Supplementary Information

Volatile atmospheric pressure chemical ionisation mass spectrometry headspace analysis of *E. coli* and *S. aureus*

Kerry Rosenthal[1], Eugenie Hunsicker[2], Elizabeth Ratcliffe[3], Martin R. Lindley[1,4], Joshua Leonard[5], Jack R. Hitchens[5], Matthew A Turner[4,5].

[1] School of Sport, Exercise & Health Sciences, Loughborough University, Loughborough, UK
[2] Department of Mathematical Sciences, Loughborough University, Loughborough, UK
[3] Department of Chemical Engineering, Loughborough University, Loughborough, UK
[4] Translational Chemical Biology Research Group
[5] Department of Chemistry, Loughborough University, Loughborough, UK

E-mail: m.r.lindley@lboro.ac.uk

Contents

Figure S1. Total intensity count over the 5-min sampling period for three randomly chosen samples for each class in each experiment; dataset 1 5-min samples on normal agar, dataset 2 5-min samples on normal agar, and dataset 2 5-min samples on blood agar. Colours are simply to show the three separate samples for each class.
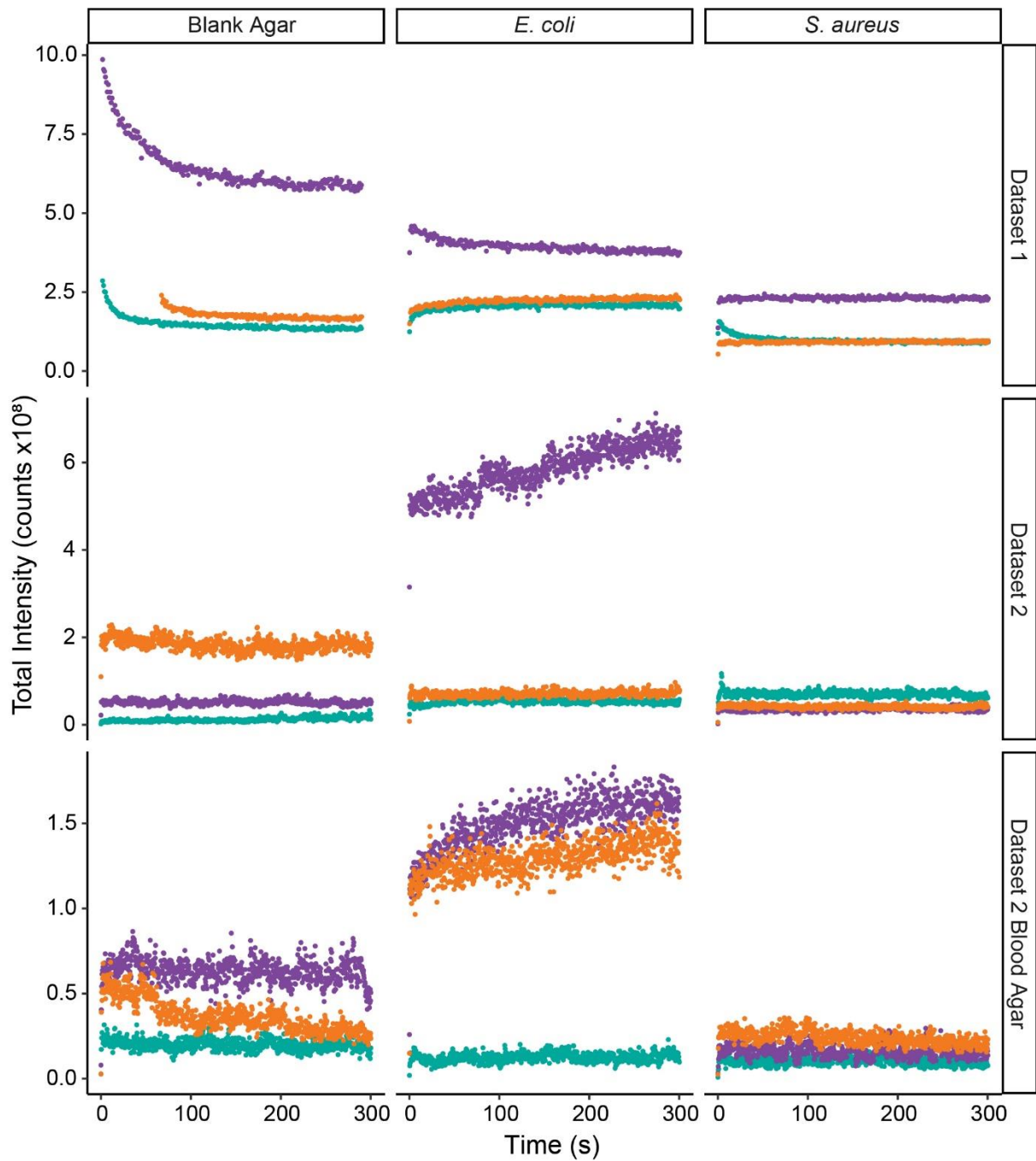
Figure S2. The *m/z* 118 intensity count for each scan over the 5-min sampling period for an example *E. coli* sample. One-minute's worth of scans (60/scan rate) were summed. The turquoise points show the first set of scans selected. The values for those same scans were summed separately for each *m/z* bin. Then the process was repeated 200 times (orange points, purple points; this figure only shows three times). The scans were selected entirely at random each time and the same scans could be selected more than once.
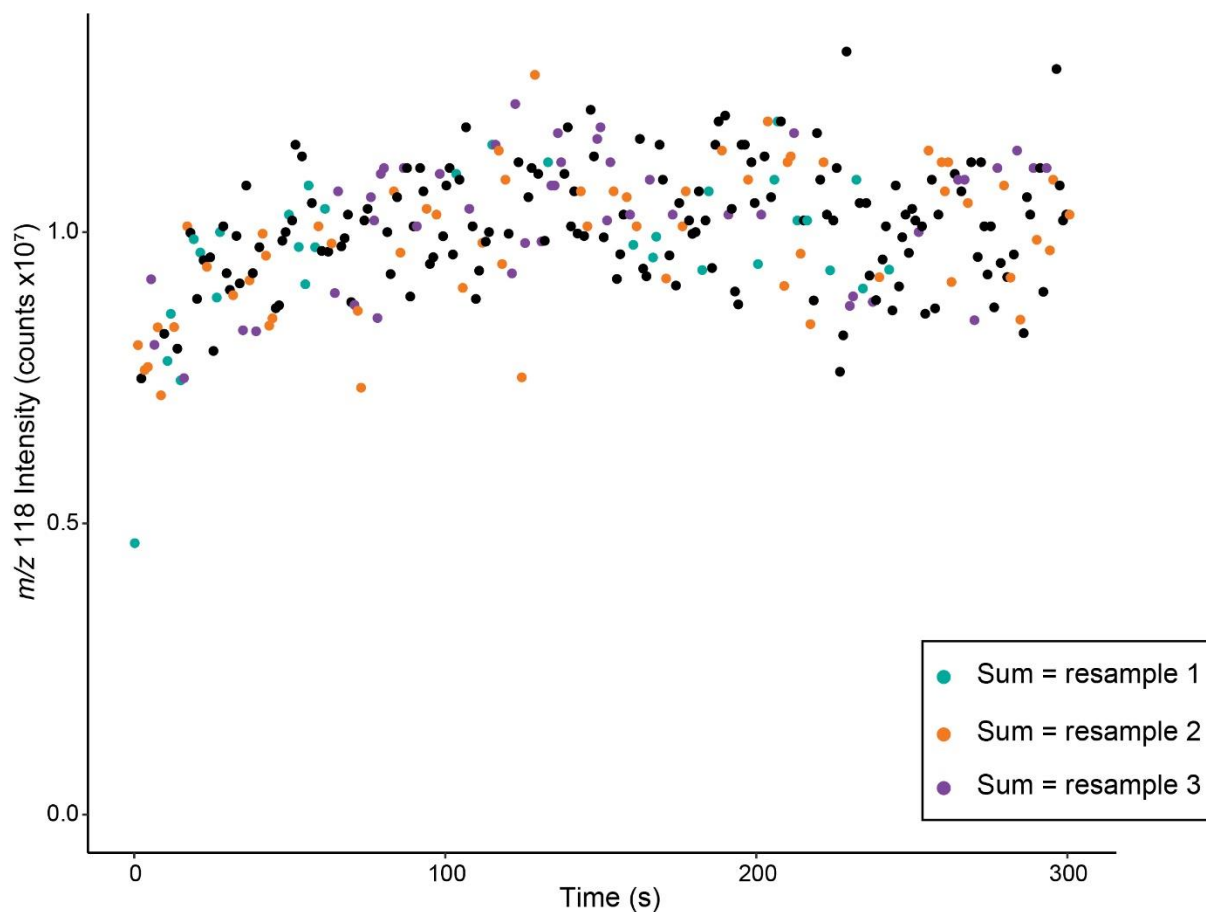
Figure S3. Mean (SD) of E. coli (black triangles), S. aureus (black crosses), and blank agar (white circles) intensity counts of dataset 2 samples for m/z 118 and m/z 231.
