Computational Biology

Team 22

12 Raw Files

Under supervision :

Dr / Ibrahim el semman

Table of contents

Team Members

Ashrqat Maher

Rewaa Ashraf

Fatma Osama

Yasmin Ahmed

Yara Essam

Yara Aony

Load mzMl File and obtain spectrum for peptide

```
from pyopenms import *
exp = MSExperiment()
MzMLFile().load("000_VTA_PRM1_C19.mzML", exp)
spectra = exp.getSpectra()
observed_spectrum = spectra[4]
```

```
observed_spectrum = spectra[18]
observed_spectrum
```

```
<pyopenms.pyopenms_5.MSSpectrum at 0x18be08f4930>
```

Digestion and store all fragment peptide

```
from pyopenms import *
dig = ProteaseDigestion()
dig.getEnzymeName() # Trypsin

entries = []
f = FASTAFile()
f.load("Theo.fasta", entries)
bsa2=[]
for e in entries:
    bsa2.append(AASequence.fromString( e.sequence))
resultAll = []
resultOneSeq=[]
result=[]
for u in bsa2:
    resultOneSeq.append( dig.digest(u, result))
    for re in result:
        print(re)
        resultAll.append(re)
```

```
MAALDSLSLFTSLGLSEQK
AR
ETLK
NSALSAQLR
EAATQAQQTLGSTIDK
ATGILLYGLASR
LR
DTR
R
LSFLVSYIASK
K
IHTEPQLSAALEYVR
SHPLDPIDTVDFER
ECGVGVIVTPEQIEEAVEAAINR
HRPQLLVER
YHFNMGLLMGEAR
AVLK
WADGK
MIK
NEVDMQVLHLLGPK
LEADLEK
K
FK
VAK
AR
LEETDR
```

Generate theoretical spectrum

```
tsg = TheoreticalSpectrumGenerator()
theo_spectrum = MSSpectrum()
p = tsg.getParameters()
p.setValue("add_y_ions", "true")
p.setValue("add_b_ions", "true")
p.setValue("add_metainfo", "true")
tsg.setParameters(p)
```

```
tsg
```

```
<pyopenms.pyopenms_8.TheoreticalSpectrumGenerator at 0x2544ed4dbd0>
```

```
entries =[]
f = FASTAFile()
f.load(r"resultAll.fasta", entries)
fasta_file=entries[0]
peptide = AASequence.fromString(fasta_file.sequence)
tsg.getSpectrum(theo_spectrum, peptide, 1, 2)
for ion, peak in zip(theo_spectrum.getStringDataArrays()[0], theo_spectrum):
    print(ion, peak.getMZ())
```

