

# The Development of an LC-FAIMS-MS Metabolomics Workflow and its Application to the Discrimination of Benign from Malignant Renal Cell Masses

Alasdair Edge<sup>1</sup>; Grant Stewart<sup>2</sup>; Kayleigh Arthur<sup>1</sup>; Lauren Brown<sup>1</sup>; Chris Hodgkinson<sup>1</sup>; Aditya Malkar<sup>1</sup>; Paul Nasca<sup>1</sup>; Marc van der Schee<sup>1</sup>

<sup>1</sup>Owlstone Medical Ltd., 162 Cambridge Science Park, Cambridge, CB4 0GH, UK, <sup>2</sup>Academic Urology Group, University of Cambridge, Box 43, Addenbrookes Hospital, Cambridge Biomedical Campus, Hills Road, Cambridge, CB2 0QQ

For further information, email: [ultrafaims@owlstone.co.uk](mailto:ultrafaims@owlstone.co.uk)

## 1. Introduction

In non-targeted 'omics applications liquid or gas chromatography (LC or GC) is typically combined with mass spectrometry (MS) for analysis of complex biological matrices.

Molecular features can be missed by LC-MS due to:

- Trace component suppression by chemical noise.
- Chromatographically unresolved isomeric species.

Orthogonality between field asymmetric ion mobility spectrometry (FAIMS), LC and MS provides additional unique compound identifiers with detection of features based on (Figure 1):

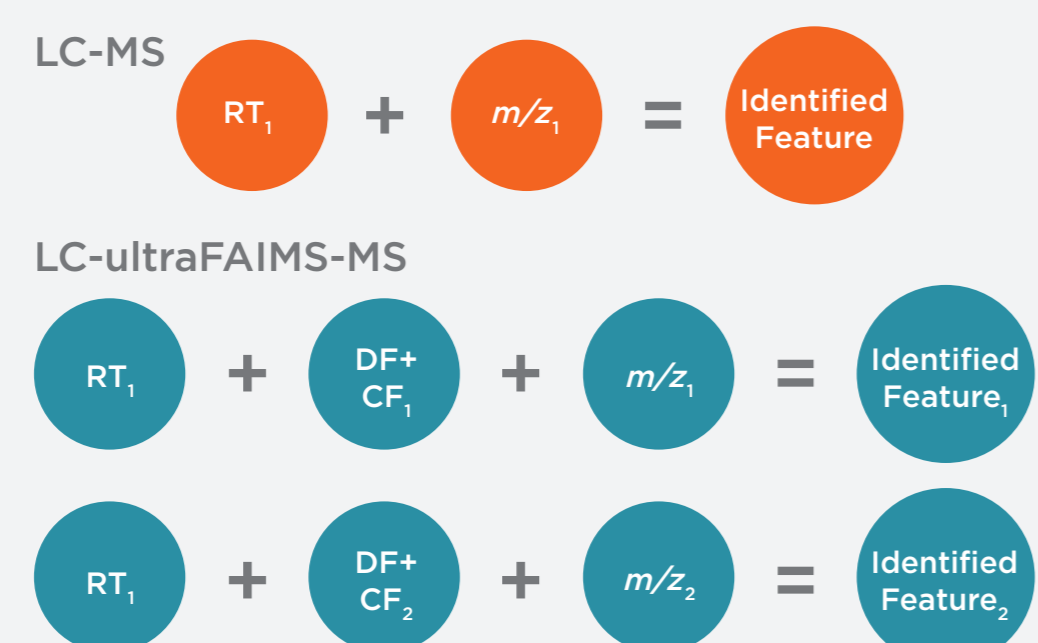


Figure 1. LC-MS and LC-ultraFAIMS-MS feature determination

- Retention time
  - FAIMS dispersion and compensation fields (DF and CF)
  - Mass-to-charge ( $m/z$ )
- A recent study<sup>1</sup> reported a threefold increase in features detected in non-targeted profiling of human urine with the addition of FAIMS to LC-MS analysis.

Here we describe an optimised workflow to produce three-dimensional metabolomics data sets and its application to the metabolic assessment of indeterminate masses using urine samples from a renal cell cancer (RCC) cohort (DIAMOND study; NRES East Of England O3/O18).

## 2. Methods

Analysis was performed on a 6230 TOF-MS and a 1290 series LC (Agilent, Santa Clara, US) combined with an ultraFAIMS device (Owlstone Medical Ltd, Cambridge, UK).

