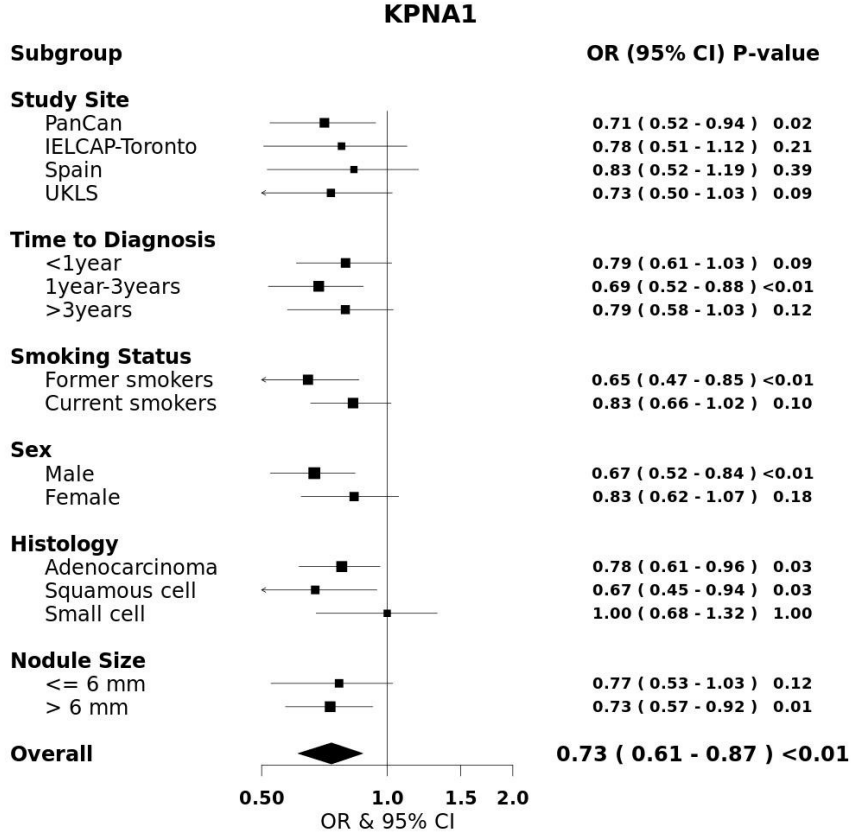


### FCRL5

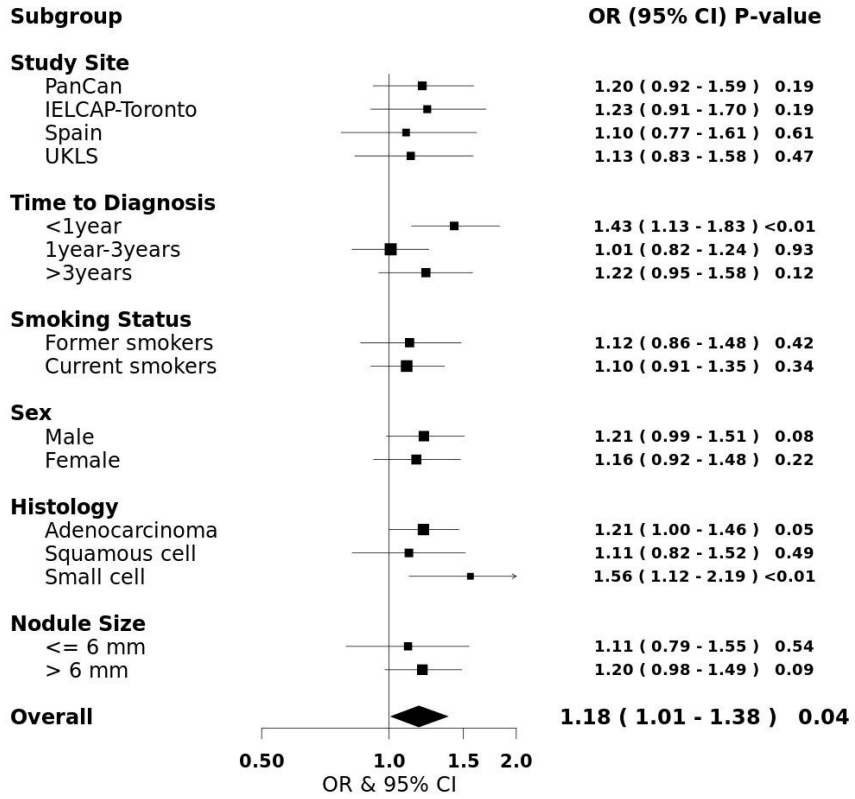


### KRT19

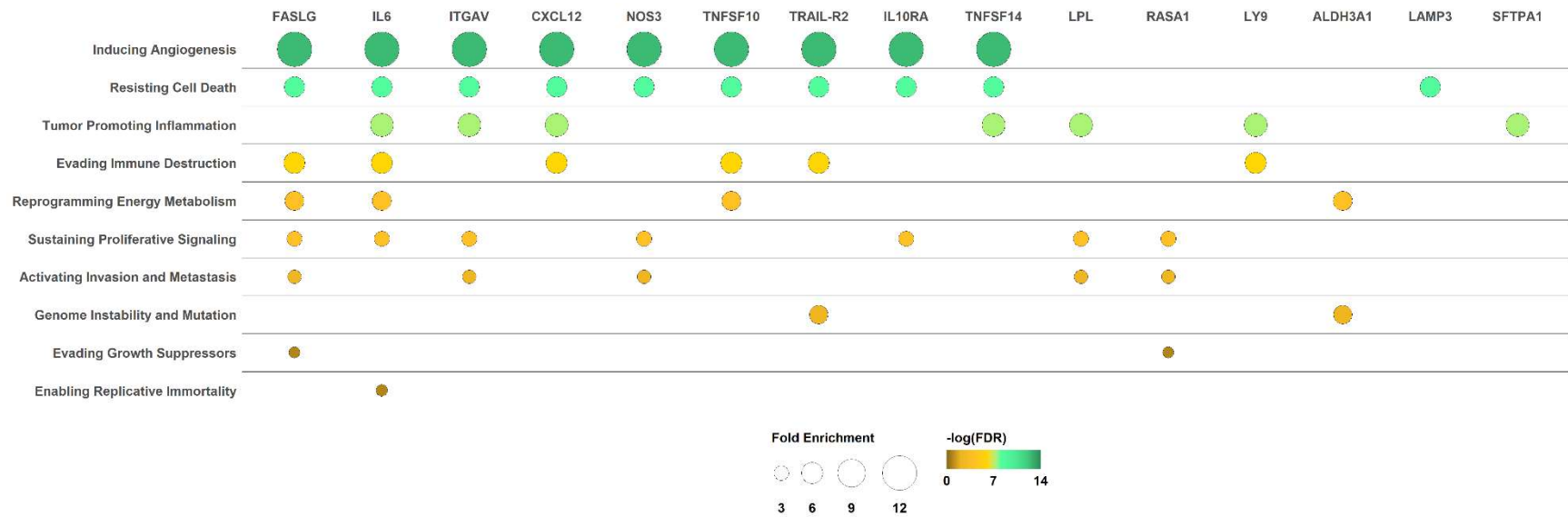