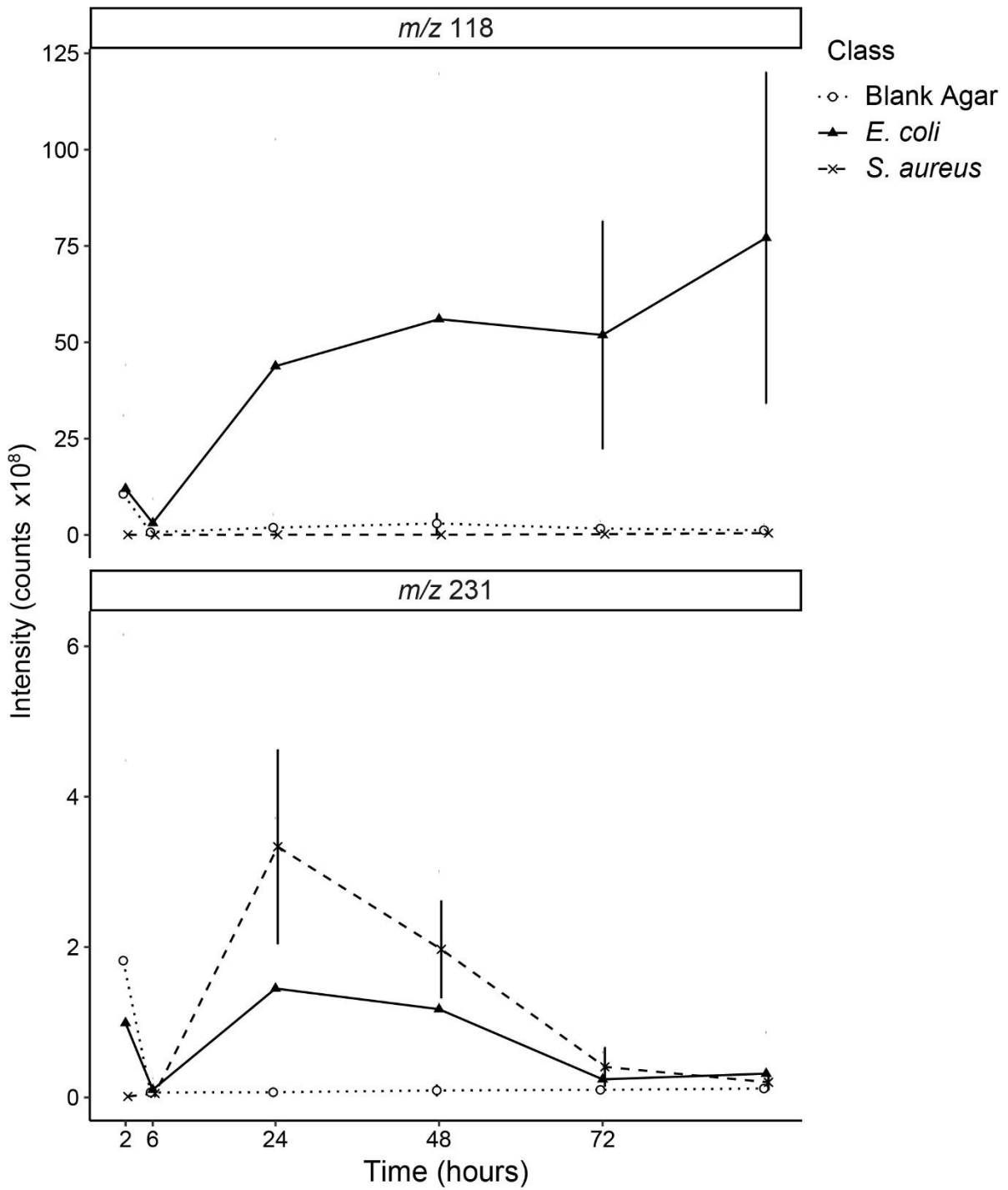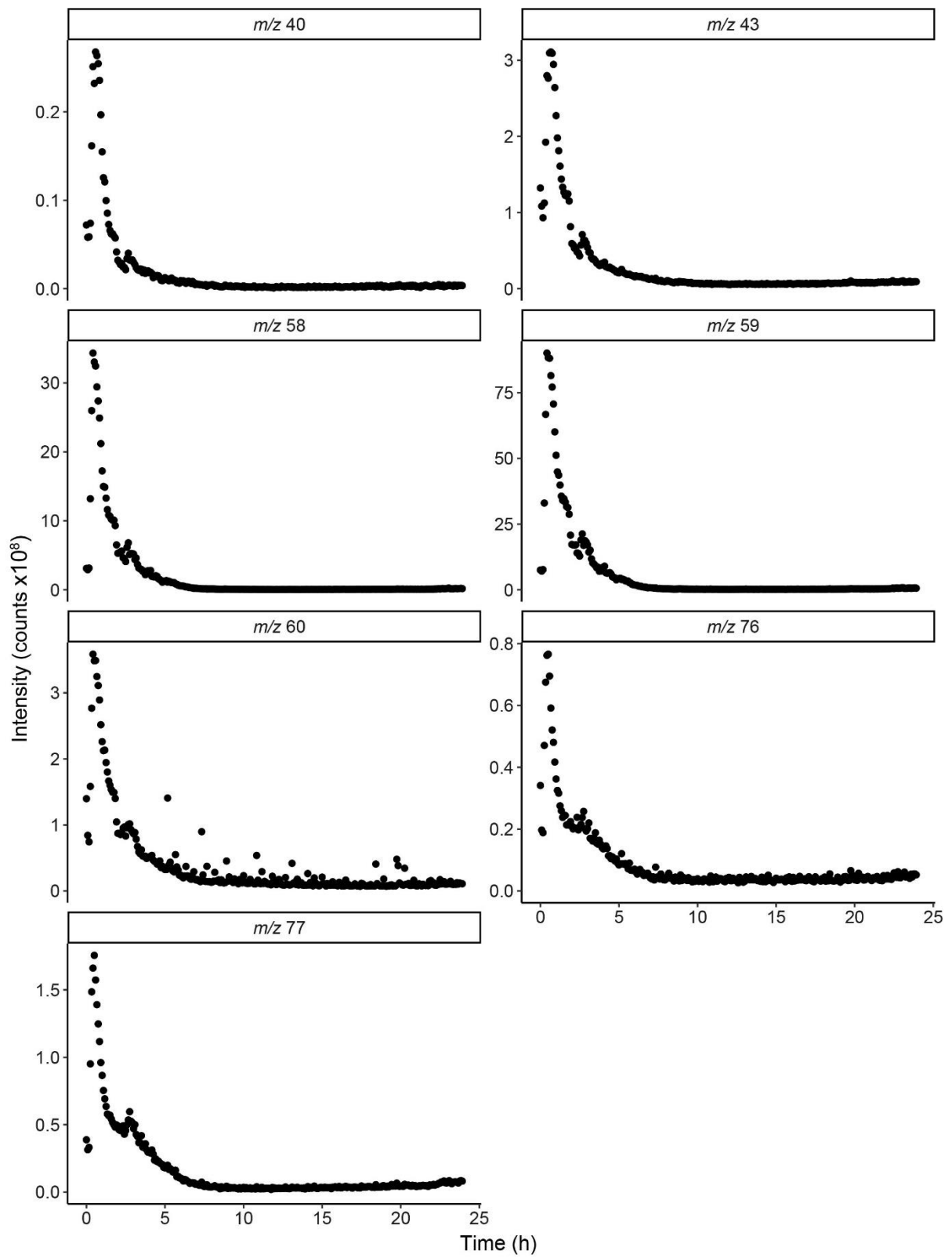
Figure S4a. Selected variables (*) for each of the four *E. coli* datasets, with the highest median signal-to-noise ratio across time points indicated by the gradient.