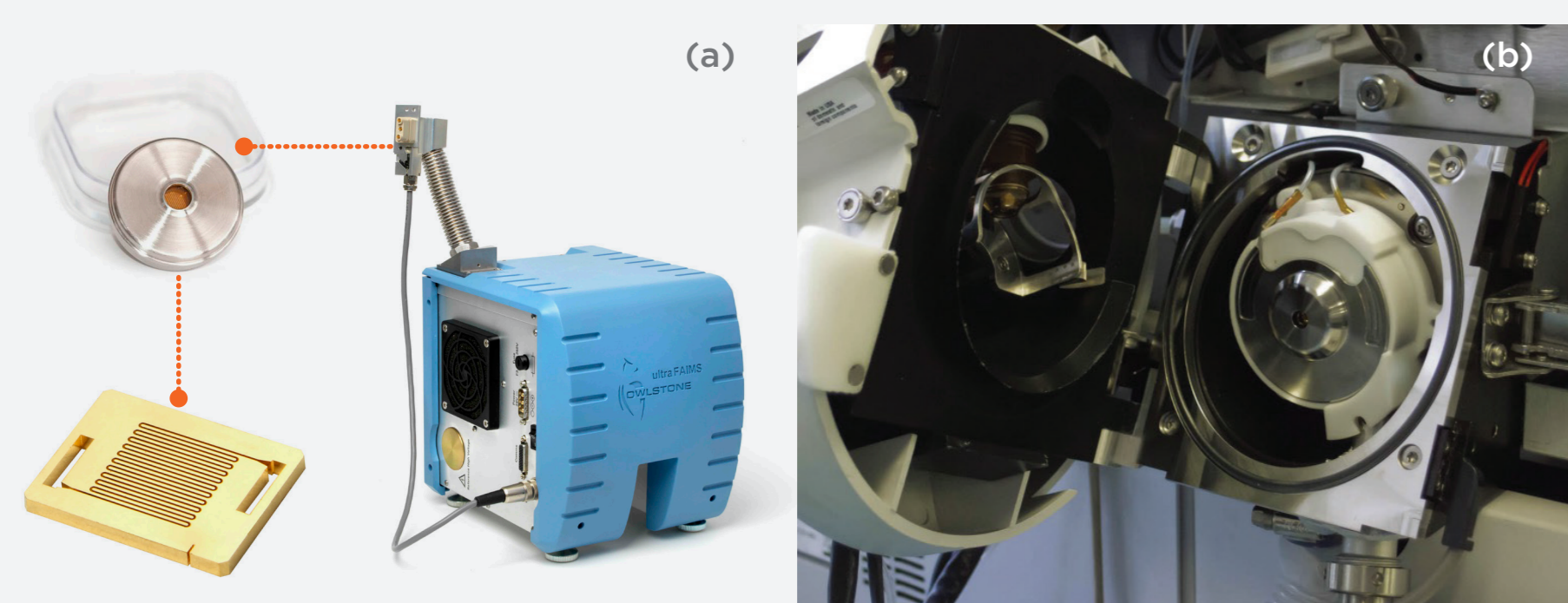


Figure 2. (a) ultraFAIMS chip and system (b) installed on Agilent 6230 ToF-MS

- The key dimensions of the ultraFAIMS device are the 100  $\mu\text{m}$  electrode gap and 700  $\mu\text{m}$  path length
- The small scale is key to the ability to integrate into the LC-MS workflow; an entire CF scan  $s^{-1}$  can be achieved, making the scanning approach compatible with chromatographic timescales.
- The CF values are synchronised with MS acquisition so a mass spectrum is acquired for each CF (Figure 3).
- LC was performed using a reversed phase Poroshell 120 EC-C18 column, 2.1 x 100 mm, particle size 2.7  $\mu\text{m}$  (Agilent Technologies).

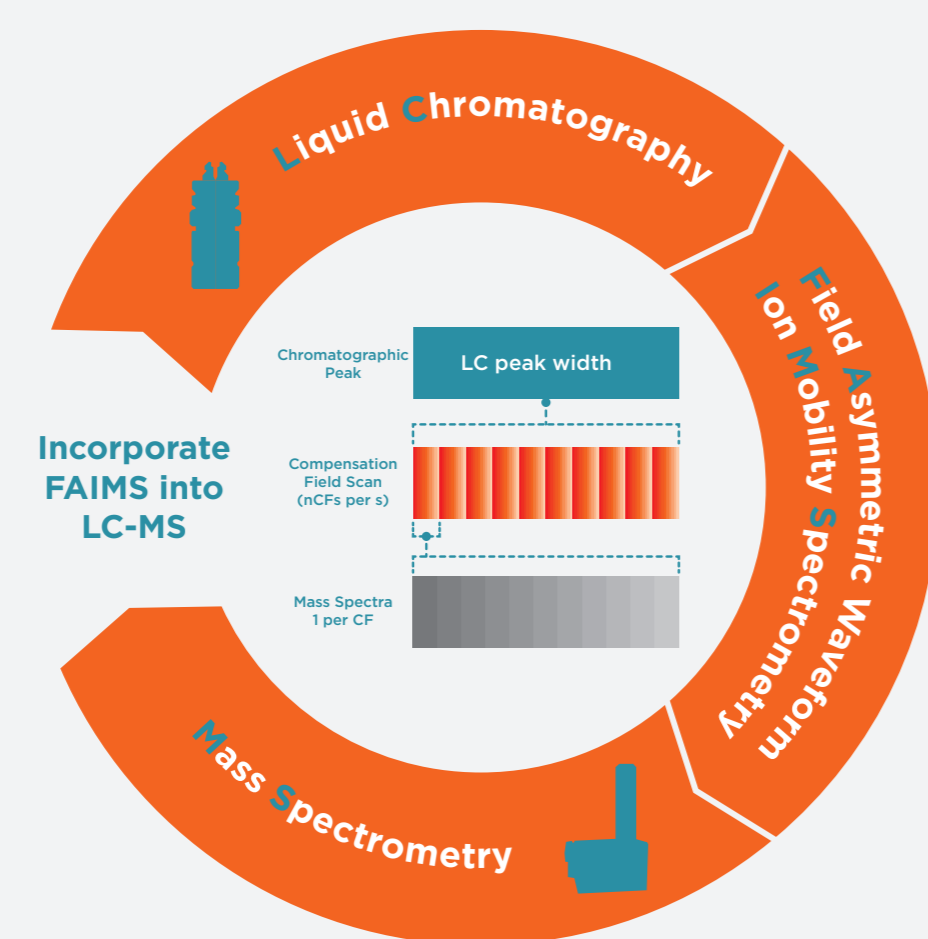


Figure 3. Integrated LC-ultraFAIMS-MS analysis

## 3. Data analysis

The workflow for feature extraction from LC-FAIMS-MS files is detailed in Figure 4.

A feature was defined as a unique identifier for each component of a  $m/z$ , a retention time (tR) and a compensation field (CF).