**CEACAM5**

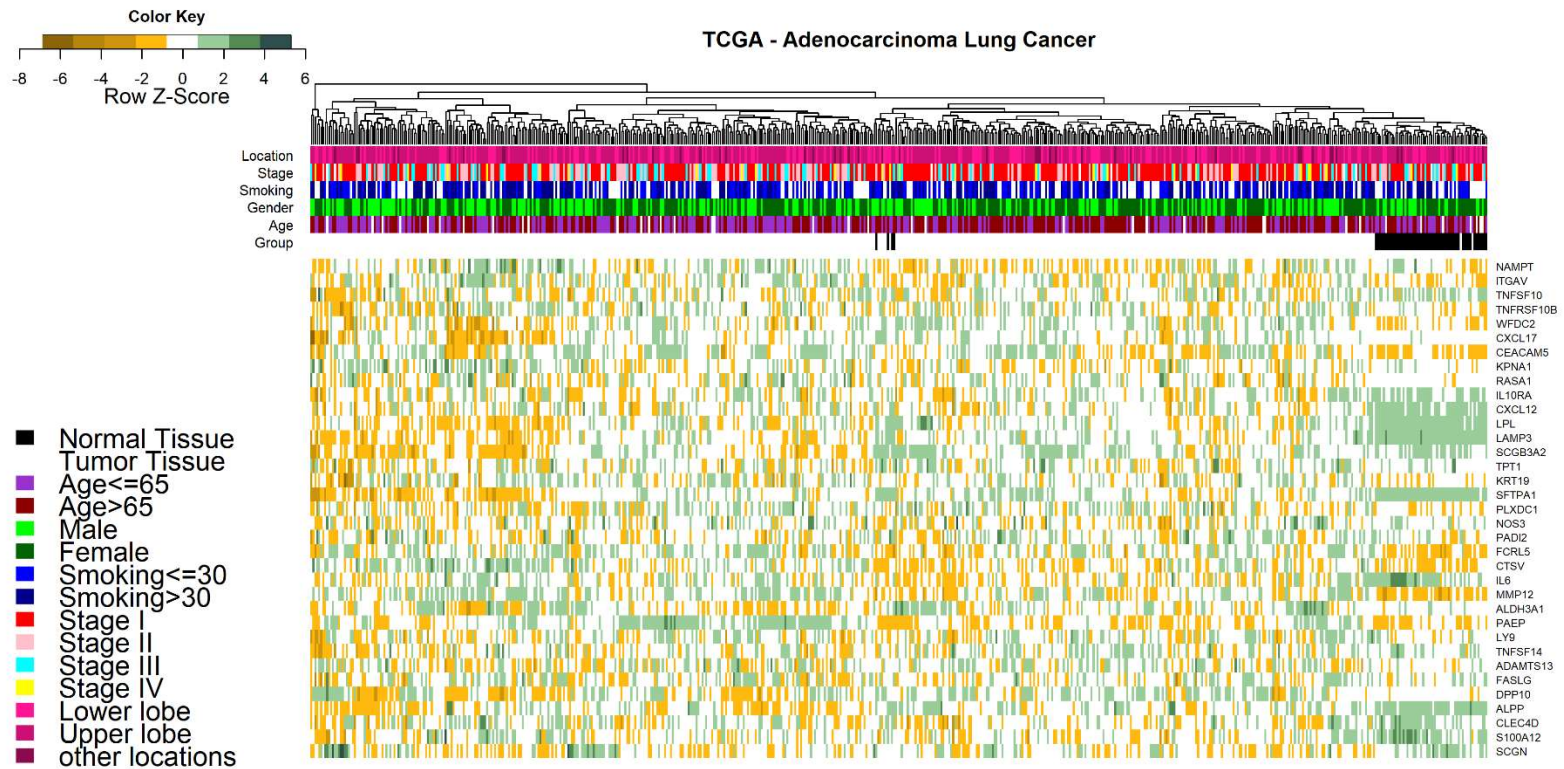


**Supplementary Figure 2: The cancer hallmarks represented by the 36 informative protein markers, and the corresponding allocation.** The fold enrichment is calculated for each hallmark separately, depicted by circle size, and the significance level ( $-\log(\text{FDR})$ ) is annotated by color gradient.

