Finding putative PTM (pPTM) Marker Ion in HCD scans

Christian Panse Paolo Nanni

April 18, 2013

Contents

1 Get Data In – Preprocessing

The minimal data structure requirement for the PTM_MarkerFinder function looks as follow.

```
> library(protViz)
> data(HexNAc)
> str(HexNAc[[1]], nchar.max = 30)

List of 12
 $ peptideSequence   : chr "STMQELNSR"
 $ mascotScore       : num 49.5
 $ modification      : chr "00000000000"
 $ MonoisotopicAAmass: num [1:9] 0 0 0 0 0 0 0 0 0
 $ proteinInformation: chr "zz|ZZ_FGCZCont0219|"
 $ title             : chr "NGlycoFASP_NHexNAc_HCDETD.265"| __truncated__
 $ pepmass           : num 533
 $ charge            : num 2
 $ scans             : num 2659
 $ rtinseconds       : num 1846
 $ mZ                : num [1:150] 101 104 105 110 112 ...
 $ intensity         : num [1:150] 369.3 2860 37.3 103.8 190.7 ...
```

Here we have listed the HexNAc data which is included in protViz.

protViz also provides and perl script mascotDat2RData.pl taking mascot server dat files as input and producing RData output.

```
$ /usr/local/lib/R/site-library/protViz/exec/mascotDat2RData.pl \
-d=/usr/local/mascot/data/20130116/F178287.dat \
-m=$HOME/mod_file
```

mascotDat2RData.pl requires the mascot server mod_file keeping all the configured modification of the mascot server.

In theory PTM_MarkerFinder can process the output of any search engine for peptide identification. It is up to the R user writing a wrapper script converting the output of any particular peptide identification search engine to the data structure listed above.

2 Finding the Marker Ions

PTM_MarkerFinder can search for any Marker ion series. The next lines define the HexNAc_MarkerIons.

```
> HexNAc_MarkerIons <- c(126.05495, 138.05495, 144.06552, 168.06552,
+     186.07608, 204.08665)
```

The lines below configure the modification information used by the search engine. The HexNAc modification below is described on unimod http://www.unimod.org/modifications_view.php?editid1=43.

```
> ptm.0<-cbind(AA="-",
+         mono=0.0, avg=0.0, desc="unmodified", unimodAccID=NA)
> ptm.1<-cbind(AA='N',
+         mono=317.122300, avg=NA, desc="HexNAc",
+         unimodAccID=2)
> ptm.2<-cbind(AA='M',
+         mono=147.035400, avg=NA, desc="Oxidation",
+         unimodAccID=1)
> m<-as.data.frame(rbind(ptm.0, ptm.1, ptm.2))
```

PTM_MarkerFinder is called.

```
> s <- PTM_MarkerFinder(data=HexNAc, modification=m$mono,
+         modificationName=m$desc,
+         minMarkerIntensityRatio=3,
+         itol_ppm=20,
+         mZmarkerIons=HexNAc_MarkerIons)
> s
```

	scans	mZ	markerIonMZ	markerIonIntensity	markerIonMzError
1	3687	126.0550	126.0550	9945.0	-0.000081
2	3687	138.0553	138.0549	1933.0	-0.000344
3	3687	144.0658	144.0655	412.3	-0.000230
4	3687	168.0659	168.0655	810.2	-0.000398
5	3687	204.0870	204.0866	3273.0	-0.000356
6	2540	126.0551	126.0550	2945.0	-0.000104
7	2540	138.0564	138.0549	759.2	-0.001432
8	2540	144.0655	144.0655	195.4	-0.000017
9	2540	168.0657	168.0655	262.9	-0.000154
10	2540	186.0766	186.0761	188.5	-0.000550
11	2540	204.0870	204.0866	998.4	-0.000310
12	4393	126.0551	126.0550	13620.0	-0.000131
13	4393	138.0550	138.0549	3798.0	-0.000058
14	4393	168.0656	168.0655	1526.0	-0.000108
15	4393	186.0763	186.0761	1014.0	-0.000183
16	4393	204.0869	204.0866	5041.0	-0.000218
17	2739	126.0550	126.0550	7327.0	-0.000087
18	2739	138.0550	138.0549	1963.0	-0.000043
19	2739	144.0656	144.0655	468.6	-0.000077
20	2739	168.0656	168.0655	624.3	-0.000108
21	2739	204.0868	204.0866	2496.0	-0.000127