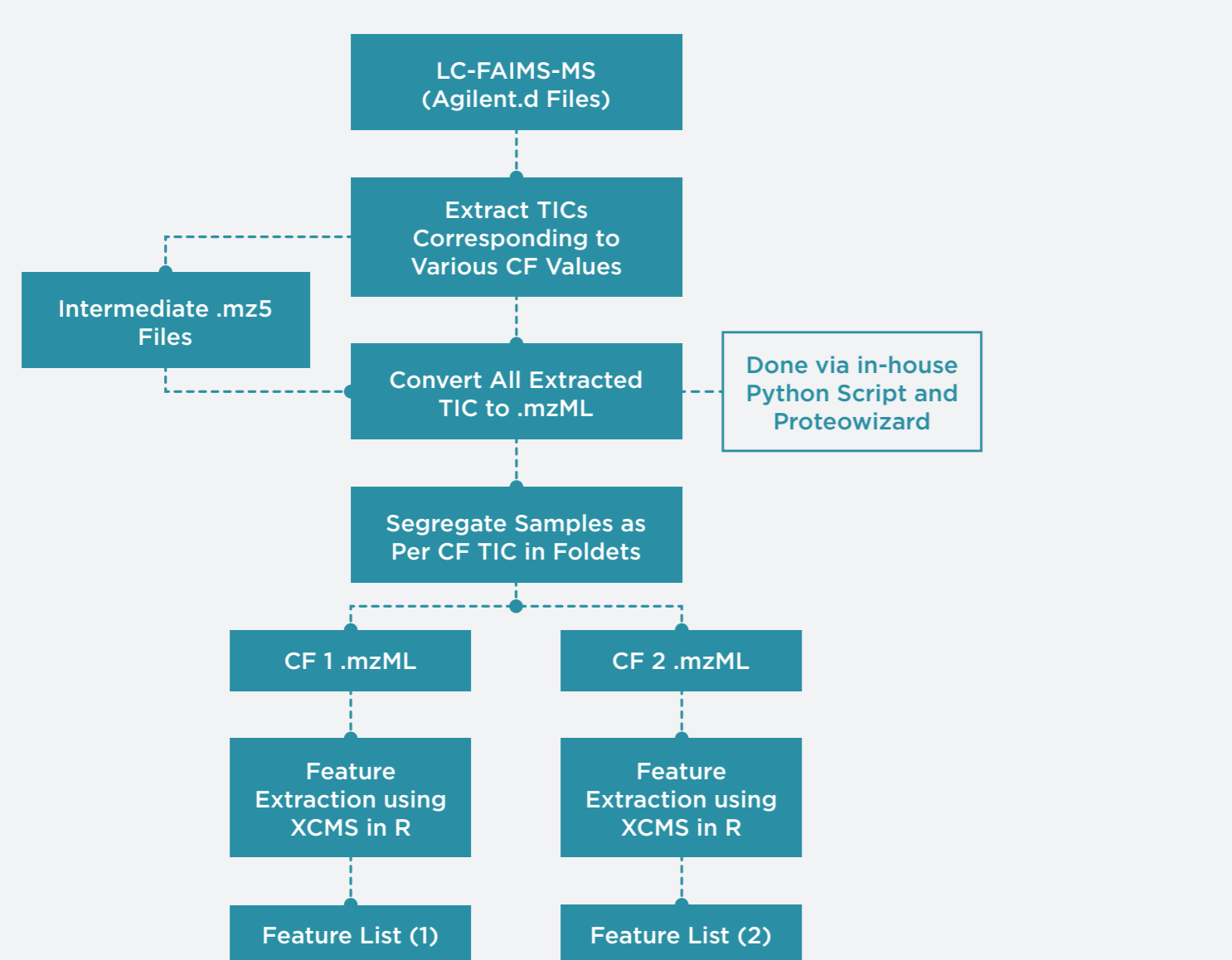


Figure 4. LC-ultraFAIMS-MS feature extraction workflow

## 4. LC-ultraFAIMS-MS optimisation

Optimisation of nested FAIMS data acquisition focussed on:

- # and range of FAIMS CF settings to give optimal FAIMS separation.
- # of data points within the timescale of the chromatographic peaks.
- Optimal sensitivity via chromatographic peak heights and # of TOF scans  $s^{-1}$ .

Table 1: LC-ultraFAIMS-MS optimisation experiments

Test	MS scan rate ( $s^{-1}$ )	per s	$m/z$ start	$m/z$ end	Index	Start CF (Td)	End CF (Td)	N CF steps	N CF actual	CF step size (Td)	N Repeats	Start DF (Td)	End DF (Td)	N DF Steps
1a	12	1	80	1500	0	-0.9	4.1	10	12	0.5	503	250	250	0
1b	12	1	80	1500	0	-0.9	4.1	10	12	0.5	503	240	240	0
2	24	2	80	1500	0	-0.9	4.1	10	12	0.5	1007	Test 1	Test 1	0
3a	10	1	80	1500	0	-0.9	3.1	8	10	0.5	503	Test 1	Test 1	0
3b	20	2	80	1500	0	-0.9	3.1	8	10	0.5	1007	Test 1	Test 1	0
3c	18	1	80	1500	0	-0.9	3.1	16	18	0.25	503	Test 1	Test 1	0
3d	6	1	80	1500	0	-0.9	3.1	4	6	1	503	Test 1	Test 1	0
3e	12	2	80	1500	0	-0.9	3.1	4	6	1	1007	Test 1	Test 1	0
3f	18	3	80	1500	0	-0.9	3.1	4	6	1	1511	Test 1	Test 1	0

Optimal DF was determined based on selectivity and sensitivity.

- The CF vs features plot (Figure 5) shows good coverage across the analytical space in the range -1 to +3 Td.
- More features were detected in the higher CF region at the higher DFs. 240 Td was selected for further experiments based on widest distribution of detected features across the CF range.