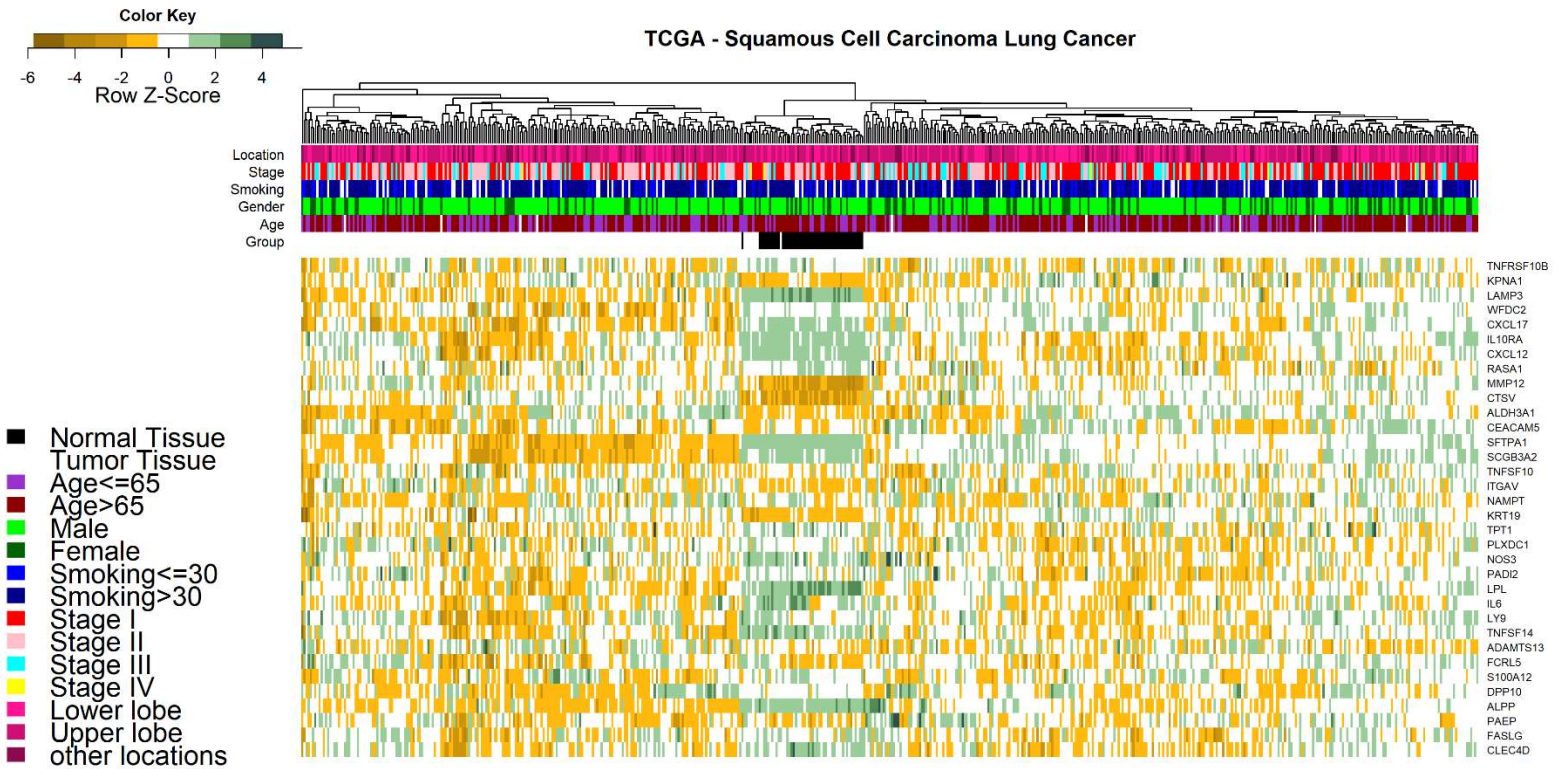


**Supplementary Figure 3: Heatmap of differential gene expressions based on TCGA data comparing lung cancer tissues vs lung normal tissues based on TCGA database by unsupervised hierarchical clustering. (A) Adenocarcinoma (B) Squamous cell carcinoma. This plot includes 35 genes, as TSHB data was not available in TCGA.**

