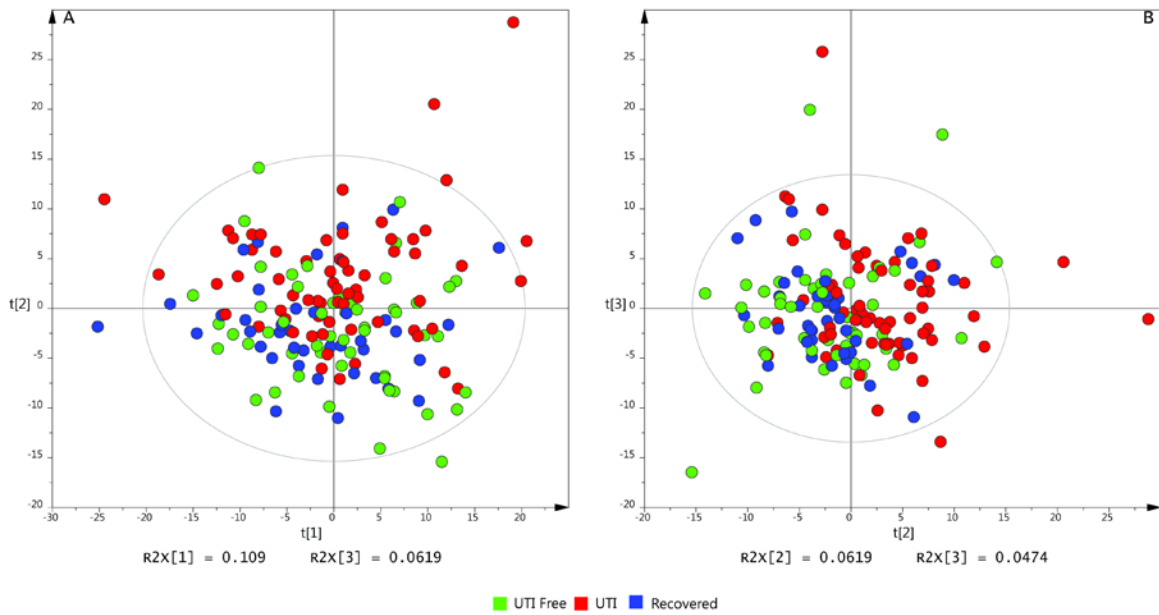
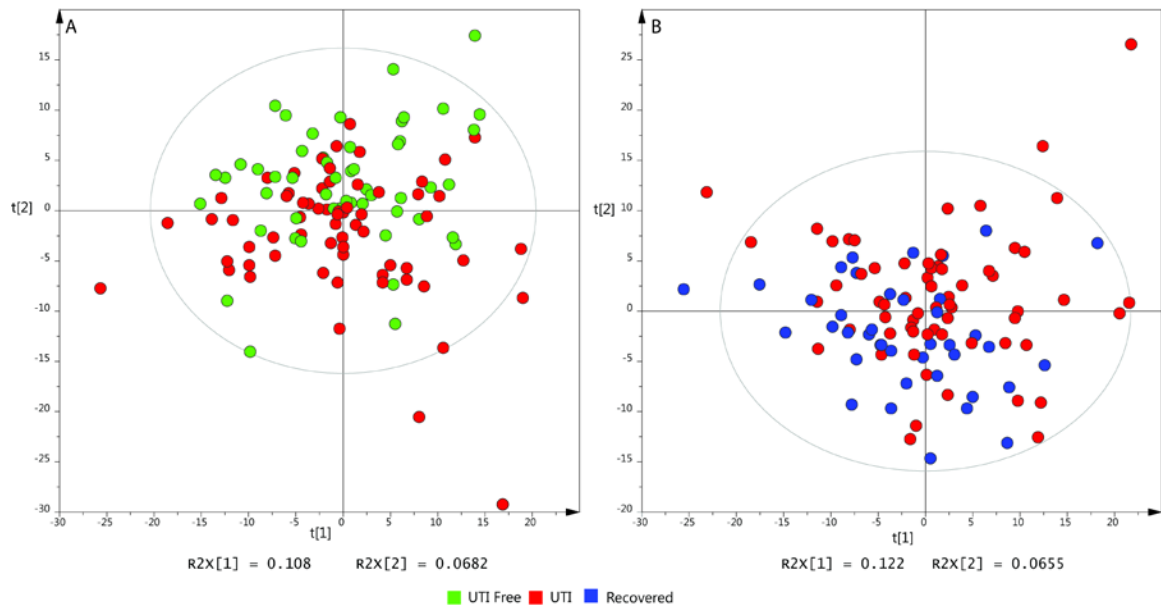


Supplementary Figure 1. The score plot of a PCA model built on the entire data set, 24 components cover 56% of variance, dots are colored according to their ID: QC and samples.