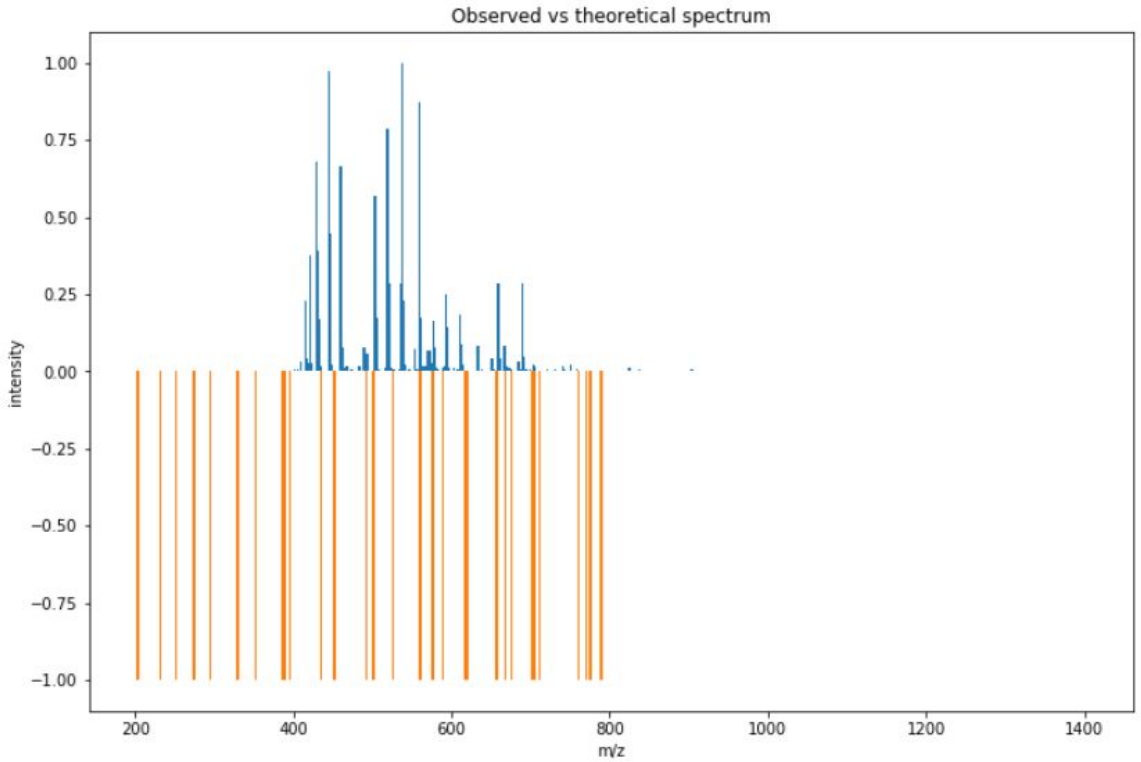
```
b'y1++' 59.546766142221
b'y2++' 88.05749819007102
b'b2++' 102.04607605507098
b'y3++' 116.568230237921
b'y1+' 118.086255817671
b'b3++' 137.56463313482098
b'y2+' 175.107719913371
b'b4++' 194.10666531027098
b'y4++' 194.61878592932104
b'b2+' 203.08487564337096
b'y3+' 232.12918400907097
b'y5++' 251.16081810477104
b'b5++' 251.62013739002097
b'b3+' 274.121989802871
b'b6++' 295.136151969771
b'y6++' 329.211373796171
b'b7++' 351.678184145221
```

Plot -Observed vs theoretical spectrum-

```
import numpy as np
from matplotlib import pyplot as plt

def mirror_plot(obs_mz, obs_int, theo_mz, theo_int, title):
    obs_int = [element / max(obs_int) for element in obs_int]
    theo_int = [element * -1 for element in theo_int]
    plt.figure(figsize=(12,8))
    plt.bar(obs_mz, obs_int, width = 3.0)
    plt.bar(theo_mz, theo_int, width = 3.0)
    plt.title(title)
    plt.ylabel('intensity')
    plt.xlabel('m/z')

obs_mz, obs_int = observed_spectrum.get_peaks()

print(min(obs_mz))
print(max(obs_mz))

theo_mz, theo_int = [], []
for mz, intensity in zip(*theo_spectrum.get_peaks()):
    if mz >= 200.0 and mz <= 800.0:
        theo_mz.append(mz)
        theo_int.append(intensity)

title = 'Observed vs theoretical spectrum'
mirror_plot(obs_mz, obs_int, theo_mz, theo_int, title)
```

Plot -Observed vs theoretical spectrum-

```
400.0001390907547
1401.399996863582
```

