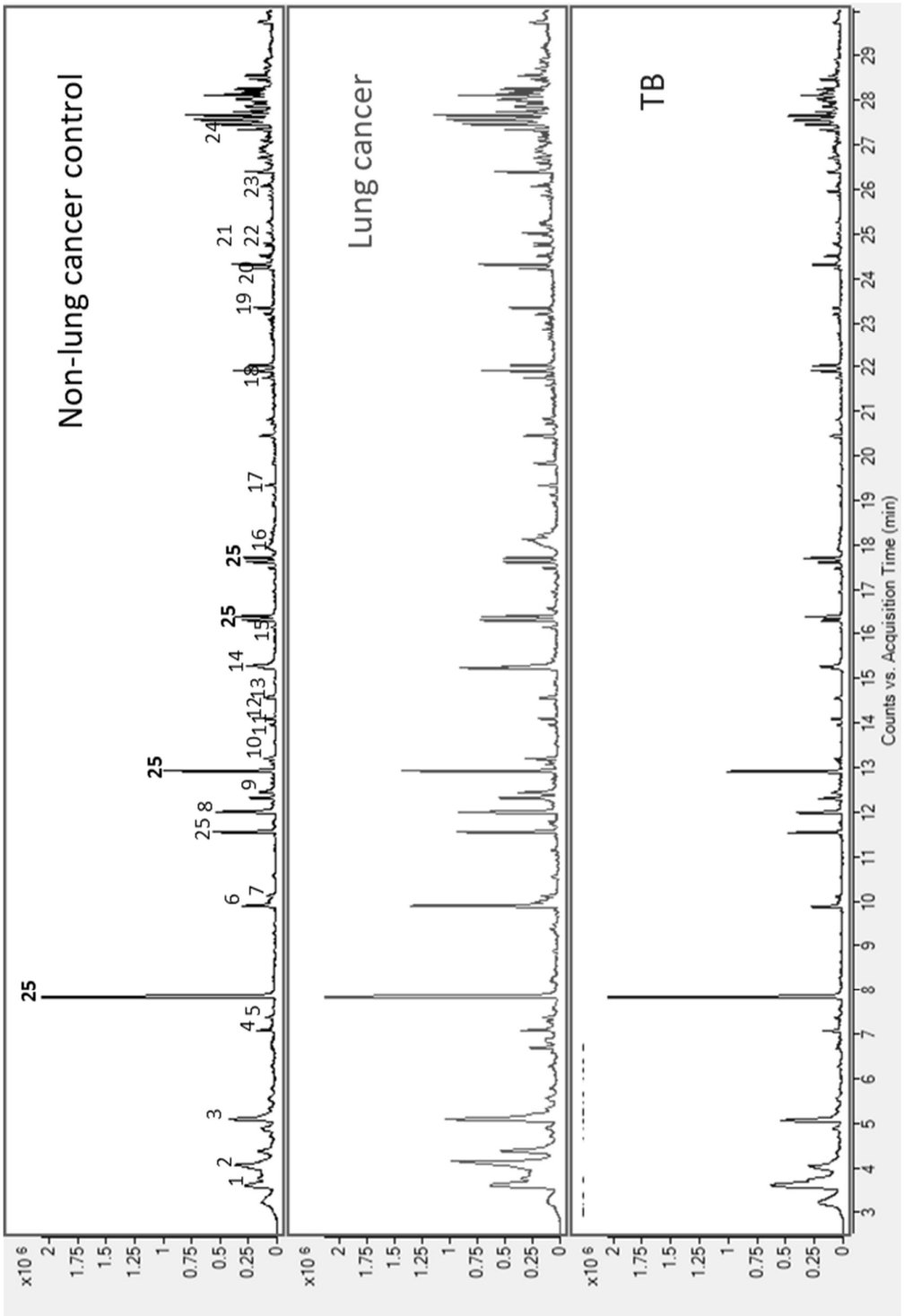
Non-lung cancer control
Lung cancer
TB
x10 6
Counts vs. Acquisition Time (min)

