Supporting Information

Distinct Dibasic Cleavage Specificities of Neuropeptide-Producing Cathepsin L and Cathepsin V Cysteine Proteases Compared to PC1/3 and PC2 Serine Proteases

Michael Yoon[1], Janneca Ames[1], Charles Mosier[1], Zhenze Jiang[1], Sonia Podvin[1]
Anthony J. O'Donoghue[1], and Vivian Hook[1, 2]

[1]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093, United States:
[2]Department of Neurosciences and Department of Pharmacology, University of California, San Diego, La Jolla, CA 92093, United States

Supplemental Information

Supplementary Figures:

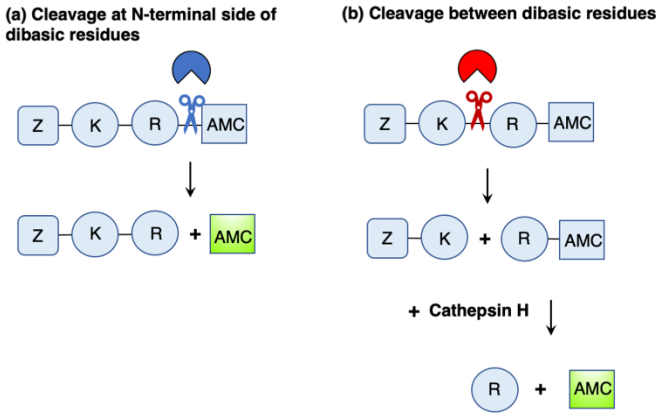

Figure S1. Coupled protease assay with cathepsin H aminopeptidase activity to monitor peptide-AMC cleavages with N-terminal residue-extended AMC products.
a. C-terminal cleavage at dibasic residues of Z-K-R-AMC, as an example of the Z-K/R-K/R-↓AMC and Z-X-K/R-K/R-↓AMC substrates, generates fluorescent AMC monitored at excitation/emission of 360/460 nm, shown here for Z-K-R-AMC as example.
b. Cleavage between or at the N-terminal side of dibasic residues of Z-K-R-AMC, as an example for the Z-K/R-K/R-↓AMC and Z-X-K/R-K/R-↓AMC substrates, generates basic residue extended AMC products that are monitored by addition of the cathepsin H aminopeptidase to generate fluorescent AMC for detection and quantitation of cleaved products.

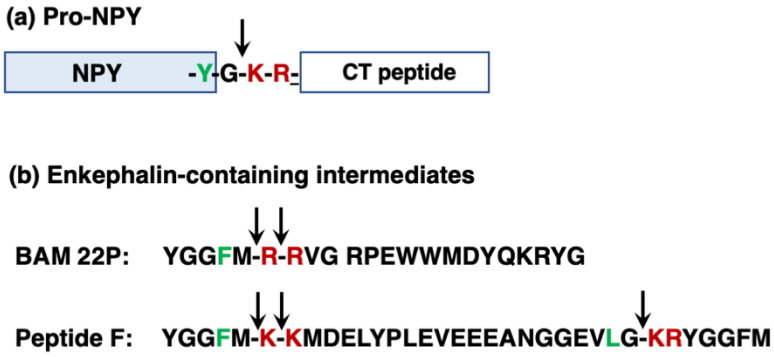

Figure S2. Cathepsin L cleavage of several pro-neuropeptides at dibasic residues, between and at the N-terminal side, reported in the literature.
a. Cathepsin L cleavage of pro-NPY. Cathepsin L cleavage of recombinant pro-NPY was determined by mass spectrometry analyses of cleavage products, as reported by our prior study of Funkelstein et al., 2008 (21). Results indicated cathepsin L cleavage at the N-terminal side of the KR (pink) processing site, with hydrophobic Tyr residue (green).
(b) Cathepsin L cleavage of enkephalin-containing intermediates BAM22P and Peptide F. Cathepsin L cleavage of BAM22P and Peptide F was determined by mass spectrometry analyses of cleavage products, as reported in our prior study of Yasothornsrikul et al., 2003 (20). Results demonstrated cathepsin L cleavage between the R↓R dibasic site, and at the N-terminal side of ↓RR with hydrophobic Tyr as P2 residue.