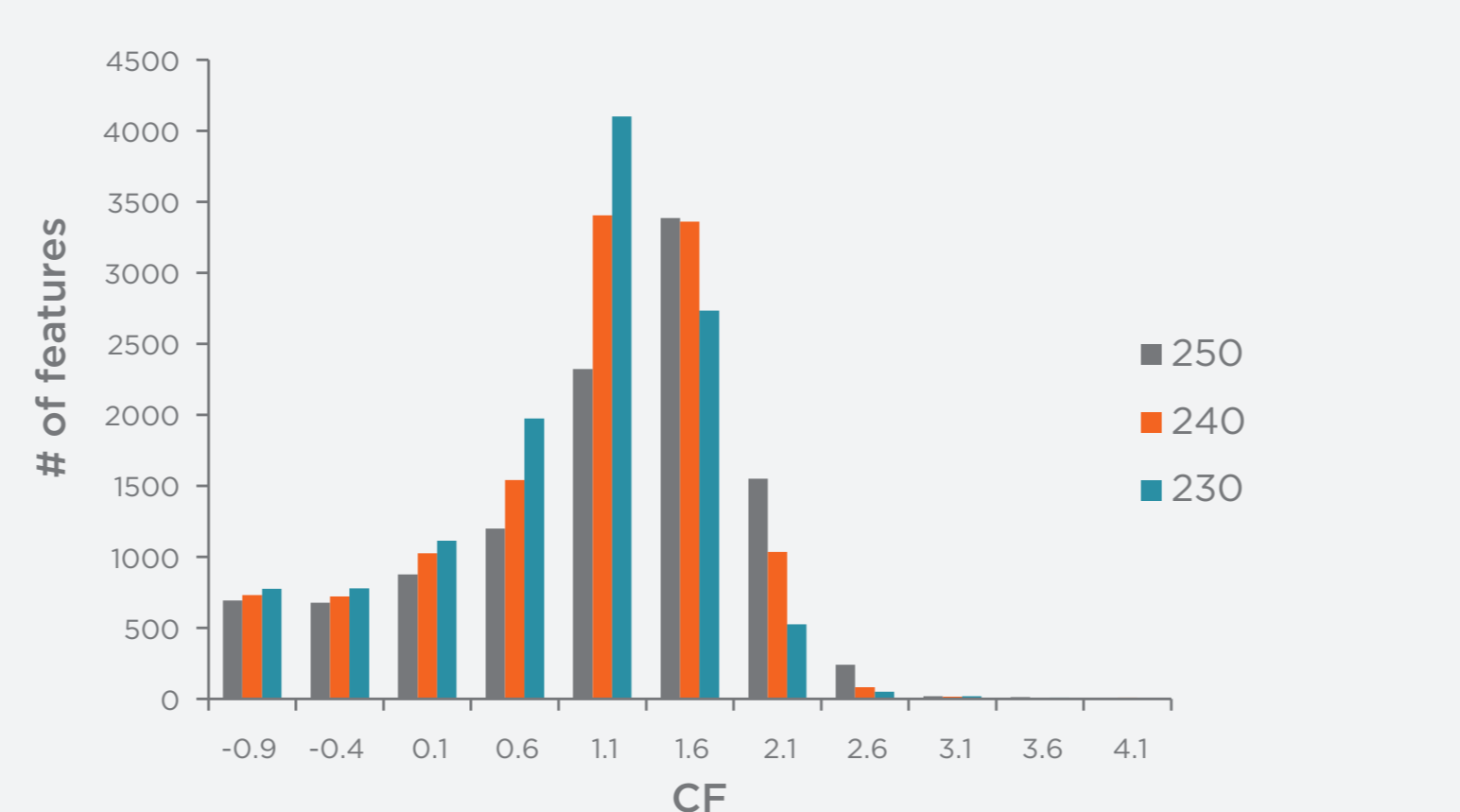


Figure 5. Detected features at DFs of 230, 240 and 250 Td

As the CF scan is synchronised with the MS acquisition, the # of data points across a chromatographic peak is dependent on the # of MS scans  $s^{-1}$  and the CF range.

- Increased MS scan rate means more data points can be acquired over a given CF range.
- Alternatively, a smaller CF step size could be applied, increasing the # of data points across the CF peak whilst maintaining the # of CF data points across LC peak.
- Faster ToF scan rates reduced peak intensity (Figure 6).

