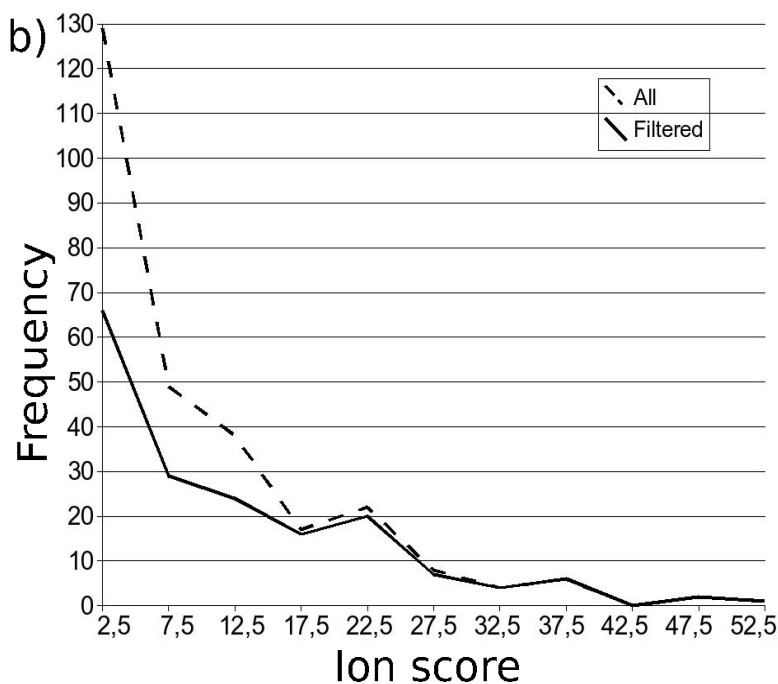
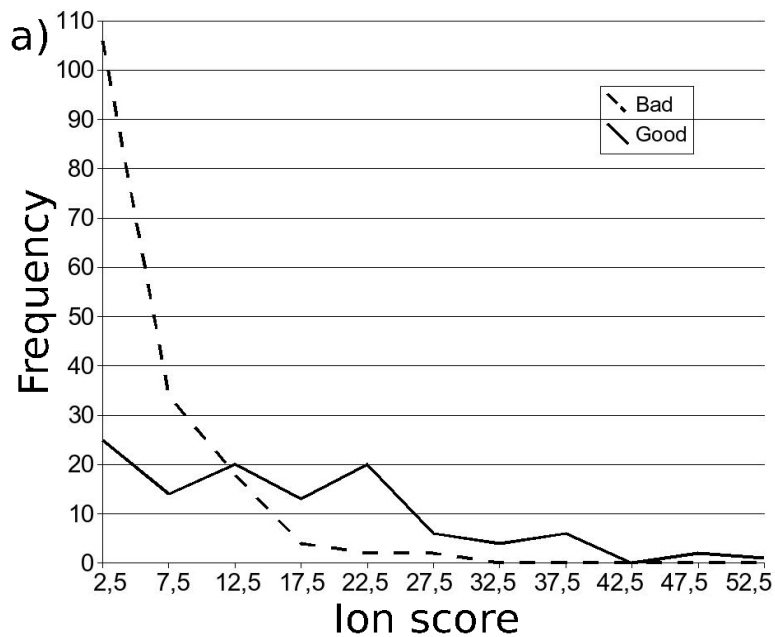


Supplementary material to article "Quality classification of tandem mass spectrometry data" by Jussi Salmi, Robert Moulder, Jan-Jonas Filén, Olli S. Nevalainen, Tuula A. Nyman, Riitta Lahesmaa and Tero Aittokallio.



Supplementary figure. (a) Mascot ion score distribution in the test material 5 for the manually validated "bad" (dotted line) and "good" (continuous line) spectra at FPR=90% and (b) for the whole (dotted line) and filtered (continuous line) set of spectra. It can be seen from the figure, that the manual validation agrees fairly well with the Mascot score, and that the filtering reduces the number of spectra with lower ion scores. The ion score was available only for spectra, which were identified by Mascot (118 out of 130 "good" spectra and 214 out of 449 "bad" spectra were identified in this material). This makes the figure somewhat biased, as compared with Figure 3, for evaluating the manual validation and filtering processes.

Rules used by the C4.5 decision tree in classifying the test materials with a combined feature set F_1 to F_7 and B_1 to B_7 .