FIGURE 1

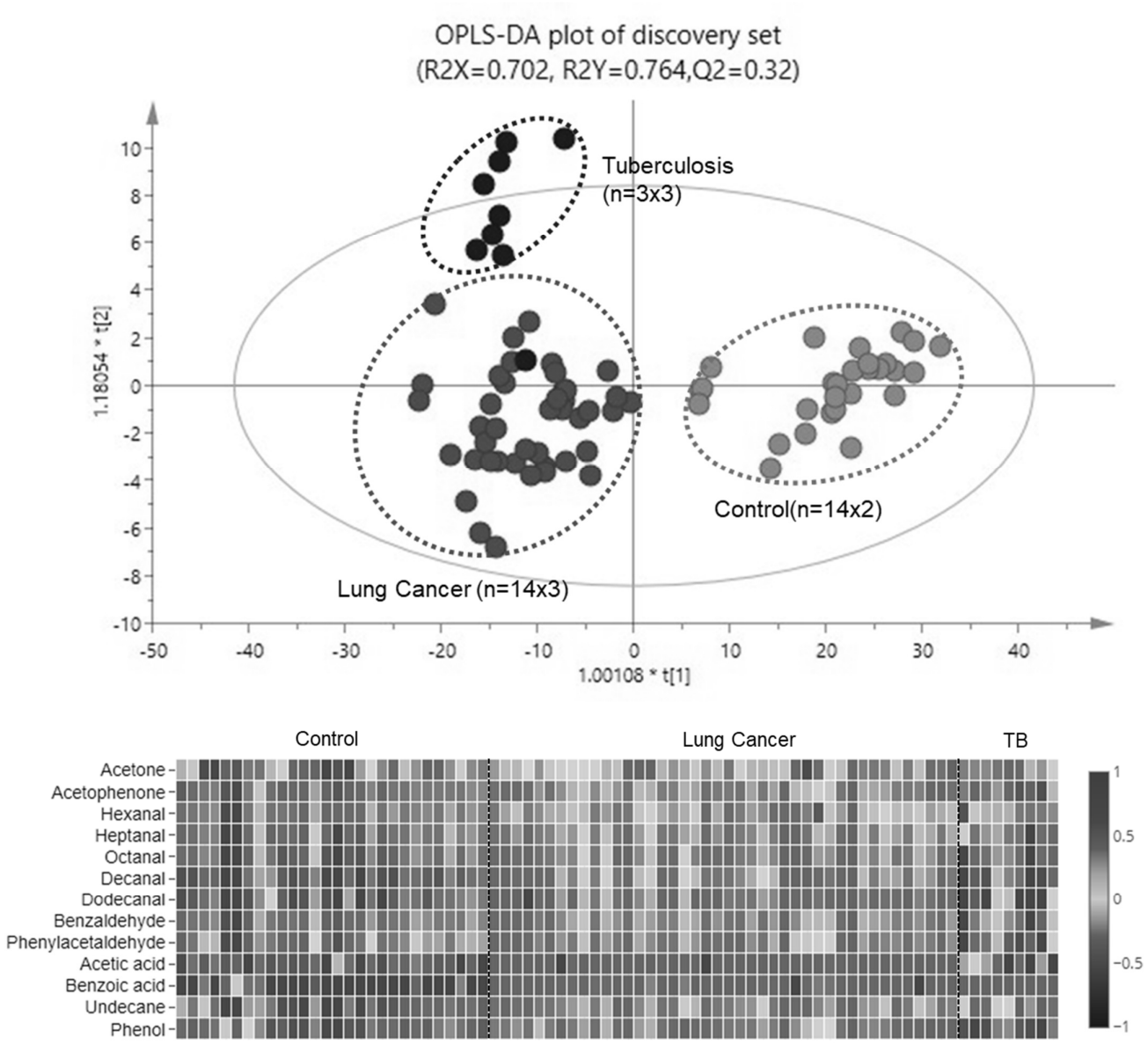
OPLS-DA plot of discovery set
(R2X=0.702, R2Y=0.764,Q2=0.32)
Tuberculosis
(n=3x3)
Control(n=14x2)
Lung Cancer (n=14x3)
1.18054 * t[2]
1.00108 * t[1]
Control
Lung Cancer
TB
Acetone
Acetophenone
Hexanal
Heptanal
Octanal
Decanal
Dodecanal
Benzaldehyde
Phenylacetaldehyde
Acetic acid
Benzoic acid
Undecane
Phenol