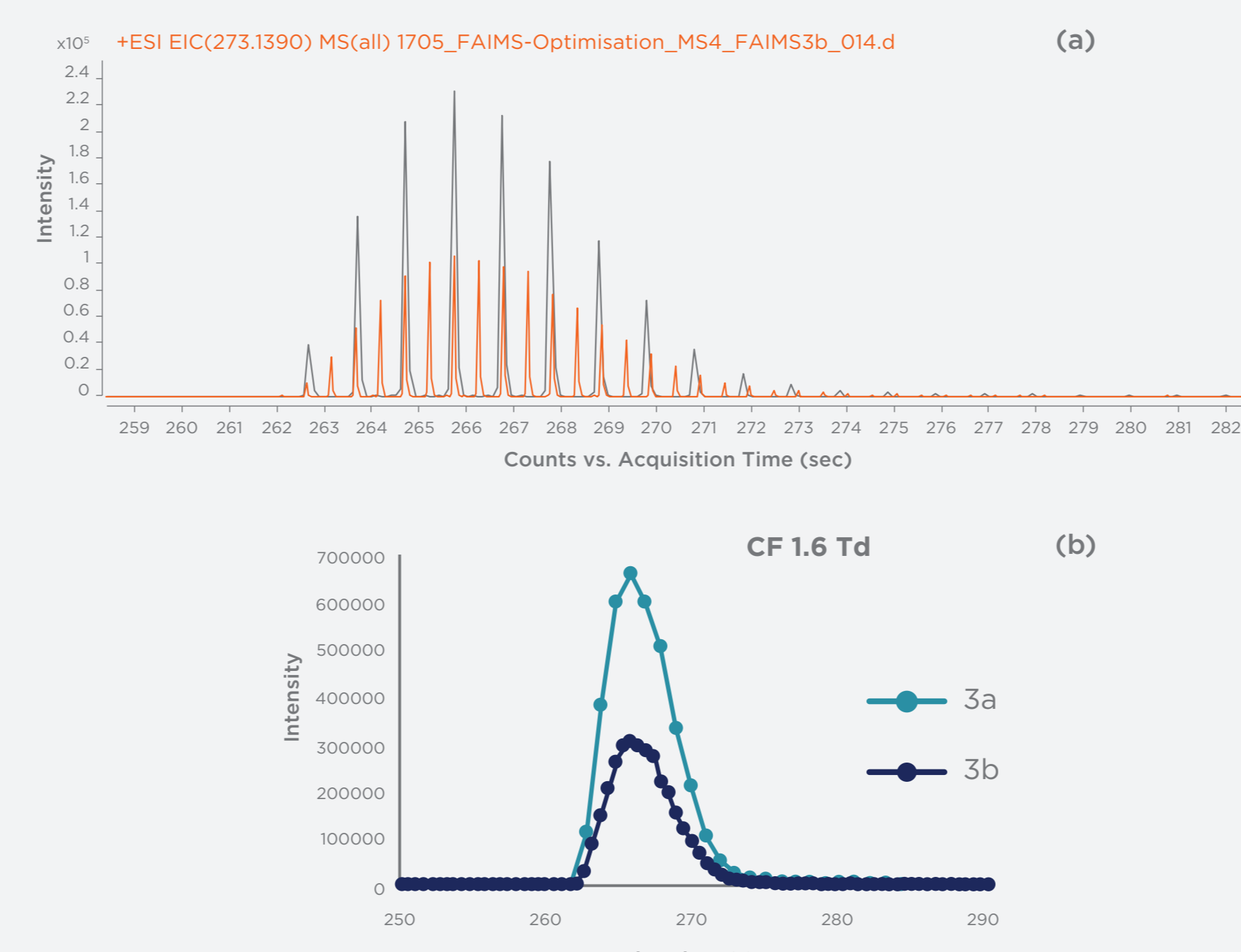


Figure 6. (a) Raw EIC for  $m/z$  273 across all CFs (black) 10 spectra  $s^{-1}$  and (red) 20 spectra  $s^{-1}$ , showing twice as many CF scans across the LC peak and (b) deconvoluted LC-MS peak at CF of 1.6 Td at 20 spectra  $s^{-1}$

The effect of the different MS scan rates and CF ranges on feature detection was investigated.

- Acquiring data at 2 CF scans  $s^{-1}$  increased the # of features detected, in all cases, despite the decrease in peak intensity associated with higher acquisition speeds (Figure 7b, c).
- Increasing to 3 CF scans  $s^{-1}$  did not further increase the # of features detected.

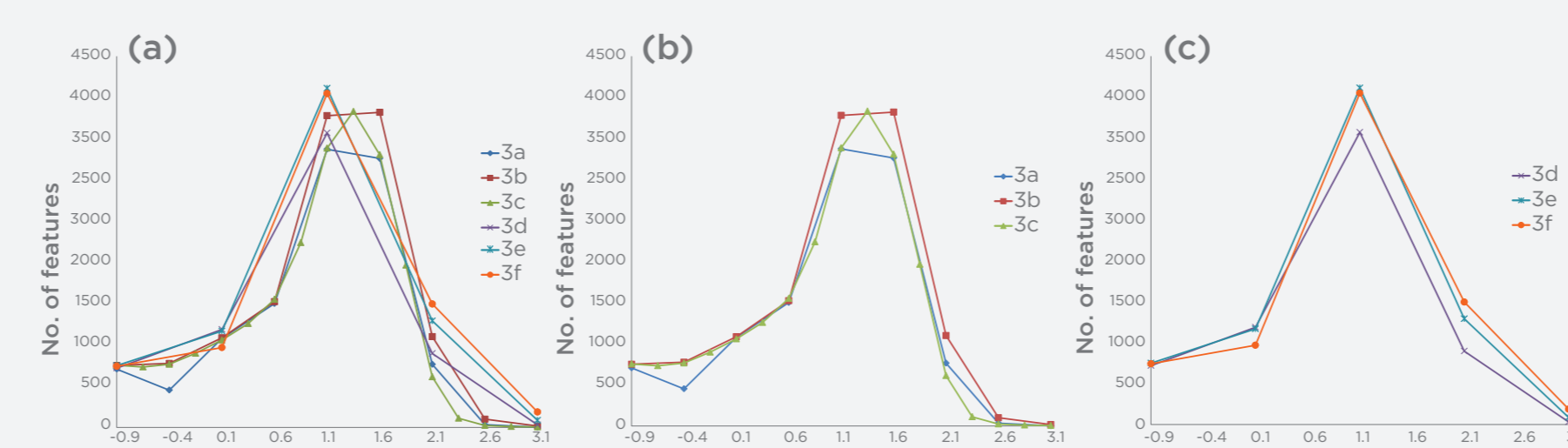


Figure 7. Comparison of # of features detected in experiments 3 a-f as described in Table 1

The # of features detected by scanning the FAIMS at a rate of 1 CF scan  $s^{-1}$  and 2 CF scans  $s^{-1}$  was compared.

- At a S:N of 3, more features were detected using a 2 CF scans  $s^{-1}$  scan rate (Figure 8a).