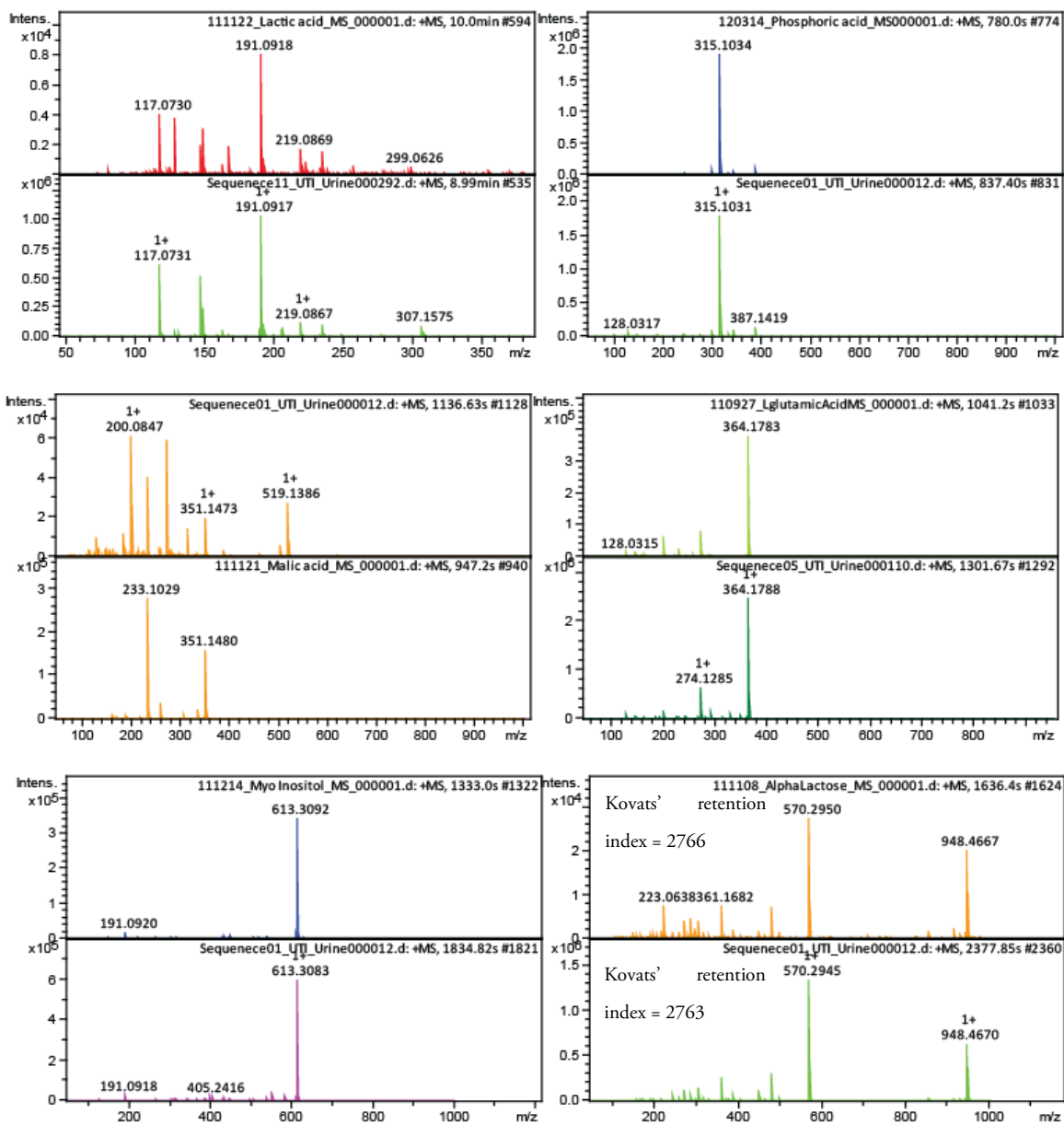


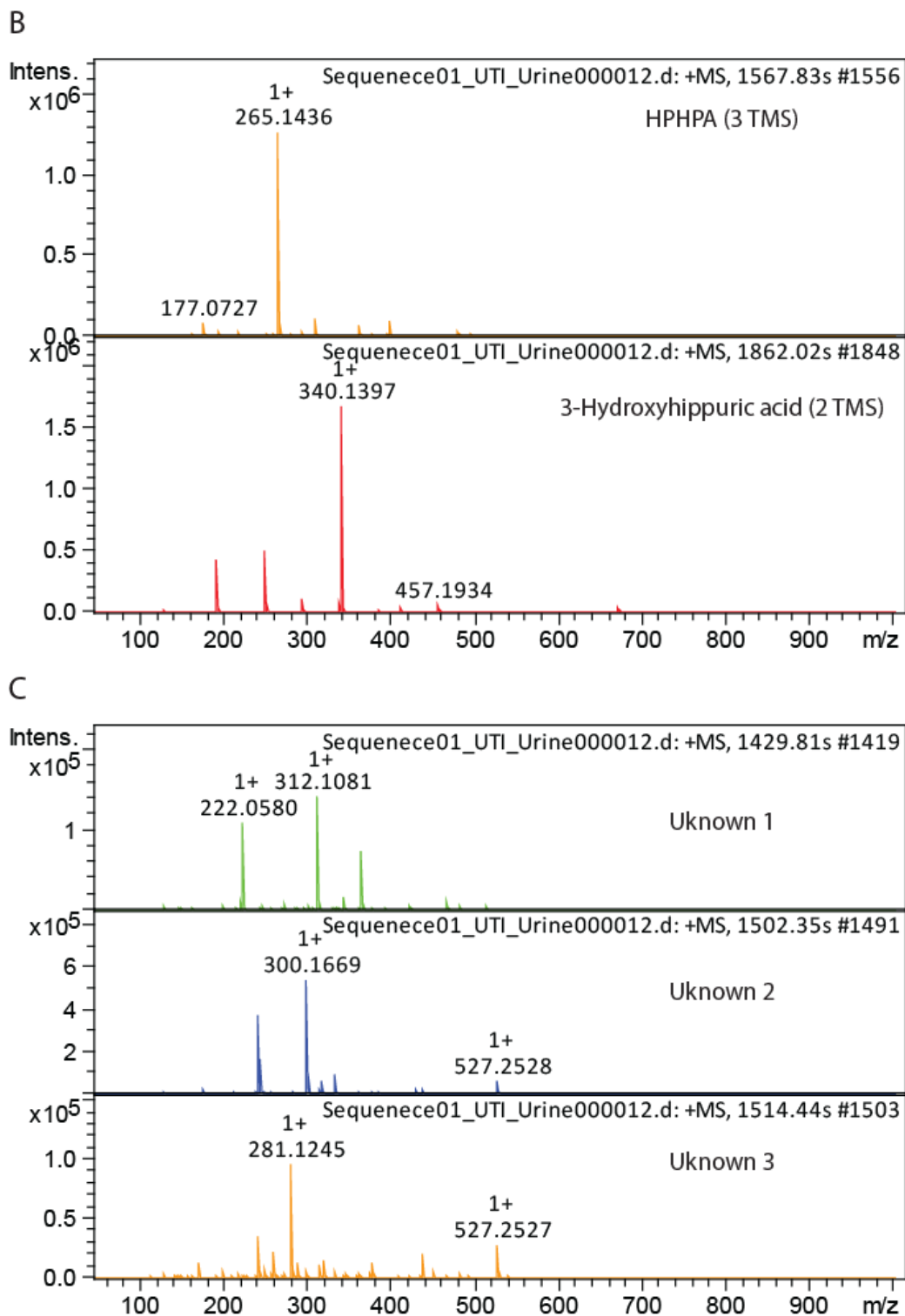
Supplementary Figure 2. A score plots showing the first and second (A) and the second and third (B) components of the PCA model (24 components cover 59% of the variance) built of 155 clinical samples. The samples are coloured according to the infectious status. The first and the second components display a major trend which cover more than 16% of the total variability (A). However, the second and the third component show a tendency in the cluster of the controls and UTI patients $t=30$ days against UTI patients $t=0$ day (B).