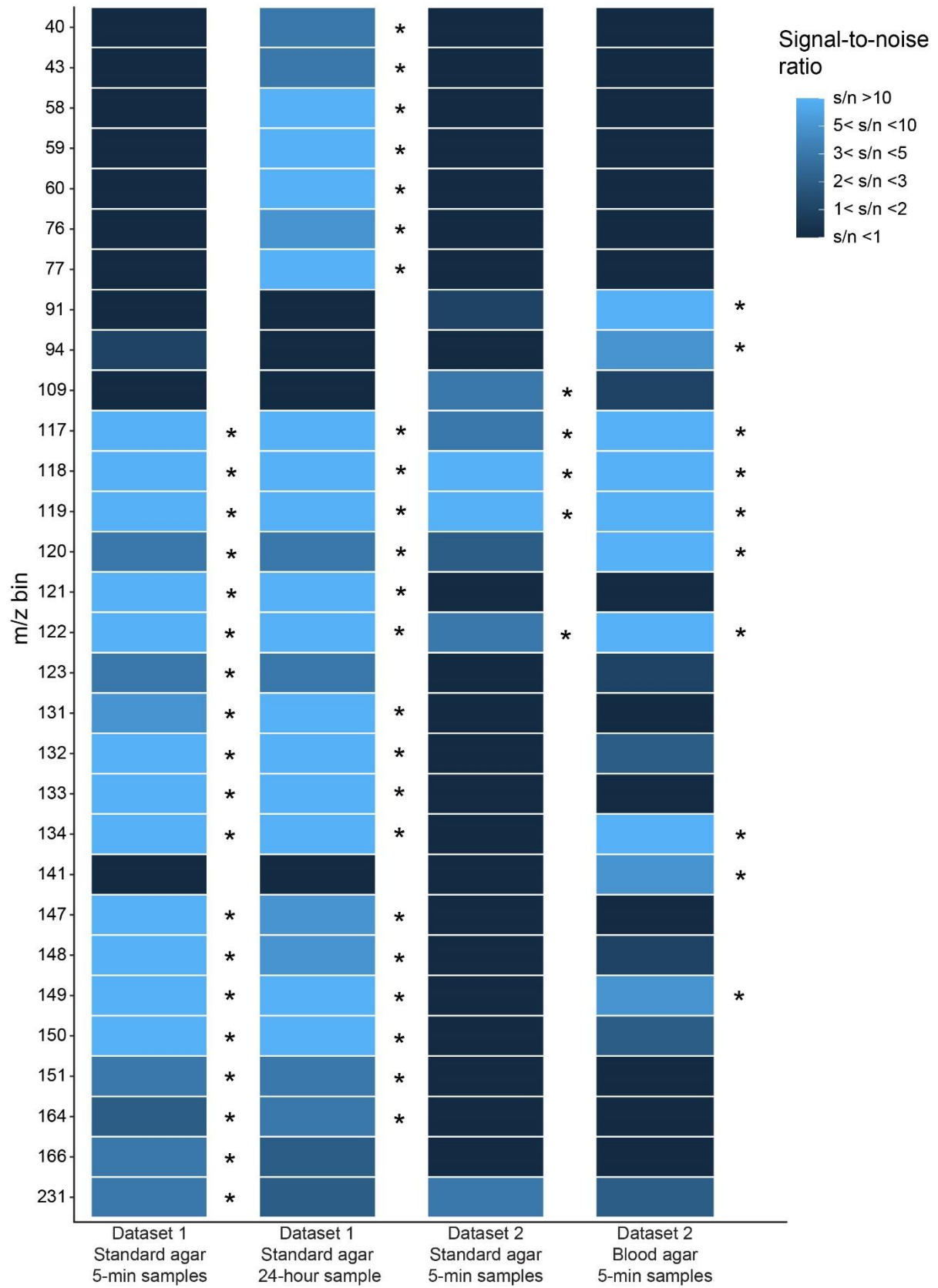
Figure S4b. Selected variables (*) for each of the four *S. aureus* datasets, with the highest median signal-to-noise ratio across time points indicated by the gradient.