FIGURE 2

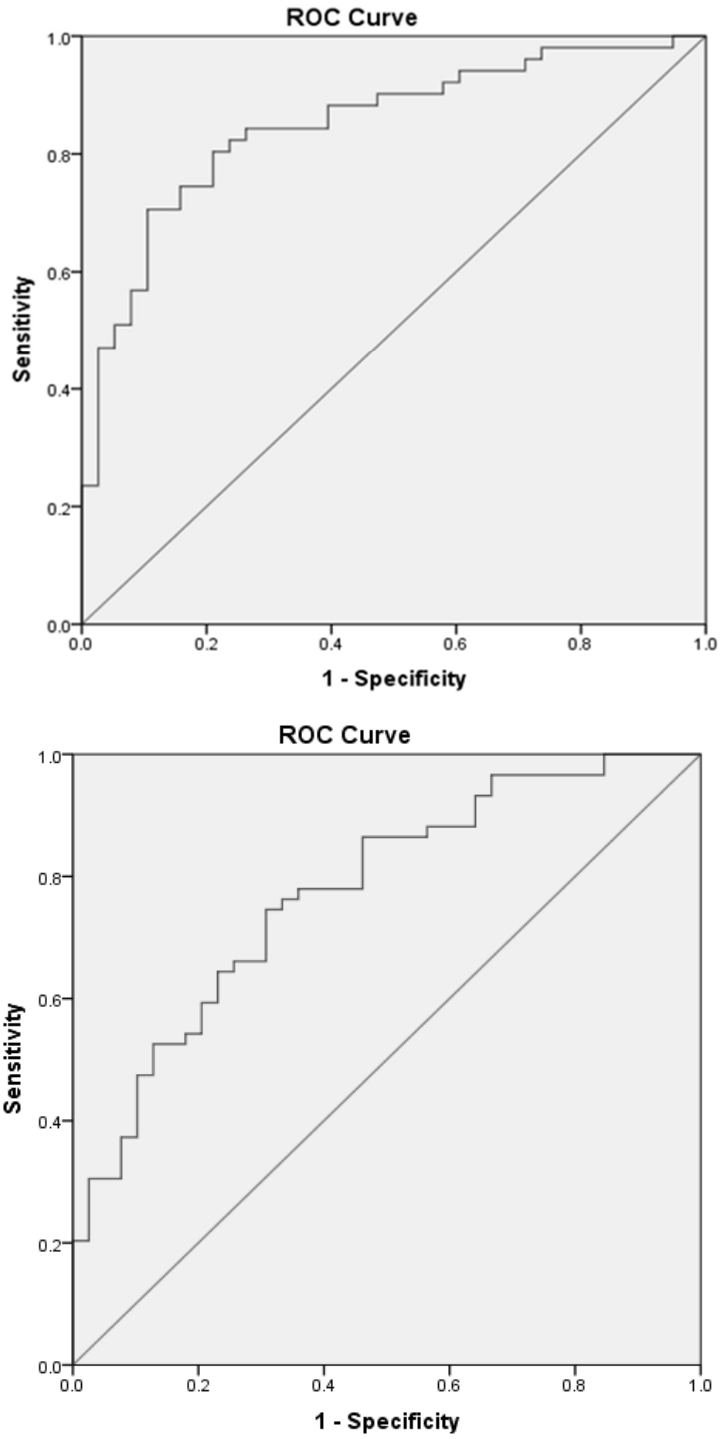
ROC Curve
Sensitivity
1 - Specificity
0.0
0.2
0.4
0.6
0.8
1.0
ROC Curve
Sensitivity
1 - Specificity
0.0
0.2
0.4
0.6
0.8
1.0