F16 SEQUEST

B7 <= 0.077324: 1 (695.48/7.74)
B7 > 0.077324
| B3 <= 0.077671: 1 (66.0/0.5)
| B3 > 0.077671: 0 (10.52)

F16 ProICAT

F5 <= 0.03369
| B7 <= 0.000652: 1 (19.54/1.44)
| B7 > 0.000652: 0 (11.97)
F5 > 0.03369: 1 (740.5/16.66)

F17 SEQUEST

F3 <= 0.090532
| B4 <= 0.59212
| | B5 <= 0.410918
| | | B2 <= 0.000809: 1 (5.0/0.39)
| | | B2 > 0.000809: 0 (6.58)
| | B5 > 0.410918: 1 (35.51/3.29)
| B4 > 0.59212: 0 (12.1)
F3 > 0.090532
| F3 <= 0.360055: 1 (338.62/16.44)
| F3 > 0.360055
| | B4 <= 0.418266: 0 (3.81)
| | B4 > 0.418266: 1 (15.39/1.58)

F17 ProICAT

F3 <= 0.067175
| B7 <= 0.045174: 1 (8.43/0.94)
| B7 > 0.045174: 0 (20.45)
F3 > 0.067175
| F3 <= 0.283934
| | F6 <= 0.011339
| | | B7 <= 0.203653
| | | | B7 <= 0.08788
| | | | | F7 <= 0.001664: 1 (7.65/0.16)
| | | | | F7 > 0.001664: 0 (3.28)
| | | | B7 > 0.08788: 1 (18.89/0.16)
| | | B7 > 0.203653: 0 (3.59)
| | F6 > 0.011339: 1 (310.84/11.08)
| F3 > 0.283934
| | F3 <= 0.451524
| | | F6 <= 0.215185: 1 (37.94/4.22)
| | | F6 > 0.215185: 0 (2.34)
| | F3 > 0.451524: 0 (3.59)

F18 SEQUEST

```
F3 <= 0.072383
| F3 <= 0.044543: 0 (7.51)
| F3 > 0.044543
| | F1 <= 0.056023: 0 (3.24)
| | F1 > 0.056023
| | | F5 <= 0.117171: 1 (21.1/1.04)
| | | F5 > 0.117171
| | | B5 <= 0.757112: 0 (4.01)
| | | B5 > 0.757112: 1 (4.14/0.13)
F3 > 0.072383
| F3 <= 0.257795: 1 (455.2/25.76)
| F3 > 0.257795
| | F1 <= 0.113127: 0 (8.16)
| | F1 > 0.113127
| | | B3 <= 0.314321: 1 (16.18/0.13)
| | | B3 > 0.314321: 0 (2.46)
```

F18 ProICAT

```
F6 <= 0.015754
| B2 <= 0.471867: 0 (39.23)
| B2 > 0.471867
| | F8ladder <= 0.009881
| | | B7 <= 0.156093
| | | | F3 <= 0.154835: 0 (3.16)
| | | | F3 > 0.154835: 1 (2.85)
| | | | B7 > 0.156093: 1 (11.39)
| | | F8ladder > 0.009881: 0 (5.06)
F6 > 0.015754
| F3 <= 0.258804
| | B3 <= 0.086317
| | | B7 <= 0.424875
| | | | B6 <= 0.722367: 1 (19.61/2.53)
| | | | B6 > 0.722367: 0 (5.06)
| | | | B7 > 0.424875: 0 (10.12)
| | | B3 > 0.086317: 1 (377.11/32.59)
| | F3 > 0.258804
| | | B2 <= 0.347251: 0 (9.49)
| | | B2 > 0.347251
| | | | F3 <= 0.378424: 1 (26.89/4.11)
| | | | F3 > 0.378424
| | | | B4 <= 0.142638: 1 (2.85)
| | | | B4 > 0.142638: 0 (9.17)
```

F19 SEQUEST

```
F3 <= 0.05697: 0 (37.6)
F3 > 0.05697
| F7 <= 0.02754
| | B4 <= 0.654162: 0 (14.29)
| | B4 > 0.654162: 1 (17.45/0.6)
| F7 > 0.02754
| | B7 <= 0.114544: 1 (476.52/13.24)
| | B7 > 0.114544
| | | B2 <= 0.170216: 0 (9.78)
| | | B2 > 0.170216
| | | | F8 <= 0.17425: 1 (35.65/1.96)
| | | | F8 > 0.17425: 0 (2.71)
```

F19 ProICAT

```
F3 <= 0.058962: 0 (34.26)
F3 > 0.058962
|   F7 <= 0.025279
|   |   F3 <= 0.092571: 0 (6.05)
|   |   F3 > 0.092571: 1 (40.58/5.11)
|   F7 > 0.025279
|   |   B7 <= 0.339105: 1 (481.94/14.91)
|   |   B7 > 0.339105
|   |   |   F7 <= 0.097015: 0 (4.17)
|   |   |   F7 > 0.097015: 1 (27.01/3.36)
```