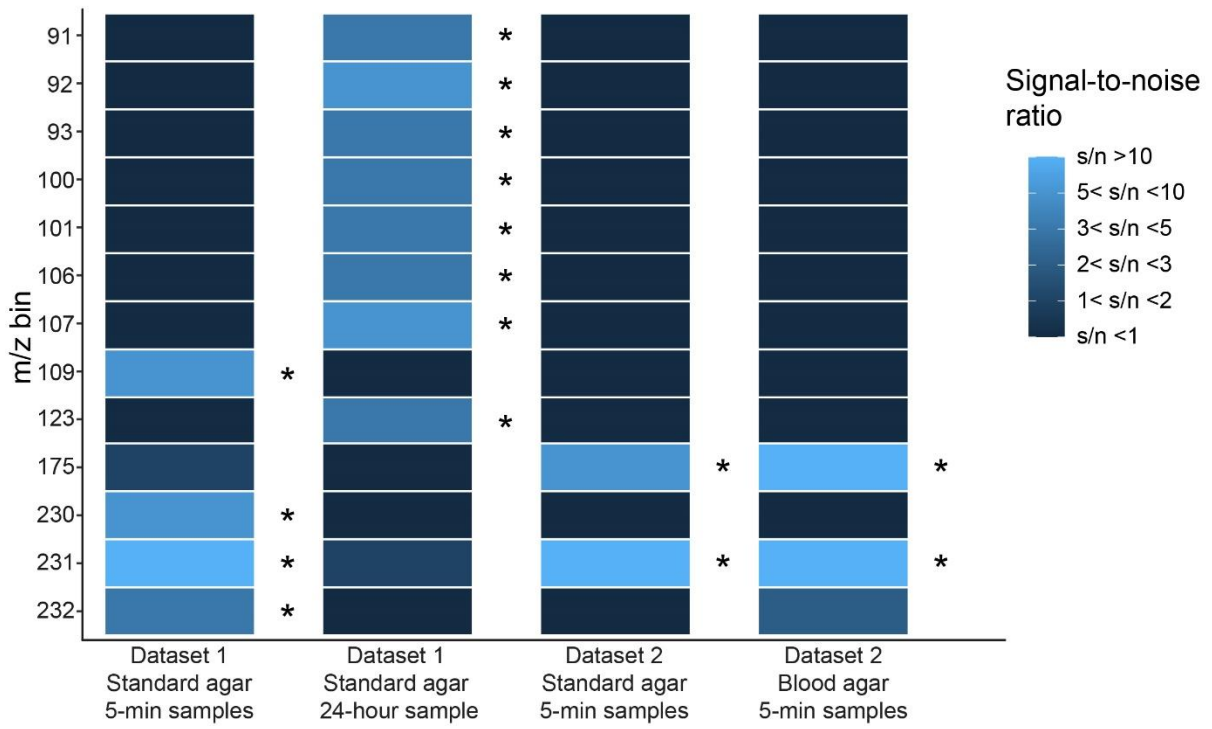
Figure S5a. Plots of the 24-hour continuous *E. coli* sample (mean of the resamples for each 5 minute section) for the variables (labelled above each plot) not selected by any other method.

Figure S5b. Plots of the 24-hour continuous *S. aureus* sample (mean of the resamples for each 5 minute section) for the variables (labelled above each plot) not selected by any other method.