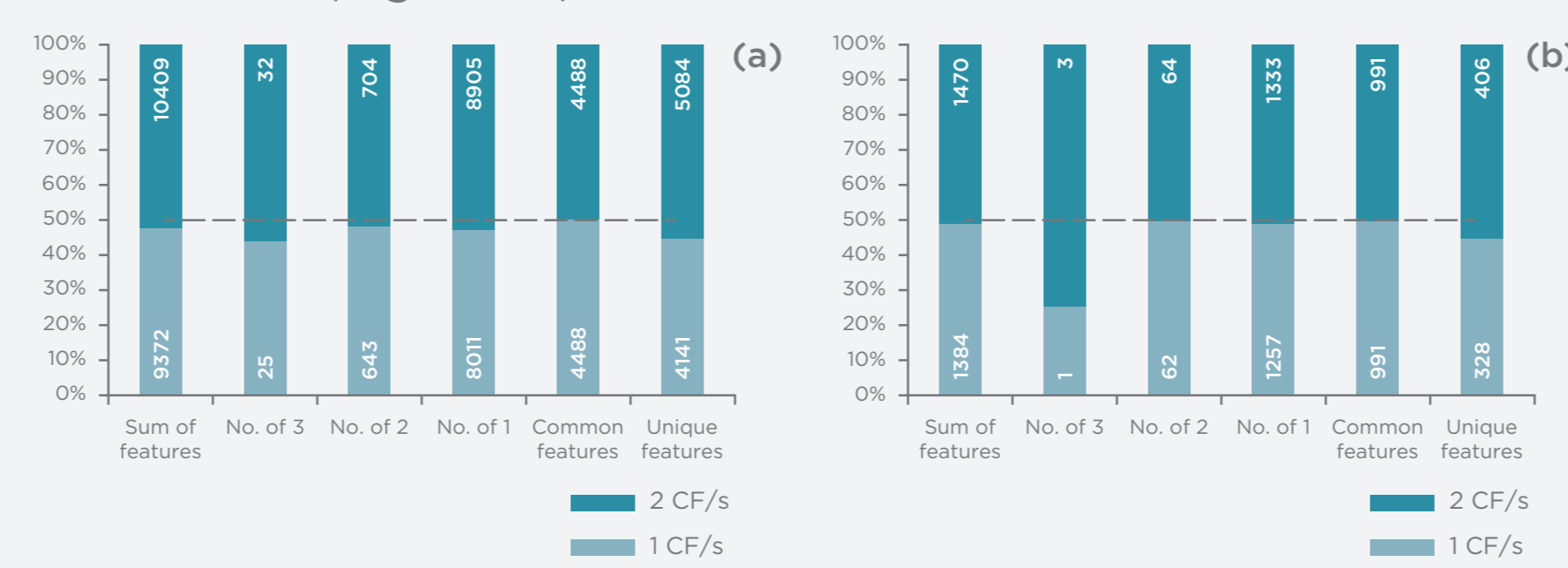


Figure 8. Comparison of features detected at 1 CF scan  $s^{-1}$  and 2 CF scans  $s^{-1}$  at (a) S:N of 3 and (b) S:N of 10

To determine if this result was accurate, or a result of increased noise using the faster scan rate, the analysis was repeated at a S:N of 10 (Figure 8b).

- Whilst total # of features decreased, the faster scan rate again produced more features than the slow scan rate.
- 2 CF scans  $s^{-1}$  scan rate was used for analysis.

## 5. LC-ultraFAIMS-MS Workflow Validation

The optimised FAIMS scan settings of CF -0.9 to 3.1 Td at DF 240 Td at 2 CF scans  $s^{-1}$  were used to analyse:

- n=10 urine samples from renal cancer patients
- n=10 urine samples from healthy individuals

Raw spectra were used to generate features lists for each CF using XCMS using parameters:

- method="centWave", scanrange=c(0,504), peakwidth=c(5,30), ppm=25, mzdiff=c(-0.001), snthresh=3, integrate=2, fitgauss=FALSE

Produced feature-extracted and feature-aligned data.

An unsupervised, pareto scaled principal component analysis (PCA) plot for data obtained at CF 1.1 Td is shown in Figure 9.

- Pareto scaling used for the analysis as consistent with published untargeted metabolomic studies<sup>2</sup>