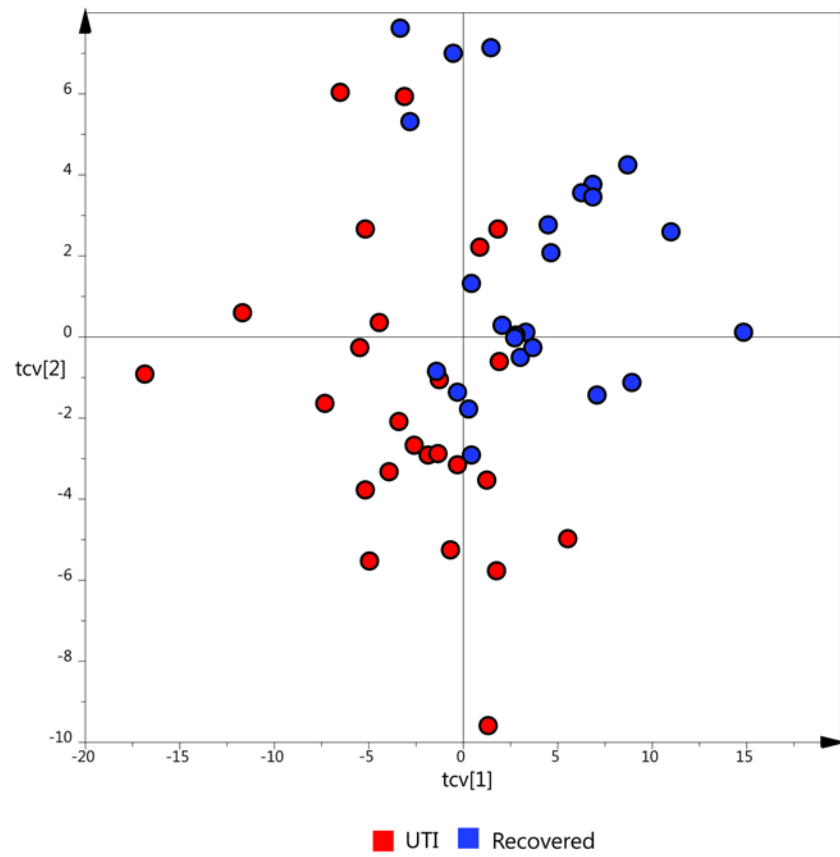
Supplementary Figure 3. PCA score plots for the models built for cross-sectional (A) and longitudinal data (B) parts separately. A) UTI Free + UTI t=0; 9 components cover 57% of the variance. B) UTI t=0 + UTI t=30; 16 components cover 55% of the variance .