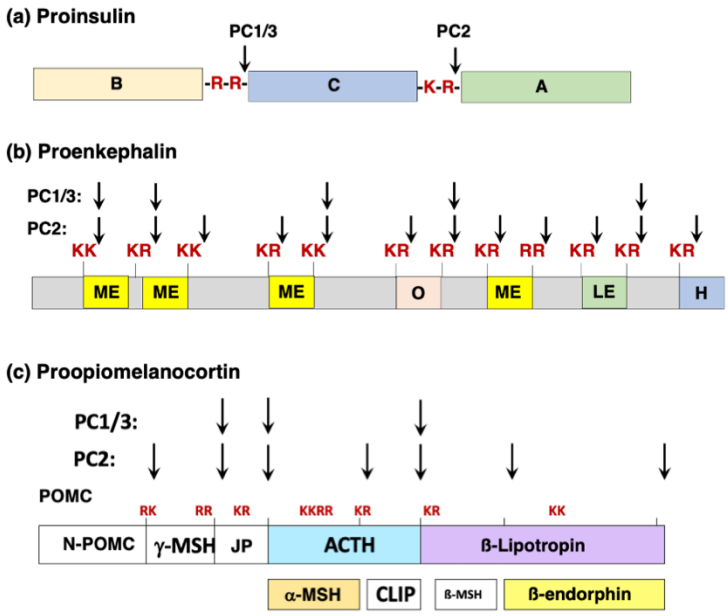

Figure S3. PC1/3 and PC2 cleavage of proinsulin and proenkephalin at the N-terminal side of dibasic residues, reported in the literature.
(a) Proinsulin processing by PC1/3 and PC2. PC1/3 (human) has been demonstrated to cleave human proinsulin at the C-terminal side of RR↓ between the B and C chains, as reported by Davidson et al., 1988 (22). PC2 (human) has been shown to cleave at the COOH-terminal side of KR↓ between the C and A chains of proinsulin (22).
(b) Proenkephalin processing by PC1/3 and PC2. Processing of proenkephalin (rat) by PC1/3 and PC2 (mouse) was conducted by analyses of peptide products by MALDI-TOF, as reported by Peinado et al., 2003 (26). Arrows indicate the cleavage sites within proenkephalin by PC1/3 and PC2.

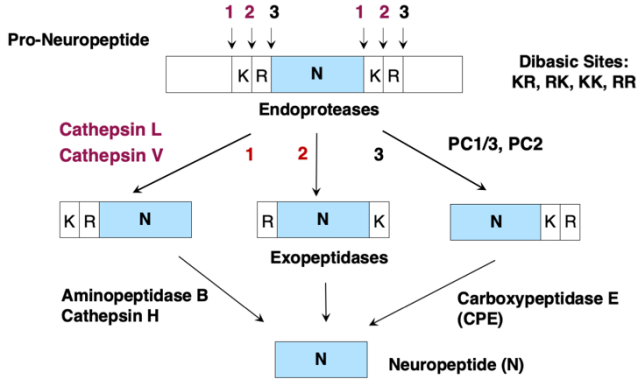

Figure S4. Hypothesis for differential dibasic cleavage specificities of the cathepsin L and cathepsin V cysteine proteases, compared to the PC1/3 and PC2 serine proteases, for neuropeptide biosynthesis. Neuropeptides comprise peptide neurotransmitters and hormones. Neuropeptides are first synthesized as pro-neuropeptide precursors with dibasic residues (KR, RK, KK, and RR) flanking the active neuropeptides. The dibasic residue regions have been found to undergo proteolytic processing by cathepsin L and cathepsin V cysteine proteases (1, 2), combined with the pro-protein convertases 1 and 2 (PC1/3 and PC2) (1-3). It is hypothesized that different cleavage specificities exist for the cathepsin L and cathepsin V proteases which cleave between and at the N-terminal side of the dibasic residues, whereas the PC1/3 and PC2 serine proteases cleavage at the C-terminal side of dibasic residues. This hypothesis is supported by the findings of this study.