Figure 3

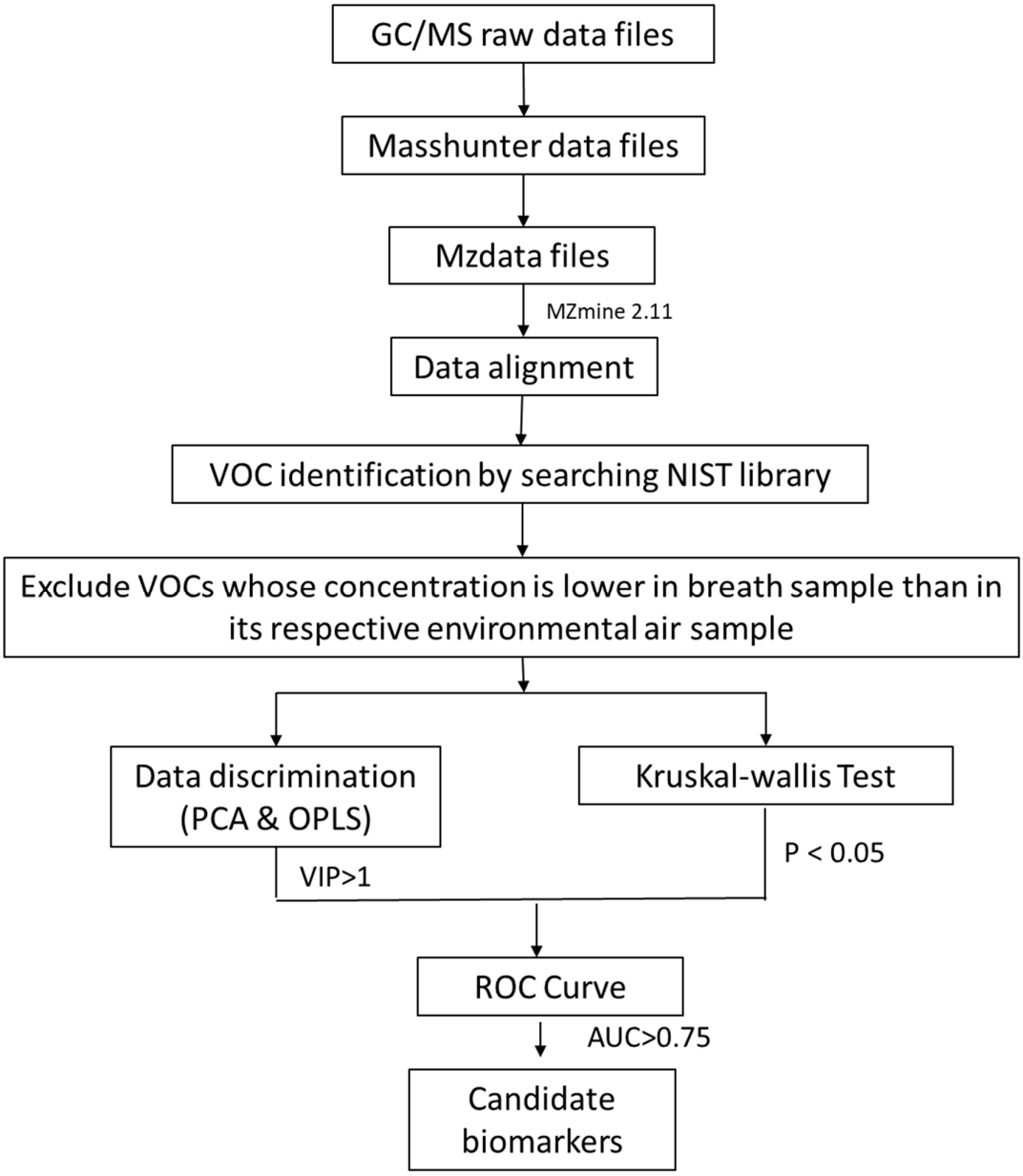
GC/MS raw data files
Masshunter data files
Mzdata files
MZmine 2.11
Data alignment
VOC identification by searching NIST library
Exclude VOCs whose concentration is lower in breath sample than in its respective environmental air sample
Data discrimination (PCA & OPLS)
Kruskal-wallis Test
P < 0.05
VIP>1
ROC Curve
AUC>0.75
Candidate biomarkers

