Supplementary

Table S1 Distribution of samples according to available sample type

Availability of samples	Sample type			
	Tissue	Blood	Sputum-supernatant	Sputum-sediment
All available	19	19	19	19
No sputum sediment	6	6	6	0
No sputum supernatant	11	11	0	11
No blood	1	0	1	1
No blood and sputum supernatant	1	0	0	1
No tissue	0	3	3	3
Total	38	39	29	35

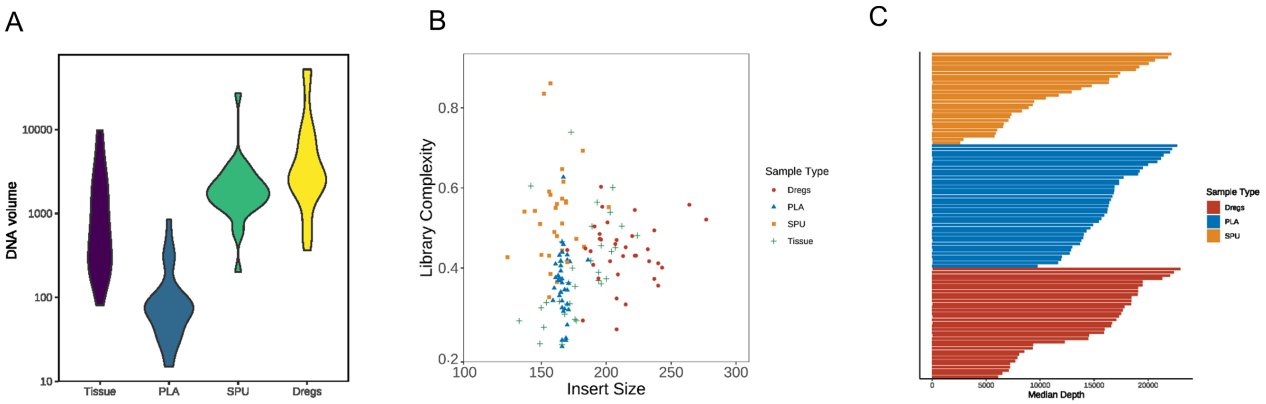

Figure S1 Quality control metrics including the DNA yield expressed in nanograms (A), the correlation between library complexity and insert size of the fragments (in base pairs) for sequencing (B), and the median sequencing depth (C) of all the samples analyzed in the study.

Table S2 By variant comparison between the variations detected from sputum supernatant and corresponding sediment samples of the 23 patients

Distribution of variation types	Total numbers of variations in sputum supernatant	Total numbers of variations in sputum sediment	Numbers of variations shared	Numbers of variations detected only in supernatant	Numbers of variations detected only in sediment	Concordance rate (%)
168 genes						
SNV + Indels	48	55	24	24	31	30.4
CNV	1	0	0	1	0	0
Fusions	5	4	4	1	0	80.0
8 driver genes						
SNV + Indels	13	7	5	8	2	33.3
CNV	0	0	0	0	0	NA
Fusions	4	4	4	0	0	100

SNV, single nucleotide variation; indels, small insertion or deletions; CNV, copy number variation.

https://dx.doi.org/10.21037/atm-22-1297

Table S3 By variant comparison between the variations detected from sputum supernatant and matched tissue samples of the 26 patients

Distribution of variation types	Total numbers of variations in sputum supernatant	Total numbers of variations in tissue samples	Numbers of variations shared	Numbers of variations detected only in sputum supernatant	Numbers of variations detected only in tissue samples	Concordance rate (%)
168 genes						
SNV + Indels	43	91	34	9	57	34.0
CNV	2	12	1	1	11	7.7
Fusions	6	8	6	0	2	75.0
8 driver genes						
SNV + Indels	16	33	14	2	19	40.0
CNV	1	3	0	1	3	0
Fusions	5	6	5	0	1	83.3

SNV, single nucleotide variation; Indels, small insertion or deletions; CNV, copy number variation.

Table S4 By variant comparison between the variations detected from sputum supernatant and matched plasma samples of the 28 patients

Distribution of variation types	Total numbers of variations in sputum supernatant	Total numbers of variations in plasma samples	Numbers of variations shared	Numbers of variations detected only in sputum supernatant	Numbers of variations detected only in plasma samples	Concordance rate (%)
168 genes						
SNV + Indels	53	57	26	27	31	31.0
CNV	3	11	1	2	10	7.7
Fusions	6	5	5	1	0	83.5
8 driver genes						
SNV + Indels	18	20	12	6	8	46.2
CNV	1	3	0	1	3	0
Fusions	5	4	4	1	0	80.0

SNV, single nucleotide variation; Indels, small insertion or deletions; CNV, copy number variation.

Table S5 Univariate analysis between clinicopathologic features and either variation detection rates or maximum allelic fraction in sputum supernatant and sediment samples of the cohort

Clinicopathologic features	Variation detection rate (P value)		Maximum allelic fraction (P value)	
	Sputum supernatant	Sputum sediment	Sputum supernatant	Sputum sediment
Age	0.922	0.944	0.299	0.285
Gender	0.390	1.000	0.060	0.269
Smoking status	0.083	0.705	0.018*	0.728
Histology	0.442	1.000	0.576	0.911
Degree of cellular differentiation of sputum cytology	0.544	0.408	0.227	0.908
Location of primary tumor	1.000	1.000	0.972	0.581
Disease stage	1.000	1.000	0.828	0.885

Asterisk (*) denotes statistical significance.

https://dx.doi.org/10.21037/atm-22-1297

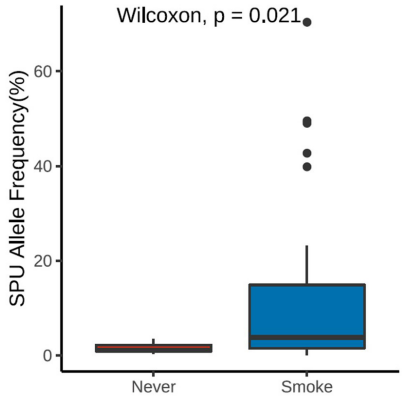

Figure S2 Smokers had significantly higher allelic fraction of sputum supernatant. Box plot illustrating the distribution of allelic frequencies in sputum supernatant (SPU) from smokers and never-smokers.

https://dx.doi.org/10.21037/atm-22-1297