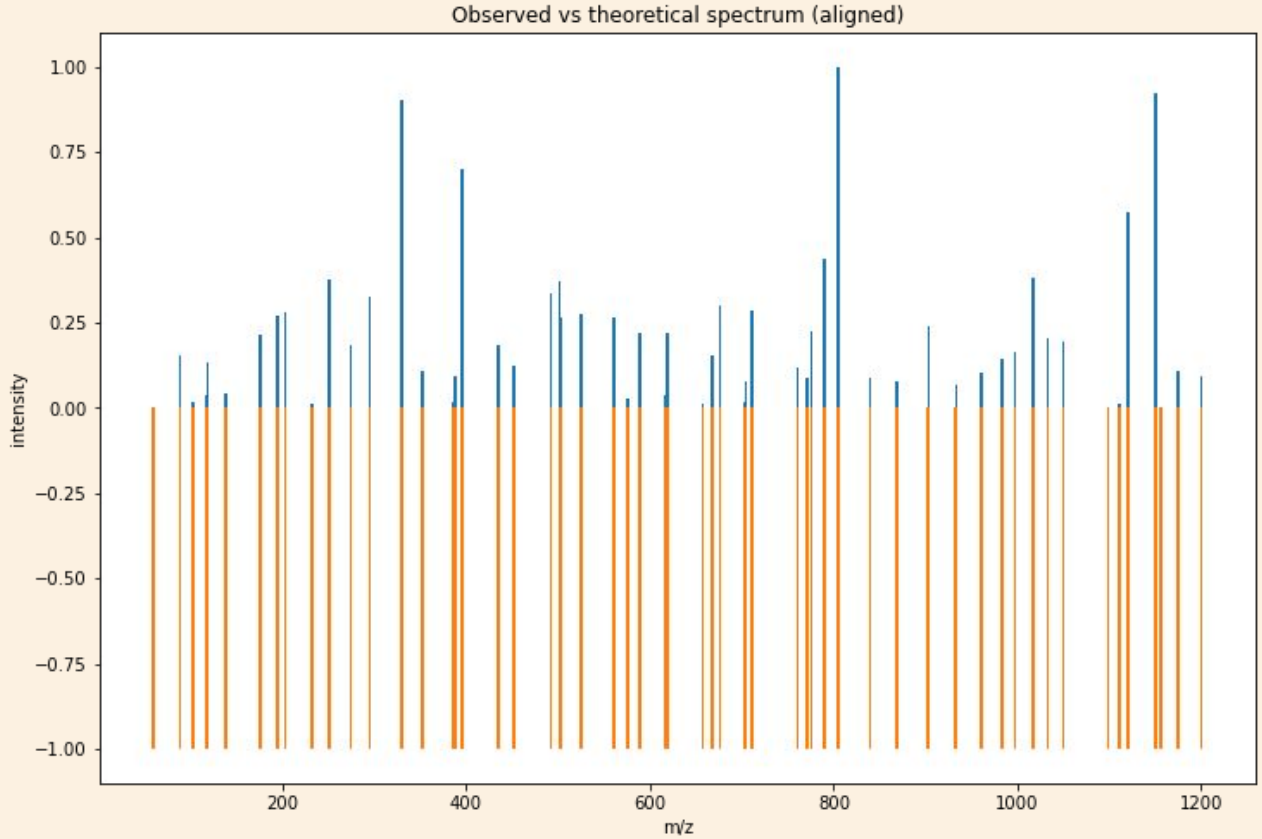


Matching ions and mz from theoretical spectrum

```
print("Number of matched peaks: " + str(len(alignment)))
print("ion\ttheo. m/z\tobserved m/z")
for theo_idx, obs_idx in alignment:
    ion_name = theo_spectrum.getStringDataArrays()[0][theo_idx].decode()
    ion_charge = theo_spectrum.getIntegerDataArrays()[0][theo_idx]
    print(ion_name + "\t" + str(ion_charge) + "\t"
      + str(theo_spectrum[theo_idx].getMZ())
      + "\t" + str(observed_spectrum[obs_idx].getMZ()))
```

```
Number of matched peaks: 62
ion     theo. m/z       observed m/z
y1++    2       59.546766142221 59.548274993896484
y2++    2       88.05749819007102       88.05718994140625
b2++    2       102.04607605507098      102.04742431640625
y3++    2       116.568230237921        116.56841278076172
y1+     1       118.086255817671        118.08643341064453
b3++    2       137.56463313482098      137.5648193359375
y2+     1       175.107719913371        175.10821533203125
b4++    2       194.10666531027098      194.10780334472656
y4++    2       194.61878592932104      194.61907958984375
b2+     1       203.08487564337096      203.08377075195312
y3+     1       232.12918400907097      232.1260223388672
y5++    2       251.16081810477104      251.15884399414062
b5++    2       251.62013739002097      251.6226348876953
b3+     1       274.121989802871        274.1199035644531
b6++    2       295.136151969771        295.1376647949219
y6++    2       329.211373796171        329.2120361328125
```