FIGURE 4

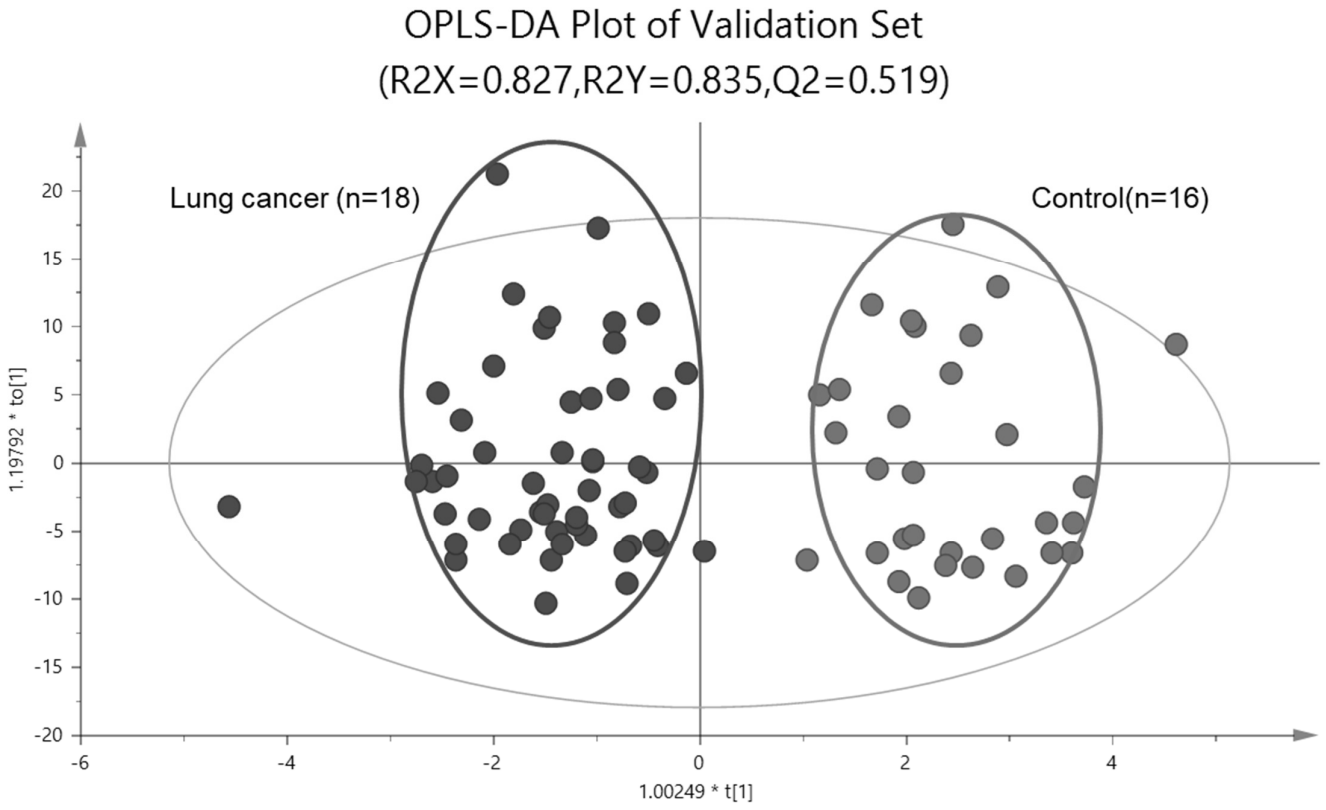
OPLS-DA Plot of Validation Set
(R2X=0.827,R2Y=0.835,Q2=0.519)
Lung cancer (n=18)
Control(n=16)
1.19792 * to[1]
1.00249 * t[1]
20
15
10
5
0
-5
-10
-15
-20
-6
-4
-2
0
2
4