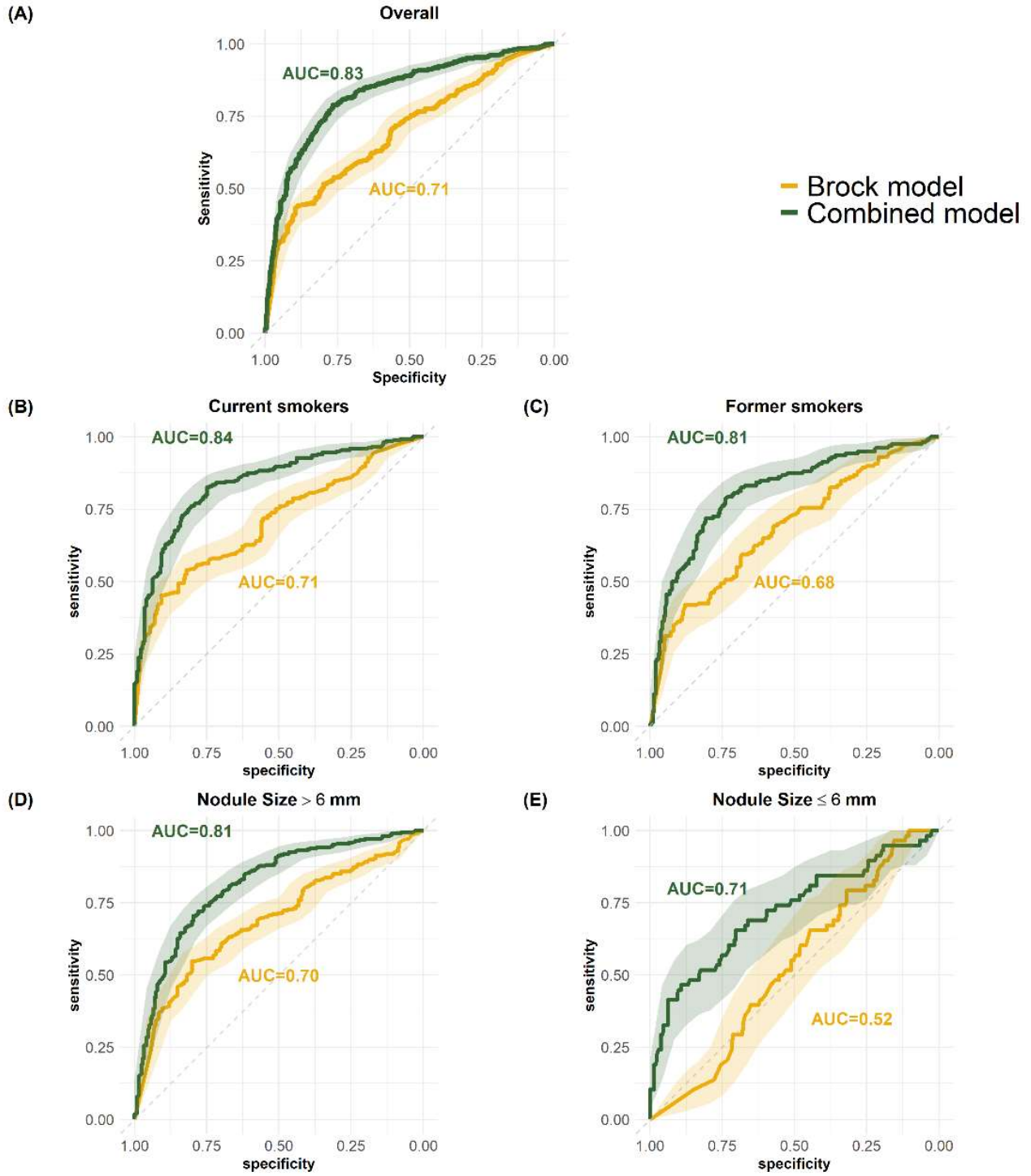
**(A) Adenocarcinoma lung cancer tissues**