	markerIonPpmError	query	pepmass
1	-0.6425765	4	713.3583
2	-2.4917552	4	713.3583
3	-1.5964933	4	713.3583
4	-2.3681184	4	713.3583
5	-1.7443541	4	713.3583
6	-0.8250363	6	490.5612

```
7        -10.3725737     6 490.5612
8         -0.1180019     6 490.5612
9         -0.9163085     6 490.5612
10        -2.9557715     6 490.5612
11        -1.5189603     6 490.5612
12        -1.0392282     9 891.4088
13        -0.4201224     9 891.4088
14        -0.6426061     9 891.4088
15        -0.9834677     9 891.4088
16        -1.0681726     9 891.4088
17        -0.6901747    10 665.5916
18        -0.3114701    10 665.5916
19        -0.5344787    10 665.5916
20        -0.6426061    10 665.5916
21        -0.6222843    10 665.5916
```

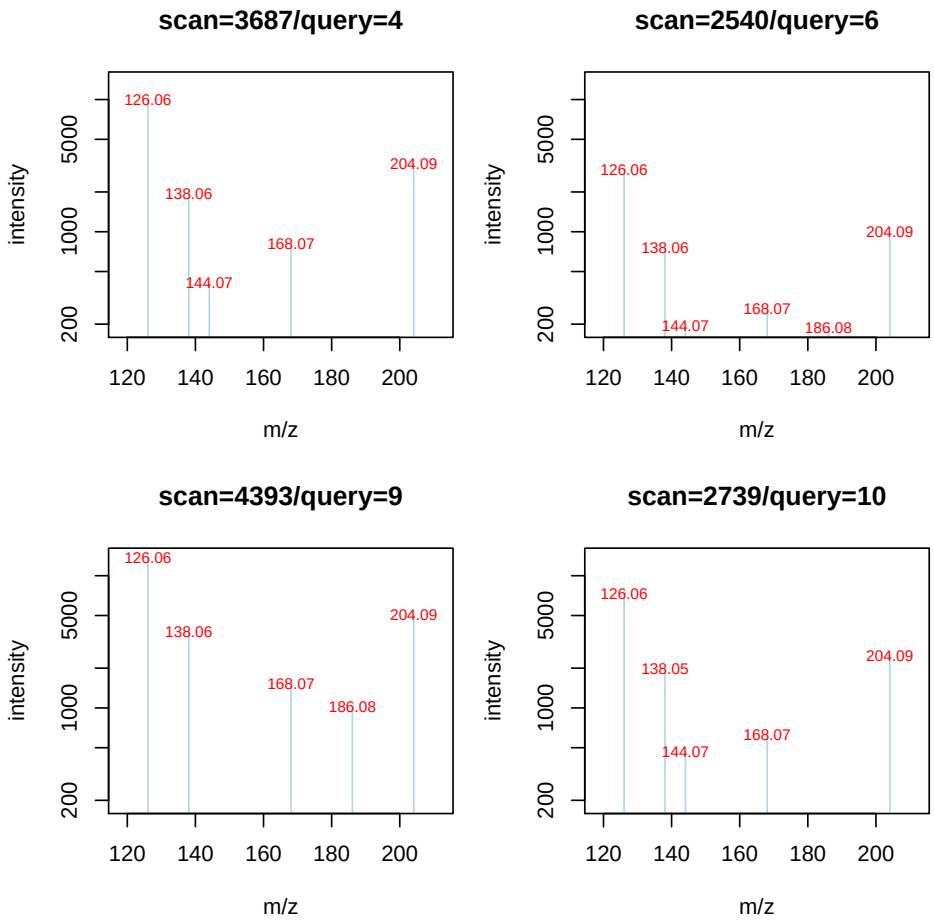
The user can call the demonstration with

```
> demo(PTM_MarkerFinder)
```