Supplemental Table:

	14-mers with dibasic sequence (n=19)
1	nLDKLnNWPQRRGn
2	GnYYKRFnAHWVGI
3	QnKKTLVnYNEWNL
4	LGWHAnFRKYPInA
5	GSQVFSWLNHYHRK
6	HTNKRISQWnWEIR
7	RKWQSPQVDLYDKS
8	HRRVYLTSPKAPES
9	VDYIEHKDQVRRnN
10	nEFHWRInQGKKAP
11	TPHHVNWYKRAPNQ
12	EGADIWYRKHSHQL
13	LRKDWGDIQFATAN
14	IEPPWVDSHAKRNn
15	YQLLTnNEIFRKWH
16	YWnSTHLAGKRRDW
17	DAWAPnVIKKESSI
18	ADARKYWNVHGTHQ
19	ANnQILDPDNFKRE

Table S1. Dibasic residue containing peptides in the library of 228 14-mer peptides used for MSP-MS analyses.
The sequences of 14-mer peptides of the library that contain dibasic residues are shown. The library also contained one tribasic peptide.

Supplemental Methods

LC-MS-MS Report:
MSP-MS of Cathepsin L, Cathepsin V, PC1/3, and PC2 at pH 5.5 and pH 7.4

Samples were resuspended in 0.1% TFA to a total peptide concentration of 28.5 µM
Each sample were injected once, 4 ul total volume per injection

The C18 column consisted of 1.7 µm bead size, 75 µm x 20 cm, heated to 65C

Solvent A – water, 0.1% formic acid
Solvent B – acetonitrile, 0.1% formic acid

Nano-LC gradient:

A	B	C	D
Time (min)	Flow (µl/min)	% solvent A	% solvent B
0	0.3	99	1
0.1	0.3	95	5
60	0.3	70	30
65	0.3	15	85
75	0.3	15	85
75.1	0.3	10	90
85	0.3	10	90
85.1	0.3	0	100

MS Sample List

MY_20180821_CatV74_60_4.raw	MY_20180821_PC1_60_4.raw
MY_20180821_CatV74_60_3.raw	MY_20180821_PC1_60_3.raw
MY_20180821_CatV74_60_2.raw	MY_20180821_PC1_60_2.raw
MY_20180821_CatV74_60_1.raw	MY_20180821_PC1_60_1.raw
MY_20180821_CatV74_30_4.raw	MY_20180821_PC1_30_4.raw
MY_20180821_CatV74_30_3.raw	MY_20180821_PC1_30_3.raw
MY_20180821_CatV74_30_2.raw	MY_20180821_PC1_30_2.raw
MY_20180821_CatV74_30_1.raw	MY_20180821_PC1_30_1.raw
MY_20180821_CatV55_60_4.raw	MY_20180821_PC1_0_4.raw
MY_20180821_CatV55_60_3.raw	MY_20180821_PC1_0_3.raw
MY_20180821_CatV55_60_2.raw	MY_20180821_PC1_0_2.raw
MY_20180821_CatV55_60_1.raw	MY_20180821_PC1_0_1.raw
MY_20180821_CatV55_30_4.raw	MY_20180821_PC2_60_4.raw
MY_20180821_CatV55_30_3.raw	MY_20180821_PC2_60_3.raw
MY_20180821_CatV55_30_2.raw	MY_20180821_PC2_60_2.raw
MY_20180821_CatV55_30_1.raw	MY_20180821_PC2_60_1.raw
MY_20180821_CatV55_0_4.raw	MY_20180821_PC2_30_4.raw
MY_20180821_CatV55_0_3.raw	MY_20180821_PC2_30_3.raw
MY_20180821_CatV55_0_2.raw	MY_20180821_PC2_30_2.raw
MY_20180821_CatV55_0_1.raw	MY_20180821_PC2_30_1.raw
MY_20180821_CatL55_60_4.raw	MY_20180821_PC2_0_4.raw
MY_20180821_CatL55_60_3.raw	MY_20180821_PC2_0_3.raw
MY_20180821_CatL55_60_2.raw	MY_20180821_PC2_0_2.raw
MY_20180821_CatL55_60_1.raw	MY_20180821_PC2_0_1.raw
MY_20180821_CatL55_30_4.raw	MY_20180821_CatL55_0_1.raw
MY_20180821_CatL55_30_3.raw	MY_20180821_CatL55_0_4.raw
MY_20180821_CatL55_30_2.raw	MY_20180821_CatL55_0_3.raw
MY_20180821_CatL55_30_1.raw	MY_20180821_CatL55_0_2.raw