A



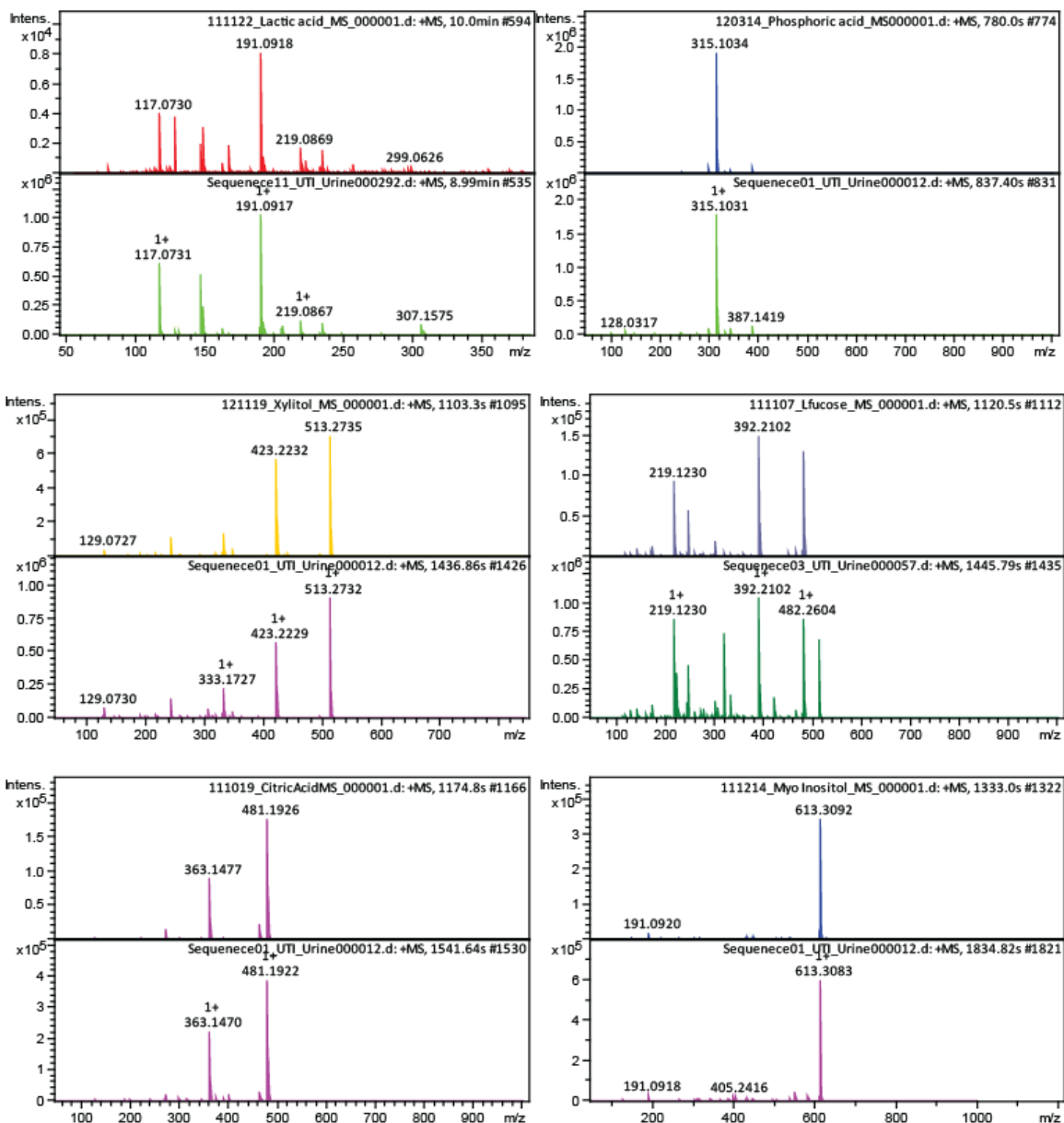


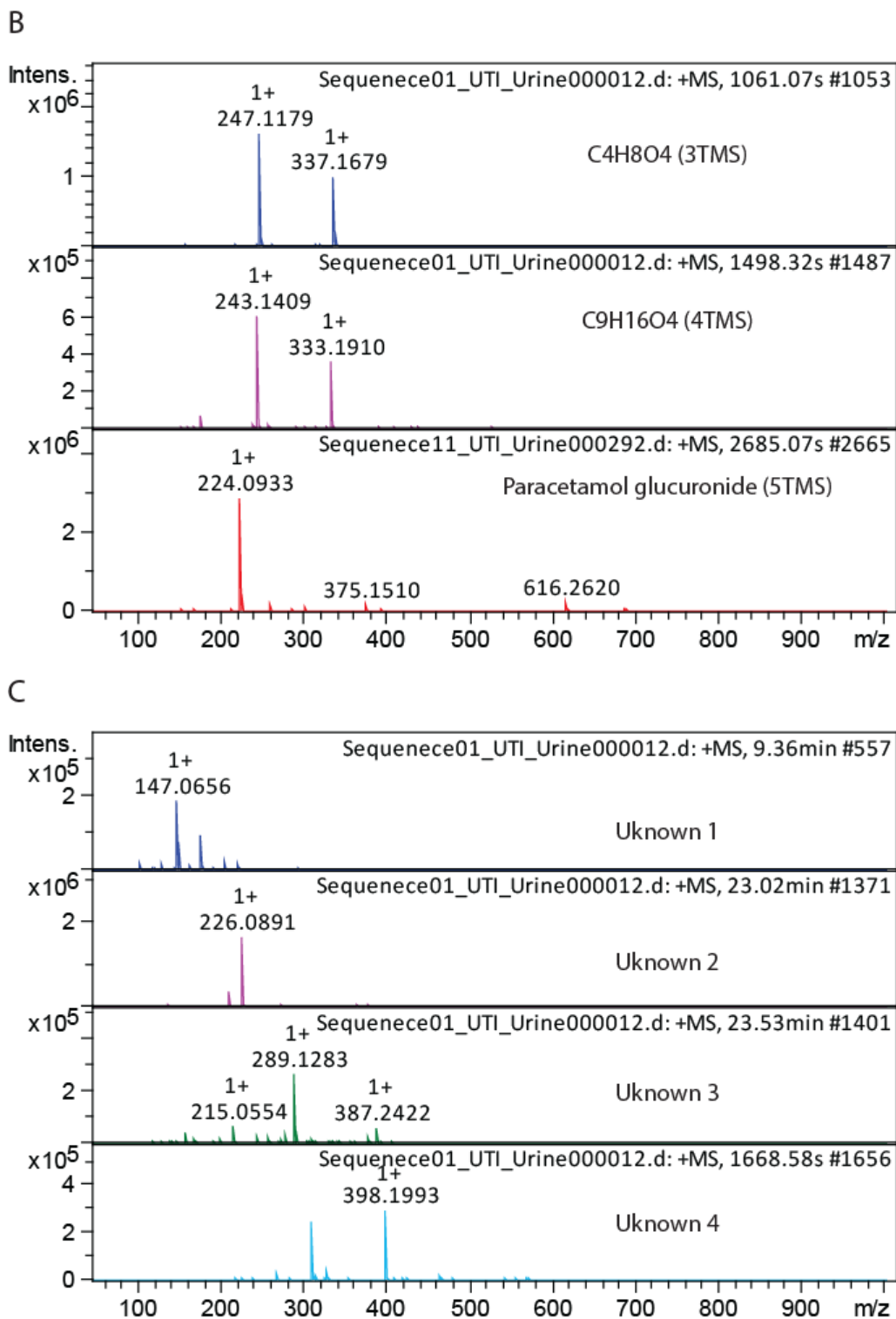
Supplementary Figure 4. MS spectra of classifiers reported in Table 2. A) Compounds identified with library matching B) compounds identified with *de novo* identification protocol and C) Unknown compounds.



Supplementary Figure 5. Cross-validated score plot of a PLS-DA built on UTI subjects (*E.coli*) who presented samples at both time point $t=0$ and $t=30$ days (50 samples in total). $R^2X=0.186$, $R^2Y=0.816$ $Q^2=0.46$.

A





Supplementary Figure 6. MS spectra of classifiers reported in Table 4. A) compounds identified with library matching B) Compounds identified with *de novo* identification protocol and C) Unknown compounds.