3 Some overview graphics

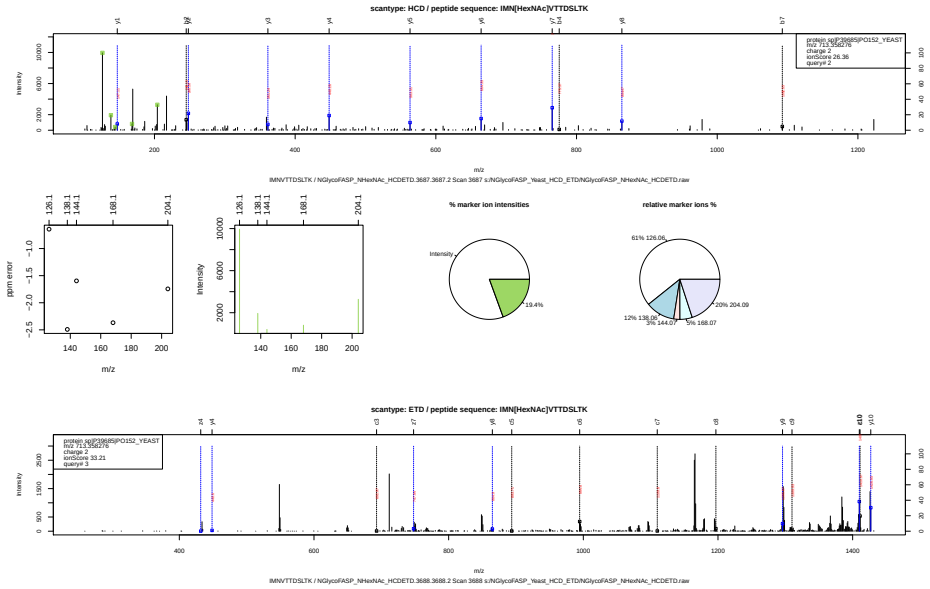
just an overview of the sample data set HexNAc.

```
> op<-par(mfrow=c(2,2), mar=c(4,4,4,1));
> dump <- lapply(split(s,s$query),
+     function(x){ plot(x$mZ, x$markerIonIntensity,
+         type='h',
+         col='lightblue',
+         cex=2,
+         ylab='intensity', xlab='m/z',
+         xlim=range(c(HexNAc_MarkerIons,
+             max(HexNAc_MarkerIons)
+                 + 0.1 * (max(HexNAc_MarkerIons) - min(HexNAc_MarkerIons)),
+             min(HexNAc_MarkerIons)
+                 - 0.1 * (max(HexNAc_MarkerIons) - min(HexNAc_MarkerIons)))),
+             ylim=range(s$markerIonIntensity),
+             log='y',
+             main=paste("scan=", unique(x$scans),
+                 "/query=", unique(x$query), sep='));
+             text(x$mZ, x$markerIonIntensity,
+                 round(x$mZ,2),col='red',cex=0.7)
+         }
+     )
> par(op)
```



The next graphics show the output of `PTM_MarkerFinder`.

```
> d<-list(); d[[1]]<-HexNAc[[3]]; d[[2]]<-HexNAc[[4]];
> d[[3]]<-HexNAc[[5]]
> ss<-PTM_MarkerFinder(data=d, modification=m$mono,
+         modificationName=m$desc,
+         minMarkerIntensityRatio=3,
+         itol_ppm=20,
+         mZmarkerIons=HexNAc_MarkerIons)
>
```



4 Reshaping the output and export

reshape the table:

```
> w<-reshape(s[,c(1,7,3,4)], direction='wide',
+     timevar="markerIonMZ", idvar=c('scans','query'))
> w
```

```
   scans query markerIonIntensity.126.05495 markerIonIntensity.138.05495
1   3687     4                         9945                        1933.0
6   2540     6                         2945                         759.2
12  4393     9                        13620                        3798.0
17  2739    10                         7327                        1963.0
   markerIonIntensity.144.06552 markerIonIntensity.168.06552
1                         412.3                        810.2
6                         195.4                        262.9
12                           NA                       1526.0
17                        468.6                        624.3
   markerIonIntensity.204.08665 markerIonIntensity.186.07608
1                        3273.0                           NA
6                         998.4                        188.5
12                       5041.0                       1014.0
17                       2496.0                           NA
```

export as comma separeted file

```
> write.table(w, file="HexNAc_PTM_markerFinder.csv",
+     sep=',', row.names=FALSE,col.names=TRUE, quote=FALSE)
>
```

References

[1] Nanni, P., Panse, C., Gehrig, P., Mueller, S., Grossmann, J., Schlapbach, R., PTM MarkerFinder, a software tool to detect and validate spectra from peptides carrying post-translational modifications. submitted to PROTEOMICS, 2013 (pmic.201300036.R1).