Standard SEQUEST

```
F8 <= 0.006151
|   F8 <= 0.003251: 0 (20.67)
|   F8 > 0.003251
|   |   F2 <= 0.356915: 0 (9.03)
|   |   F2 > 0.356915
|   |   |   F2 <= 0.424882: 1 (30.75/1.96)
|   |   |   F2 > 0.424882: 0 (3.14)
F8 > 0.006151
|   F3 <= 0.043433
|   |   F1 <= 0.053213
|   |   |   F5 <= 0.000571: 1 (74.06/4.97)
|   |   |   F5 > 0.000571: 0 (18.32)
|   |   |   F1 > 0.053213: 0 (28.13)
|   F3 > 0.043433
|   |   F4 <= 0.51631
|   |   |   F2 <= 0.307931
|   |   |   |   B7 <= 0.102249
|   |   |   |   |   F3 <= 0.141811
|   |   |   |   |   |   F5 <= 0.000867: 0 (2.62)
|   |   |   |   |   |   F5 > 0.000867: 1 (26.43/3.4)
|   |   |   |   |   |   F3 > 0.141811: 0 (6.67)
|   |   |   |   |   |   B7 > 0.102249
|   |   |   |   |   |   F3 <= 0.065411
|   |   |   |   |   |   |   F7 <= 0.000935: 1 (13.22/1.7)
|   |   |   |   |   |   |   F7 > 0.000935: 0 (8.5)
|   |   |   |   |   |   |   F3 > 0.065411
|   |   |   |   |   |   |   |   B2 <= 0.240827: 1 (307.22/19.37)
|   |   |   |   |   |   |   |   B2 > 0.240827
|   |   |   |   |   |   |   |   |   F2 <= 0.251136: 1 (12.43/0.92)
|   |   |   |   |   |   |   |   |   F2 > 0.251136: 0 (4.32)
|   |   |   |   |   |   |   |   |   F2 > 0.307931: 1 (1347.18/69.09)
|   |   F4 > 0.51631
|   |   |   B5 <= 0.304388: 0 (7.72)
|   |   |   B5 > 0.304388
|   |   |   |   B7 <= 0.282436: 1 (18.58/1.31)
|   |   |   |   B7 > 0.282436: 0 (3.01)
```

Standard ProICAT

```
F3 <= 0.043121
|   F4 <= 0.216039: 0 (97.73)
|   F4 > 0.216039
|   |   F4 <= 0.224235: 1 (9.26/0.21)
|   |   F4 > 0.224235: 0 (18.79)
F3 > 0.043121
|   F2 <= 0.307643
|   |   F3 <= 0.064168
|   |   |   F8ladder <= 0.36985: 0 (29.03)
|   |   |   F8ladder > 0.36985: 1 (4.73/0.21)
|   |   F3 > 0.064168
|   |   |   B4 <= 0.383065
|   |   |   |   B5 <= 0.536563
|   |   |   |   |   F2 <= 0.195826: 1 (5.57/1.04)
|   |   |   |   |   F2 > 0.195826: 0 (22.76)
|   |   |   |   |   B5 > 0.536563: 1 (4.52)
|   |   |   B4 > 0.383065
|   |   |   |   B3 <= 0.101356
|   |   |   |   |   F4 <= 0.230005: 0 (6.26)
|   |   |   |   |   F4 > 0.230005
|   |   |   |   |   |   B5 <= 0.235832
|   |   |   |   |   |   |   F7 <= 0.007278: 0 (9.61)
|   |   |   |   |   |   |   F7 > 0.007278: 1 (5.57/1.04)
|   |   |   |   |   |   B5 > 0.235832
|   |   |   |   |   |   |   B5 <= 0.375229: 1 (50.89/5.64)
|   |   |   |   |   |   |   B5 > 0.375229
|   |   |   |   |   |   |   |   B5 <= 0.556216: 0 (6.68)
|   |   |   |   |   |   |   |   B5 > 0.556216: 1 (10.93/1.88)
|   |   |   |   B3 > 0.101356: 1 (221.01/21.93)
|   F2 > 0.307643
|   |   F4 <= 0.27345
|   |   |   B4 <= 0.386424
|   |   |   |   F1 <= 0.072791
|   |   |   |   |   F7 <= 0.003758
|   |   |   |   |   |   F5 <= 0.000136: 1 (4.52)
|   |   |   |   |   |   F5 > 0.000136: 0 (13.57)
|   |   |   |   |   |   F7 > 0.003758: 1 (19.14/1.04)
|   |   |   |   F1 > 0.072791
|   |   |   |   |   F5 <= 0.001934: 1 (119.59/6.47)
|   |   |   |   |   F5 > 0.001934
|   |   |   |   |   |   B7 <= 0.156101: 0 (3.55)
|   |   |   |   |   |   B7 > 0.156101: 1 (4.94/0.42)
|   |   |   B4 > 0.386424
|   |   |   |   F3 <= 0.056468
|   |   |   |   |   F7 <= 0.004873
|   |   |   |   |   |   F4 <= 0.091233: 0 (2.09)
|   |   |   |   |   |   F4 > 0.091233: 1 (38.49/2.3)
|   |   |   |   |   |   F7 > 0.004873: 0 (4.18)
|   |   |   |   F3 > 0.056468: 1 (1059.2/41.14)
|   |   F4 > 0.27345
|   |   |   F2 <= 0.422156
|   |   |   |   F4 <= 0.4612: 1 (143.19/16.5)
|   |   |   |   F4 > 0.4612
|   |   |   |   |   B2 <= 0.007878: 1 (5.15/0.63)
|   |   |   |   |   B2 > 0.007878: 0 (6.06)
|   |   |   F2 > 0.422156
|   |   |   |   F6 <= 0.313815: 0 (10.44)
|   |   |   |   F6 > 0.313815: 1 (4.52)
```