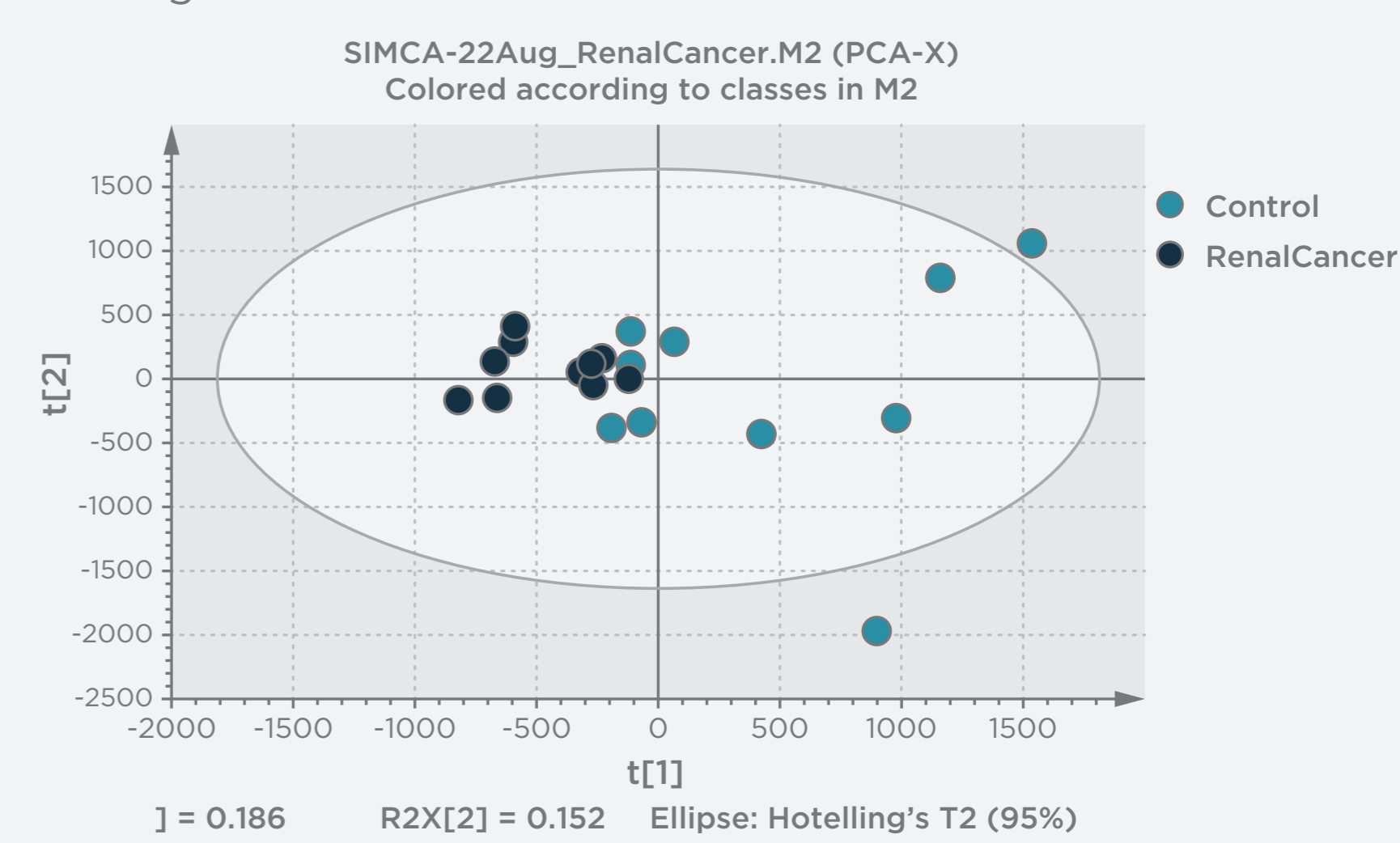


Figure 9. Unsupervised PCA (Pareto Scaling) plot at 1.1 Td CF

The PCA plot shows partial separation indicating differences in the data sets.

A supervised o-partial least squares-discriminant analysis (OPLS-DA), highlighting the within class and between class differences, also showed separation (Figure 10).

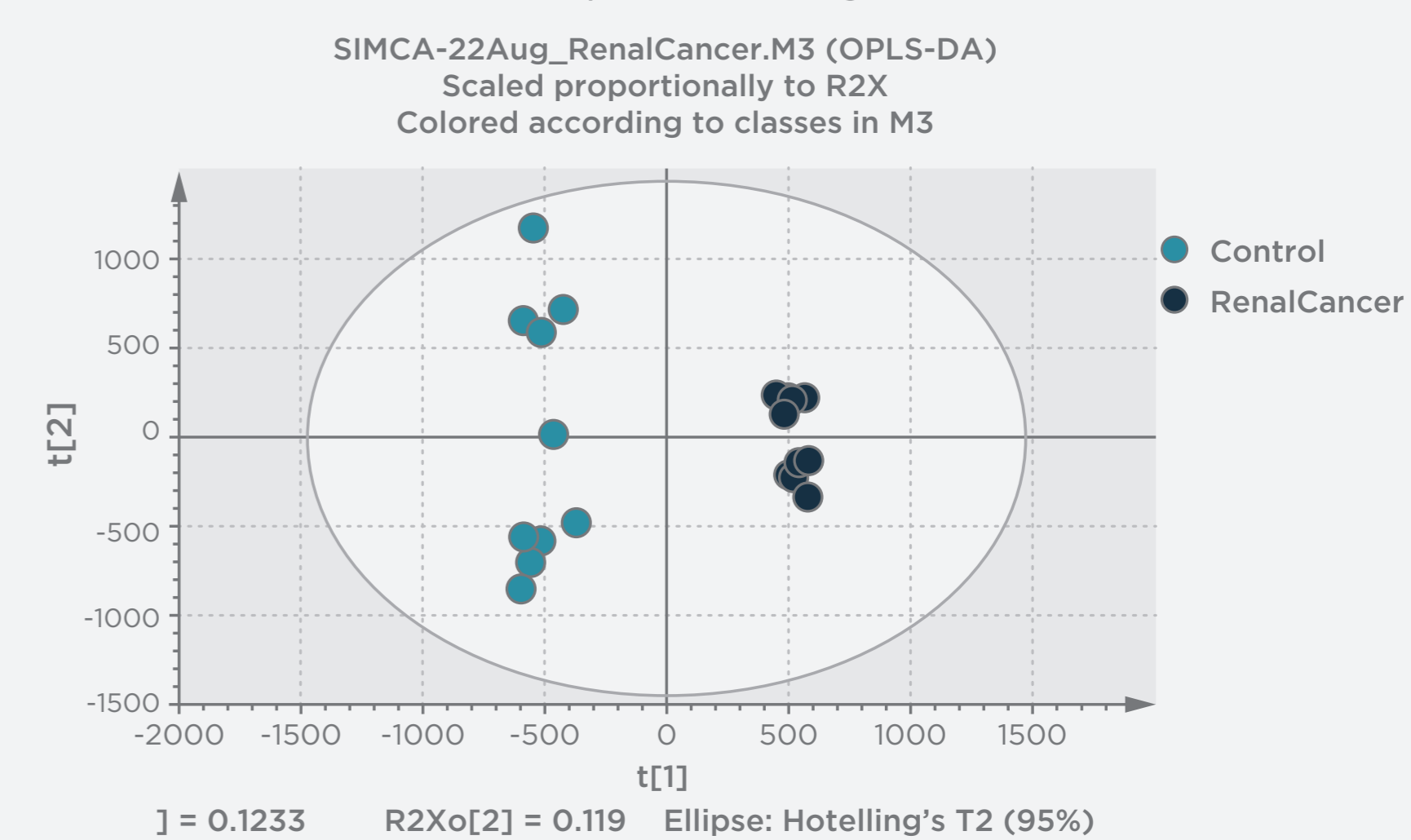


Figure 10. OPLS-DA plot at 1.1 Td CF

The S-plot, a plot of covariance and correlation based on the features making up the OPLS-DA model was also plotted.