Q-Exactive MS Report: Thermo Scientific SII for Xcalibur Method – Next Pages

Method of Q Exactive

Overall method settings

Global Settings	
Use lock masses	off
Lock mass injection	—
Chrom. peak width (FWHM)	15 s
Time	
Method duration	95.00 min
Customized Tolerances (+/-)	
Lock Masses	—
Inclusion	—
Exclusion	—
Neutral Loss	—
Mass Tags	—
Dynamic Exclusion	10.0 ppm

Experiment

Full MS / dd-MS² (TopN)

General	
Runtime	0 to 95 min
Polarity	Positive
In-source CID	0.0 eV
Default charge state	2
Inclusion	—
Exclusion	—
Tags	—
Full MS	
Microscans	1
Resolution	70,000
AGC target	3e6
Maximum IT	100 ms
Number of scan ranges	1
Scan range	250 to 1500 m/z
Spectrum data type	Profile
dd-MS² / dd-SIM	
Microscans	1
Resolution	17,500
AGC target	1e5
Maximum IT	50 ms
Loop count	12
MSX count	1
TopN	12
Isolation window	1.5 m/z
Isolation offset	0.0 m/z
Scan range	200 to 2000 m/z
Fixed first mass	150.0 m/z
(N)CE / stepped (N)CE	nce: 28
Spectrum data type	Centroid
dd Settings	
Minimum AGC target	3.00e2
Intensity threshold	6.0e3
Apex trigger	—
Charge exclusion	unassigned, 1

Peptide match		Preferred
Exclude isotopes	fc	on
Dynamic exclusion		20.0 s
If idle ..		do not pick others

Setup

Tunefiles
General
Switch Count 0
Base Tunefile C:\Xcalibur\methods\CL_nanoTune_20180409.mstune

Contact Closure
General
Used False
Start in Closed True
Switch Count 0

Syringe
General
Used False
Start in OFF True
Stop at end of run False
Switch Count 0
Pump setup
Syringe type Hamilton
Flow rate 3.000 µL/min
Inner diameter 2.303 mm
Volume 250 µL

Divert Valve A
General
Used False
Start in 1-2 True
Switch Count 0

Divert Valve B
General
Used False
Start in 1-2 True
Switch Count 0

Lock Masses
(no entries)

Inclusion List
(no entries)

Exclusion List
(no entries)

Neutral Losses
(no entries)

Mass Tags
(no entries)

1. Notes

2. Result Statistics

Figure 1. False discovery rate (FDR) curve. X axis is the number of peptide-spectrum matches (PSM) being kept. Y axis is the corresponding FDR.

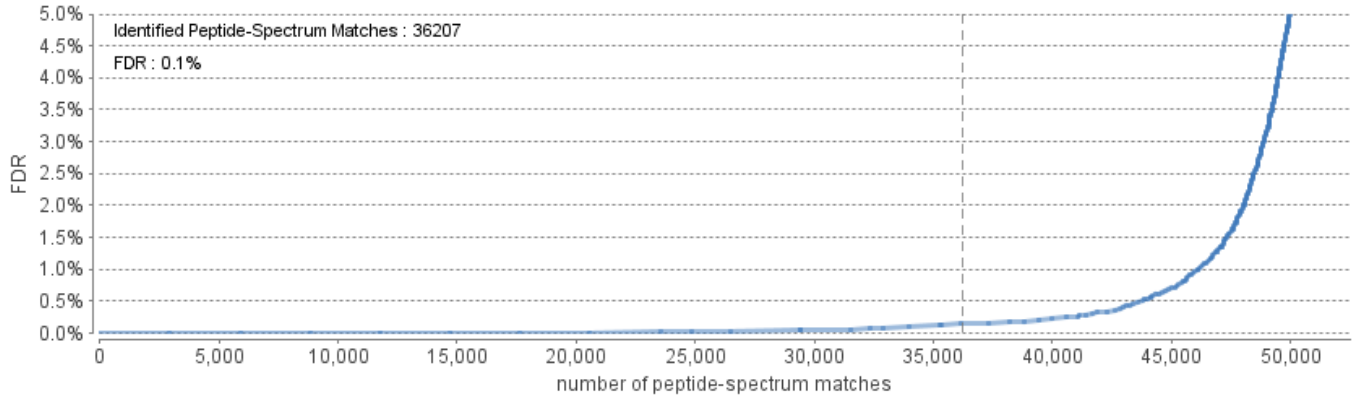

Figure 2. PSM score distribution. **(a)** Distribution of PEAKS peptide score; **(b)** Scatterplot of PEAKS peptide score versus precursor mass error.