Mirror -Aligned Spectrum-



Oxytocin response and social behavior

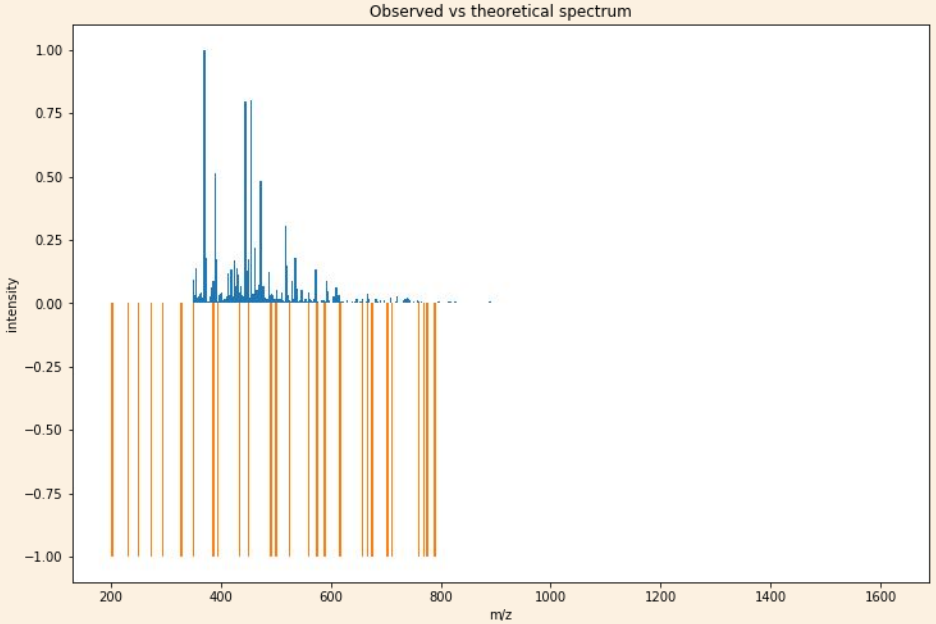
000_VTA_PRM1_C19

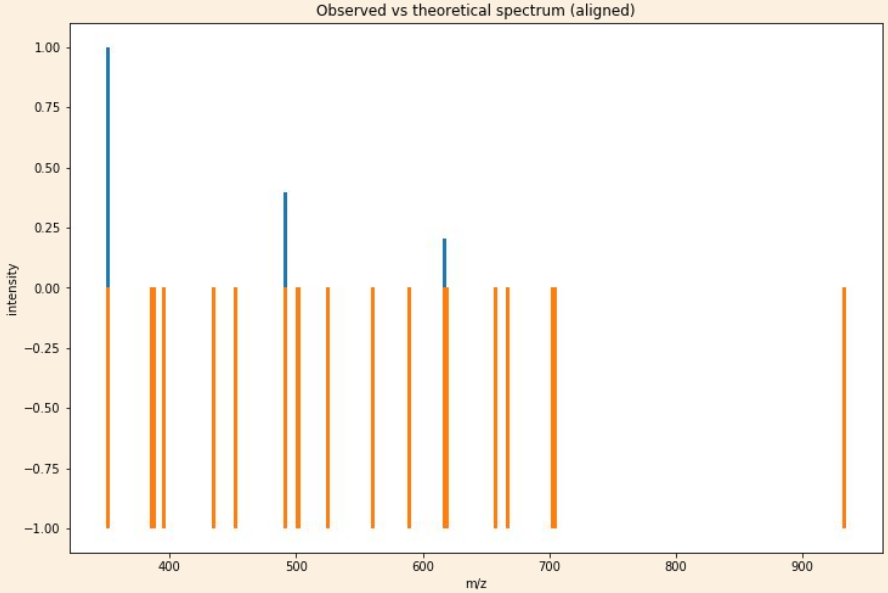
Number of matched peaks: 20

ion		theo. m/z	observed m/z
b7++	2	351.678184145221	351.678372684875
y7++	2	385.753405971621	385.79908804540554
b4+	1	387.206054153771	387.2055277787431
y4+	1	388.23029539187104	388.2309235306787
b8++	2	395.194198724971	395.184898724151
y8++	2	435.28761311517104	435.34520202133683
b9++	2	451.736230900421	451.7161914887091
y9++	2	491.82964529062104	491.8295126352524
y5+	1	501.31435974277105	501.28898378810305
b5+	1	502.23299831327097	501.94869051486586
b10++	2	525.270438043971	525.2799598068392
y10++	2	560.359101402271	560.8410040240839
b6+	1	589.265027472771	588.8723983374019
y11++	2	616.901133577721	616.9025175637059
b12++	2	619.3102922353711	619.7410662560254
y6+	1	657.415471125571	657.7652897288622
y12++	2	667.424973189371	667.2208366959128
b7+	1	702.349091823671	702.7616866498944
b14++	2	704.363056458671	704.821393840089
b18++	2	932.9716894531211	932.7006071662397