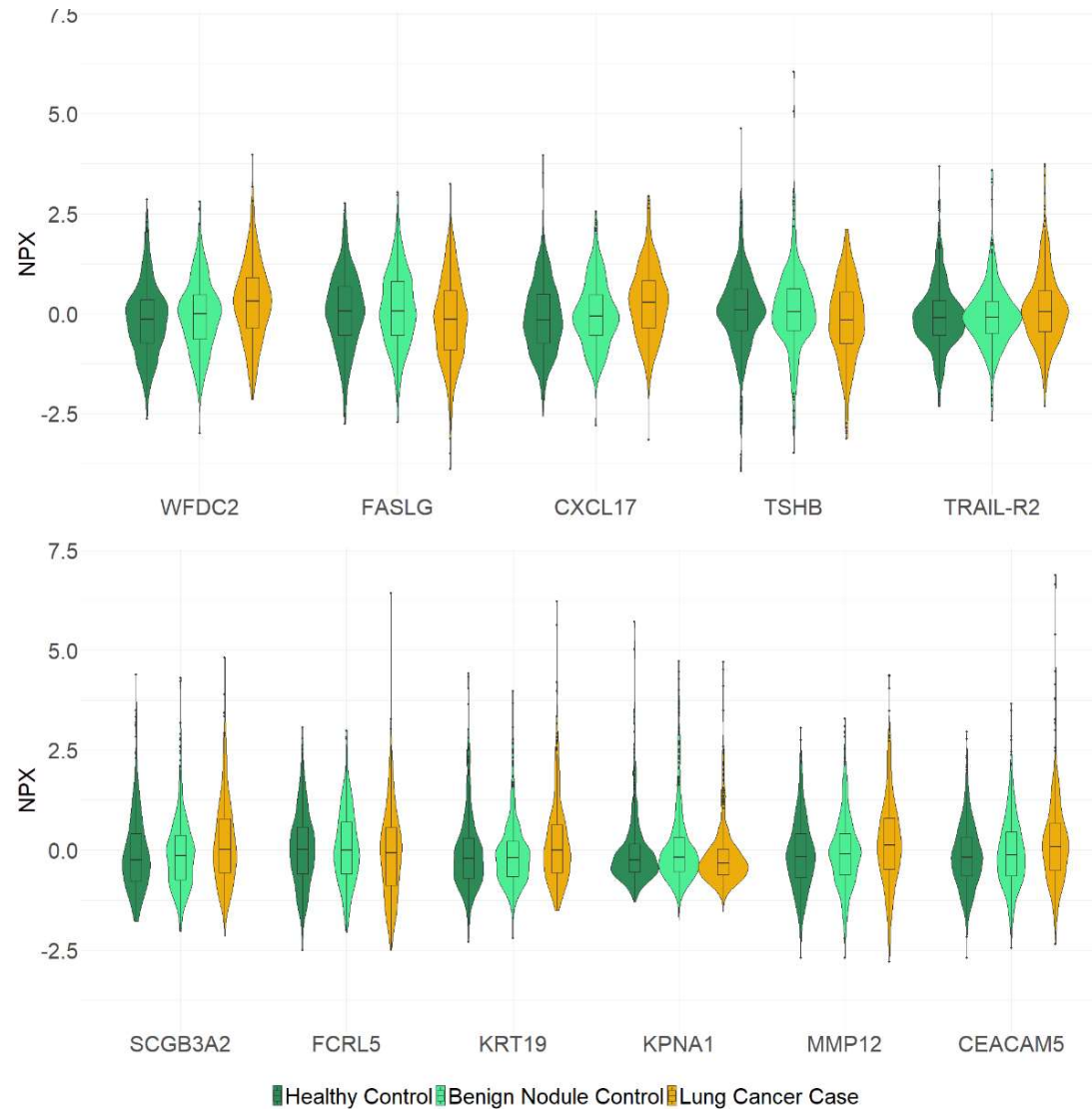
**(B) Squamous lung cancer tissues**



**Supplementary Figure 4: The ROC-AUC of the combined model including Brock score and the top 9 circulating protein markers in comparison to the Brock model, overall and in different subgroups by smoking and nodule size.**



**Supplementary Figure 5: Violin plots of the top circulating protein markers illustrating differential expression of the top protein markers in lung cancer subjects, subjects with benign nodules and subjects with no nodule.**





## REFERENCES

1. Tammemagi MC, Schmidt H, Martel S, *et al.* Participant selection for lung cancer screening by risk modelling (the Pan-Canadian Early Detection of Lung Cancer [PanCan] study): a single-arm, prospective study. *The Lancet. Oncology* 2017;18(11):1523-1531.
2. Field JK, Duffy SW, Baldwin DR, *et al.* The UK Lung Cancer Screening Trial: a pilot randomised controlled trial of low-dose computed tomography screening for the early detection of lung cancer. *Health Technology Assessment (Winchester, England)* 2016;20(40):1-146.
3. Field JK, Duffy SW, Baldwin DR, *et al.* UK Lung Cancer RCT Pilot Screening Trial: baseline findings from the screening arm provide evidence for the potential implementation of lung cancer screening. *Thorax* 2016;71(2):161-170.
4. Chung M, Tam K, Wallace C, *et al.* International Early Lung Cancer Action Program: update on lung cancer screening and the management of CT screen-detected findings. *AME Medical Journal* 2017;2(8).
5. Menezes RJ, Roberts HC, Paul NS, *et al.* Lung cancer screening using low-dose computed tomography in at-risk individuals: the Toronto experience. *Lung Cancer (Amsterdam, Netherlands)* 2010;67(2):177-183.
6. Sanchez-Salcedo P, Berto J, de-Torres JP, *et al.* Lung cancer screening: fourteen year experience of the Pamplona early detection program (P-IELCAP). *Archivos De Bronconeumologia* 2015;51(4):169-176.
7. The Human Protein Atlas. *The Human Protein Atlas*. <https://www.proteinatlas.org/> (May 31, 2022; date last accessed).
8. Schmidt T, Samaras P, Frejno M, *et al.* ProteomicsDB. *Nucleic Acids Res* 2018;46(D1):D1271-d1281.
9. Samaras P, Schmidt T, Frejno M, *et al.* ProteomicsDB: a multi-omics and multi-organism resource for life science research. *Nucleic Acids Research* 2020;48(D1):D1153-D1163.
10. ProteomicsDB. ProteomicsDB. 2021, <https://www.proteomicsdb.org/>.