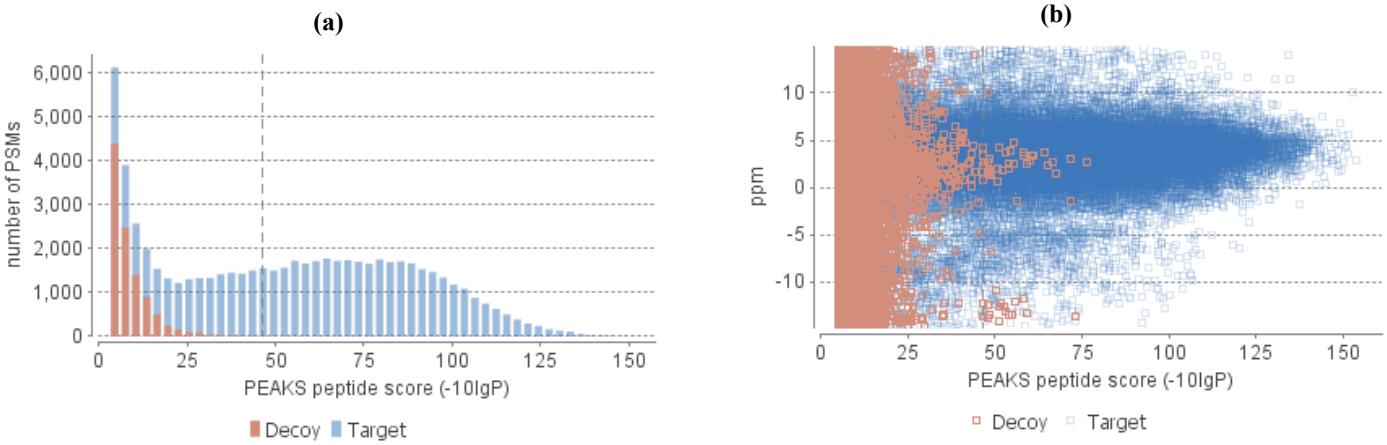

Figure 3. De novo result validation. Distribution of residue local confidence: **(a)** Residues in de novo sequences validated by confident database peptide assignment; **(b)** Residues in "de novo only" sequences.

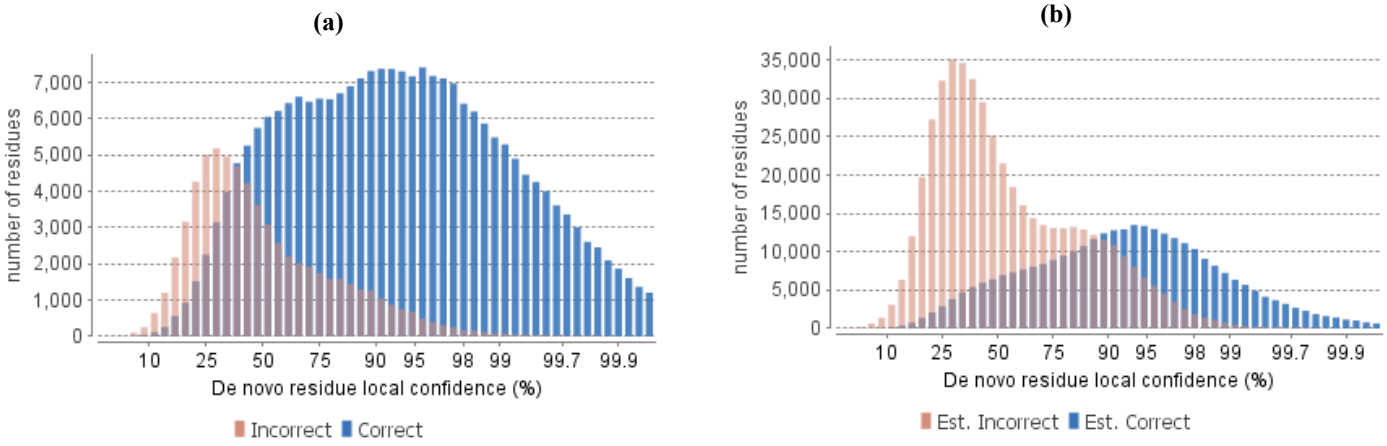

Table 1. Statistics of data.