- Trend plots for features contributing towards class separation with the most confidence were plotted (Figure 11).

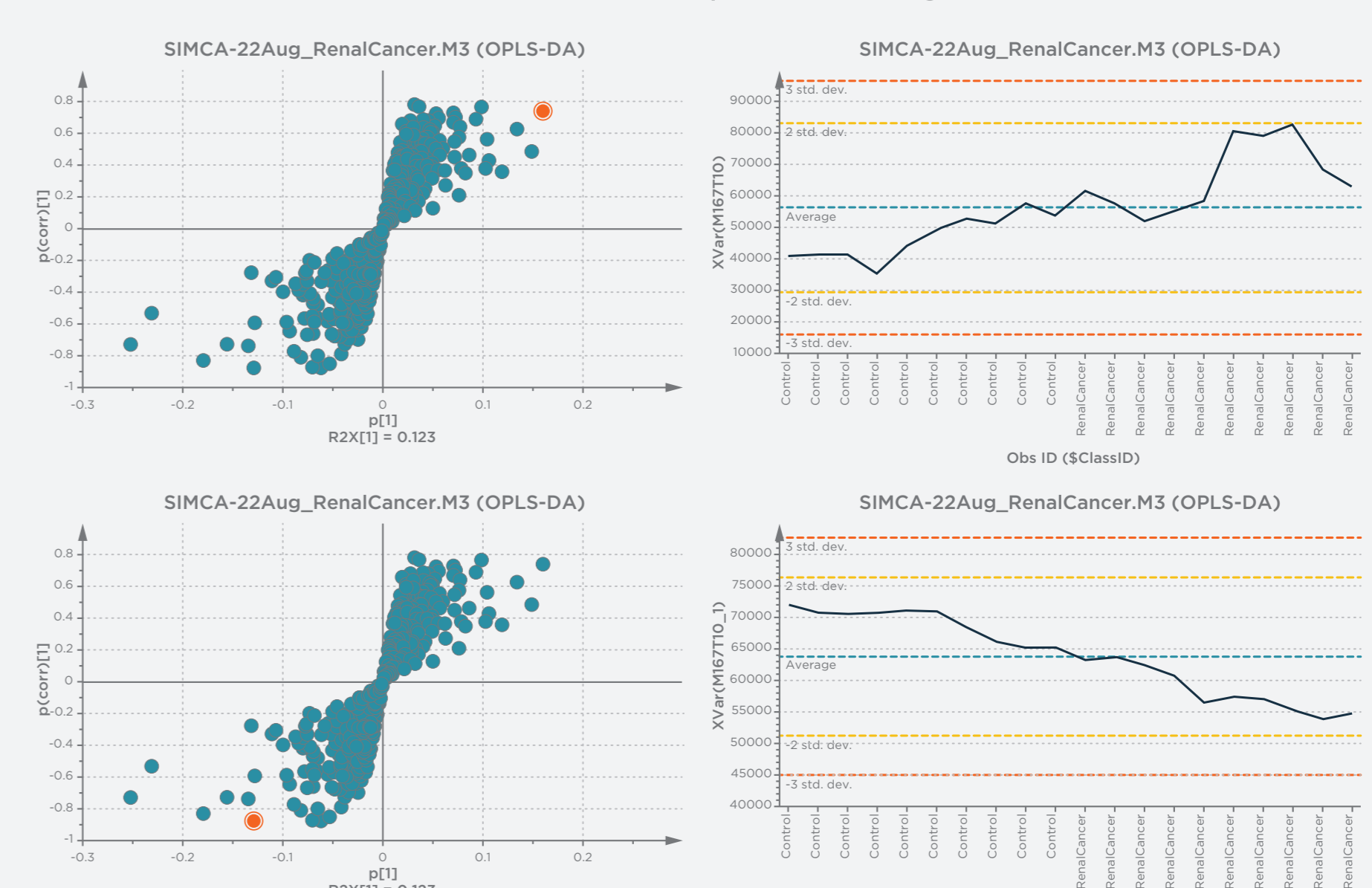


Figure 11. Example Trend plots of features contributing towards class separation

Next step is to expand the data set to a statistically significant # of patients and validate the identified features by correlating multivariate analysis to raw data, as well as cross referencing to patient data to identify potential confounders.

## 5. Conclusions

- The optimisation of untargeted profiling of the urinary metabolome using LC-FAIMS-MS and the data analysis workflow required to produce 3D nested data sets has been described.
- The optimised workflow has been applied to demonstrate proof of principle metabolic assessment of urine samples from a renal cell cancer cohort vs. healthy controls.
- Following further validation, the LC-ultraFAIMS-MS features list will be used to build a predictive model, or classifier, from which a probability of disease can be predicted.
- Combining FAIMS into a metabolomic workflow offers increased peak capacity for untargeted metabolite profiling to diagnose disease or stratify an individual's treatment.

## References

1. Arthur, KA; Turner, MA; Reynolds, JC; Creaser, CS; *Anal. Chem.*, **2017**, *89*, 3452
2. van den Berg, RA; Hoefsloot, H CJ; Westerhuis, JA; Smilde, AK; van der Werf, MJ; *BMC Genomics*, **2006**, *7*, 142