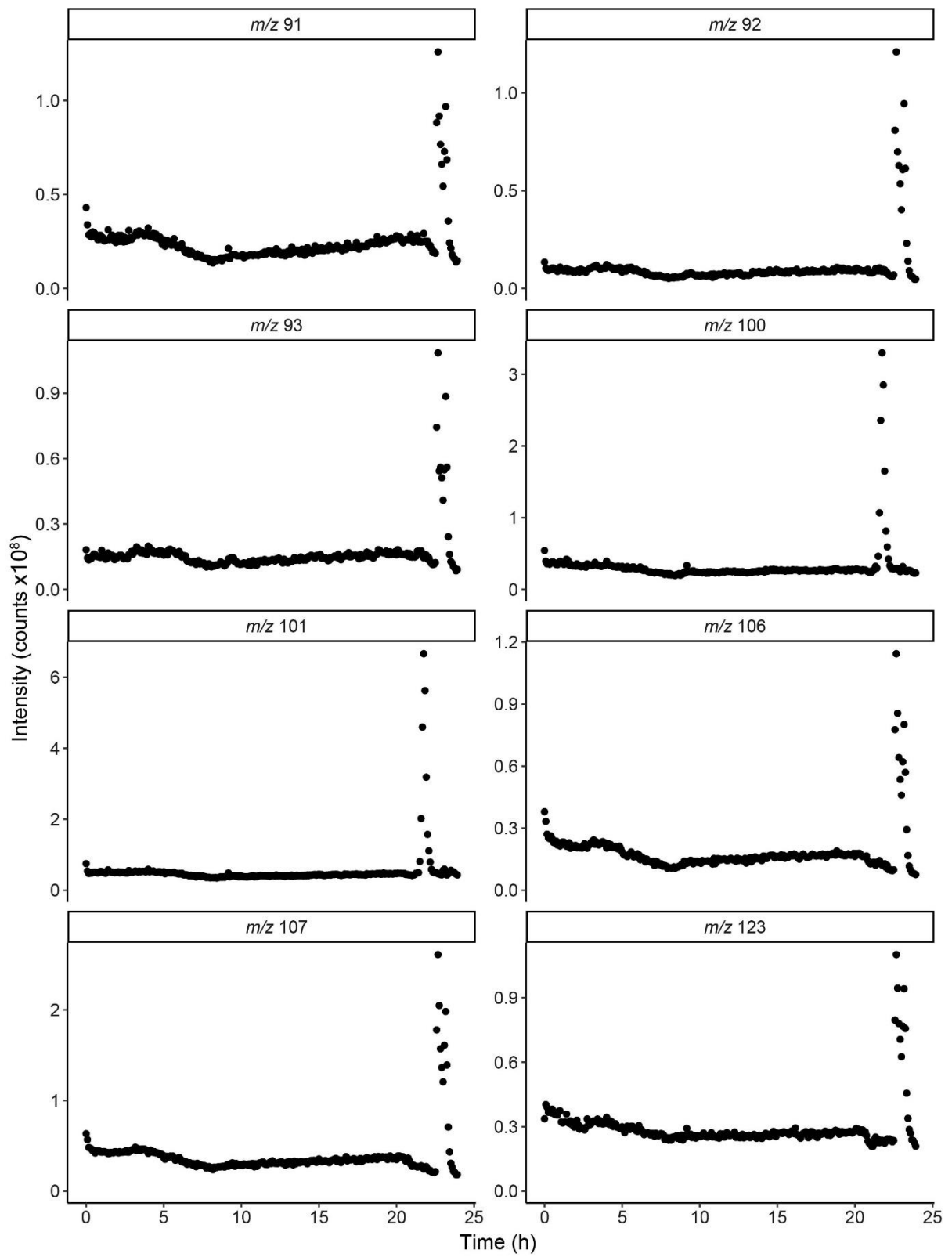
Table S1. Coefficients for the classification model. Values are rounded to two decimal places.

	Intercept	*m/z* 118	*m/z* 231
E. coli	4.17	5.36	7.84
S. aureus	-726.07	-1308.57	27.58

Figure S6. Youden's index for the classification model, separated by set, time and class (white circles: blank agar, black triangles: *E. coli*, black crosses: *S. aureus*) with a value of one indicating perfect identification, with no false negatives or false positives.