FIGURE 5

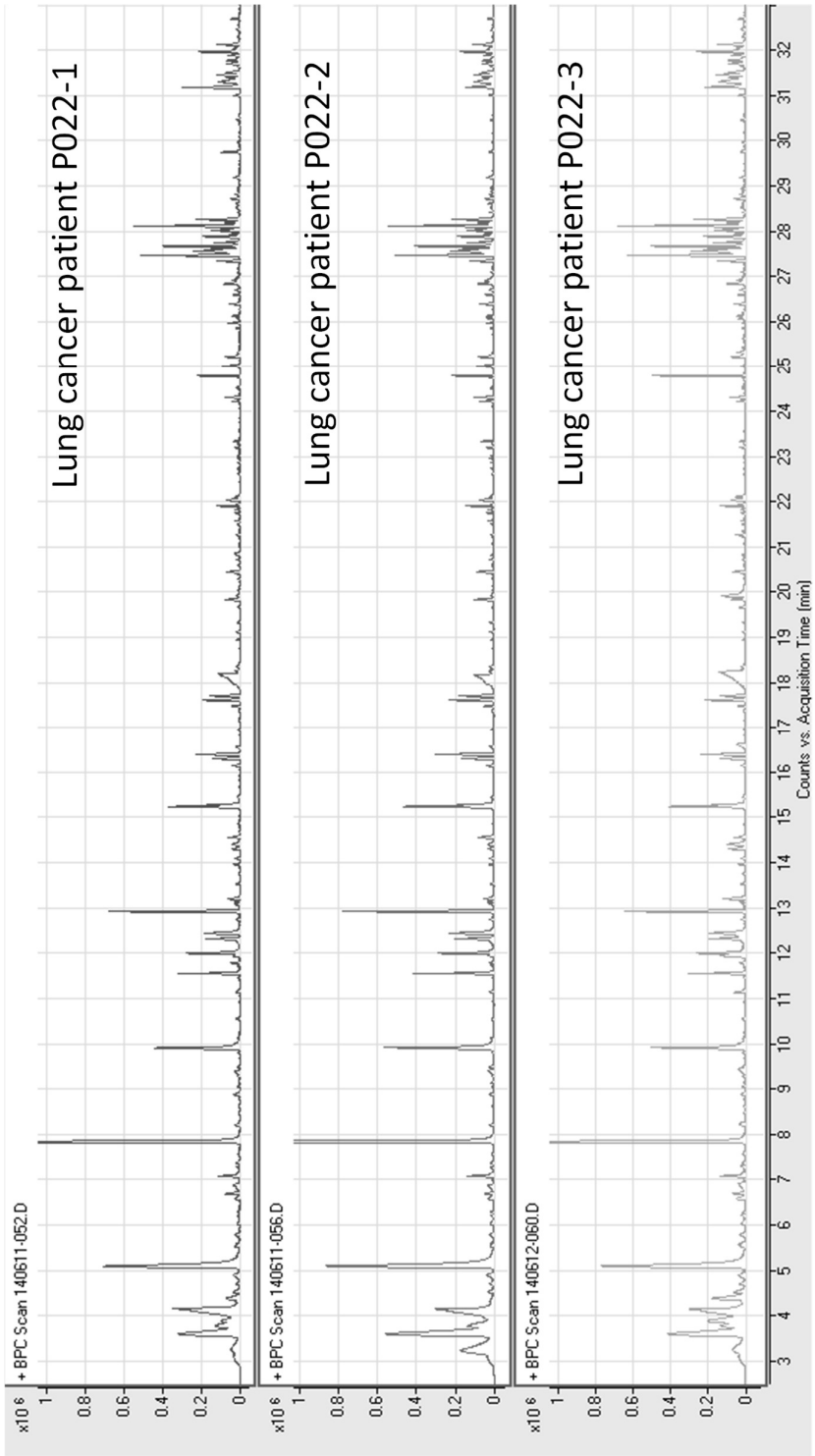
Lung cancer patient P022-1
Lung cancer patient P022-2
Lung cancer patient P022-3
+ BPC Scan 140611-052.D
+ BPC Scan 140611-056.D
+ BPC Scan 140612-060.D
x10 6
Counts vs. Acquisition Time (min)

FIGURE 6