Oxytocin response and social behavior

000_VTA_PRM1_C19





Oxytocin response and social behavior

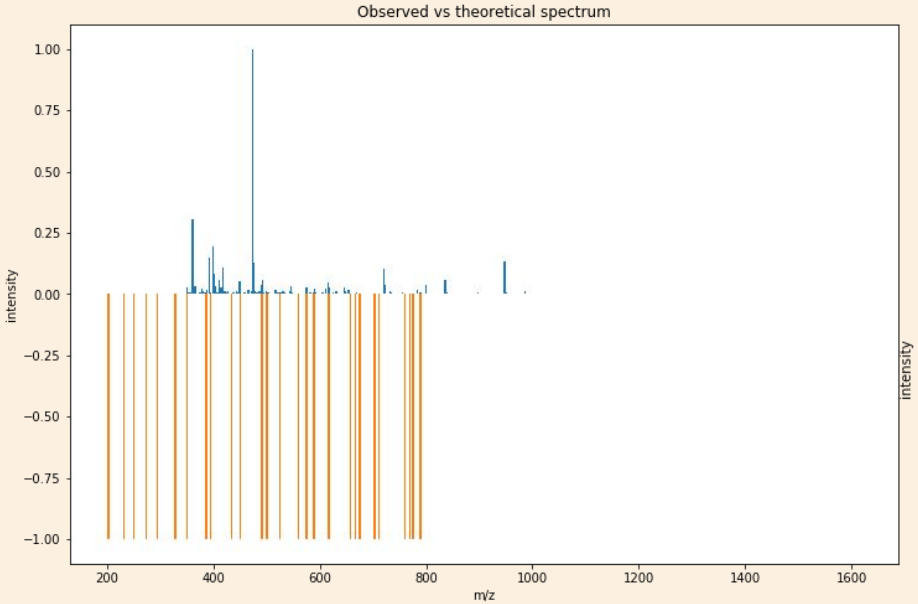
022_VTA_PRM_190122102547_C19

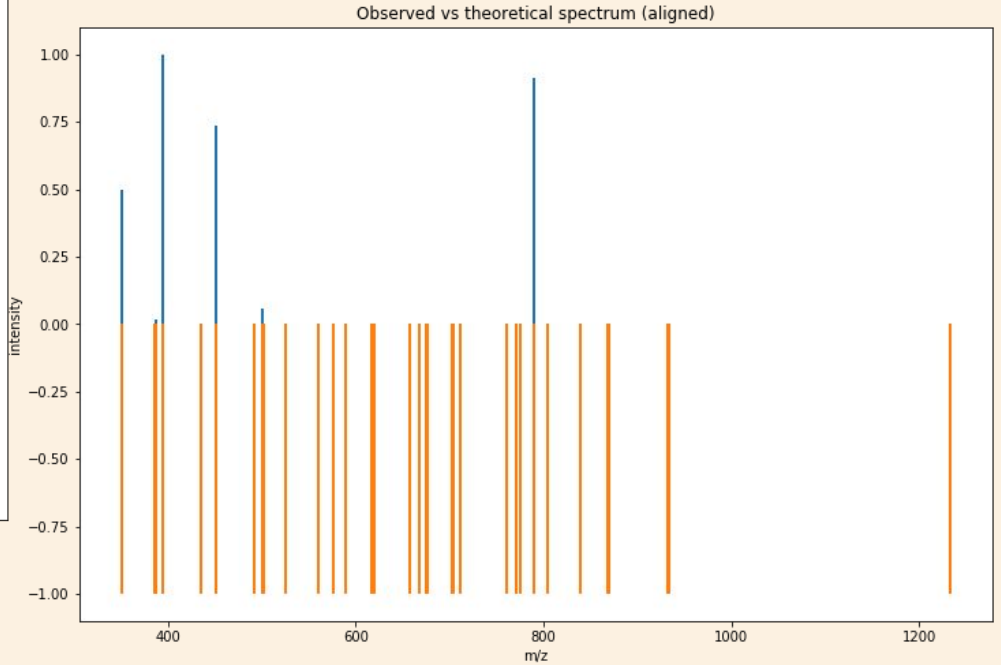
Number of matched peaks: 32

ion	theo. m/z	observed m/z	
b7++	2	351.678184145221	351.6776885629277
y7++	2	385.753405971621	385.72269428314235
b4+	1	387.206054153771	387.2051978306316
y4+	1	388.23029539187104	388.2297674126516
b8++	2	395.194198724971	395.19426508306617
y8++	2	435.28761311517104	435.2610229694314
b9++	2	451.736230900421	451.7360746340657
y9++	2	491.82964529062104	491.75436548863973
y5+	1	501.31435974277105	501.3138243479935
b5+	1	502.23299831327097	502.23415469047086
b10++	2	525.270438043971	525.2557773526281
y10++	2	560.359101402271	560.3583351040395
b11++	2	575.794277655621	576.2498117086093
b6+	1	589.265027472771	589.2735842545871
y11++	2	616.901133577721	616.8387600657122
b12++	2	619.3102922353711	619.292952541795
y6+	1	657.415471125571	657.3650509607287
y12++	2	667.424973189371	667.3085196583243
b13++	2	675.8523244108211	675.8315372317325
b7+	1	702.349091823671	702.3486943491313
b14++	2	704.363056458671	704.3725095302223
y13++	2	710.9409877691211	711.261726283785
b15++	2	760.9050886341211	760.4495378481092
y7+	1	770.499535476471	770.4134237963619
y14++	2	775.4622848807711	775.3936641239696
b8+	1	789.381120983171	789.3811410768251
b16++	2	804.4211032138711	804.4159358407297
y15++	2	839.5097665721711	839.4979331959191
b17++	2	868.9424003255211	869.3019847053566
y8+	1	869.567949763571	869.3983058978838
b18++	2	932.9716894531211	933.3940539899982
y11+	1	1232.794990688671	1232.678470394013

Oxytocin response and social behavior

022_VTA_PRM_190122102547_C19





Thanks!