# of MS scans	183611
# of MS/MS scans	1009554

Table 2. Result filtration parameters.

Peptide -10lgP	≥46.4
Peptide Ascore	≥0
Protein -10lgP	≥20
Proteins unique peptides	≥0
De novo ALC Score	≥50%

Table 3. Statistics of filtered result.

Peptide-Spectrum Matches	36207
Peptide sequences	1144
Protein groups	402
Proteins	402
Proteins (#Unique Peptides)	189 (>2); 12 (=2); 14 (=1);
FDR (Peptide-Spectrum Matches)	0.1%
FDR (Peptide Sequences)	0.9%
FDR (Protein)	85.3%
De Novo Only Spectra	59514

Table 4. PTM profile.

Name	ΔMass	Position	#PSM	-10lgP	Area	AScore

3. Experiment Control

Figure 4. Precursor mass error of peptide-spectrum matches (PSM) in filtered result. **(a)** Distribution of precursor mass error in ppm; **(b)** Scatterplot of precursor m/z versus precursor mass error in ppm.

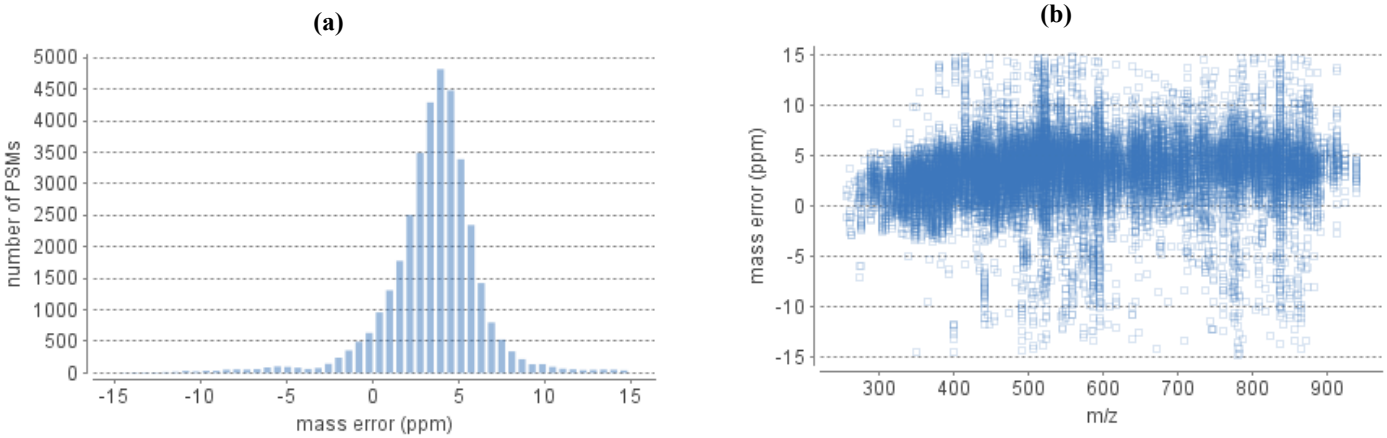

Table 5. Number of identified peptides in each sample by the number of missed cleavages

Missed Cleavages	0	1	2	3	4+
L_0_55_1	0	0	0	0	23
L_0_55_2	0	0	0	0	29
L_0_55_3	0	0	0	0	92
L_0_55_4	0	0	0	0	21
L_30_55_1	0	0	0	0	19
L_30_55_2	0	0	0	0	9
L_30_55_3	0	0	0	0	26
L_30_55_4	0	0	0	0	118
L_60_55_1	0	0	0	0	166
L_60_55_2	0	0	0	0	65

...

4. Other Information

Table 6. Search parameters.

Search Engine Name: PEAKS
Parent Mass Error Tolerance: 15.0 ppm
Fragment Mass Error Tolerance: 0.01 Da
Precursor Mass Search Type: monoisotopic
Enzyme: None
Max Missed Cleavages: 100
Non-specific Cleavage: both
Max Variable PTM Per Peptide: 3
Database: TDP_237library_07202017
Taxon: All
Searched Entry: 228
FDR Estimation: Enabled
Different data refine parameters are used for this search:

Table 7. Instrument parameters.

Fractions: MY_20180821_CatL55_0_1.raw, MY_20180821_CatL55_0_2.raw, MY_20180821_CatL55_0_3.raw, MY_20180821_CatL55_0_4.raw, MY_20180821_CatL55_30_1.raw, MY_20180821_CatL55_30_2.raw, MY_20180821_CatL55_30_3.raw, MY_20180821_CatL55_30_4.raw, MY_20180821_CatL55_60_1.raw, MY_20180821_CatL55_60_2.raw, MY_20180821_CatL55_60_3.raw, MY_20180821_CatL55_60_4.raw, MY_20180821_CatV55_0_1.raw, MY_20180821_CatV55_0_2.raw, MY_20180821_CatV55_0_3.raw, MY_20180821_CatV55_0_4.raw, MY_20180821_CatV55_30_1.raw, MY_20180821_CatV55_30_2.raw, MY_20180821_CatV55_30_3.raw, MY_20180821_CatV55_30_4.raw, MY_20180821_CatV55_60_1.raw, MY_20180821_CatV55_60_2.raw, MY_20180821_CatV55_60_3.raw, MY_20180821_CatV55_60_4.raw, MY_20180821_CatV74_30_1.raw, MY_20180821_CatV74_30_2.raw, MY_20180821_CatV74_30_3.raw, MY_20180821_CatV74_30_4.raw, MY_20180821_CatV74_60_1.raw, MY_20180821_CatV74_60_2.raw, MY_20180821_CatV74_60_3.raw, MY_20180821_CatV74_60_4.raw
Ion Source: ESI(nano-spray)
Fragmentation Mode: high energy CID (y and b ions)
MS Scan Mode: FT-ICR/Orbitrap
MS/MS Scan Mode: FT-ICR/Orbitrap

1. Notes

2. Result Statistics

Figure 1. False discovery rate (FDR) curve. X axis is the number of peptide-spectrum matches (PSM) being kept. Y axis is the corresponding FDR.

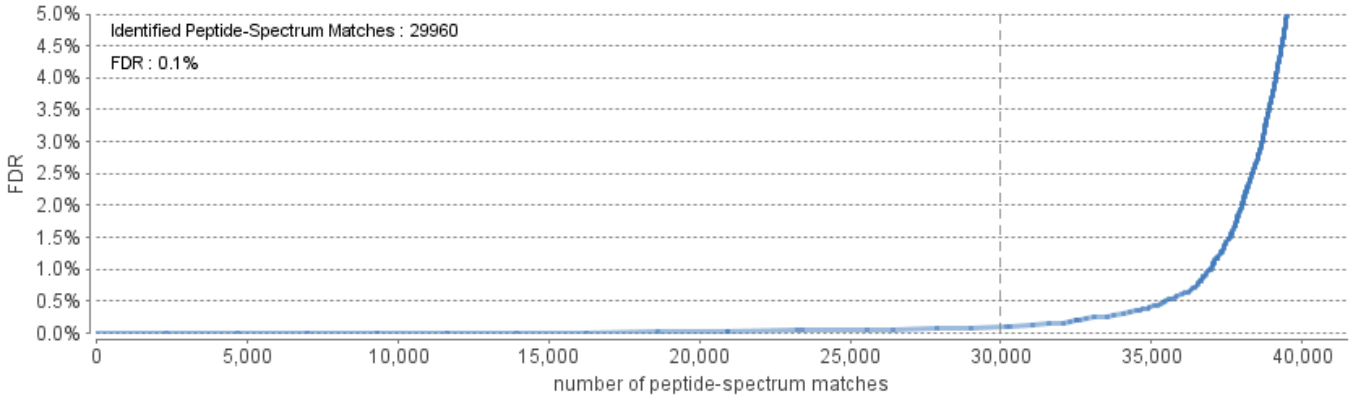

Figure 2. PSM score distribution. **(a)** Distribution of PEAKS peptide score; **(b)** Scatterplot of PEAKS peptide score versus precursor mass error.

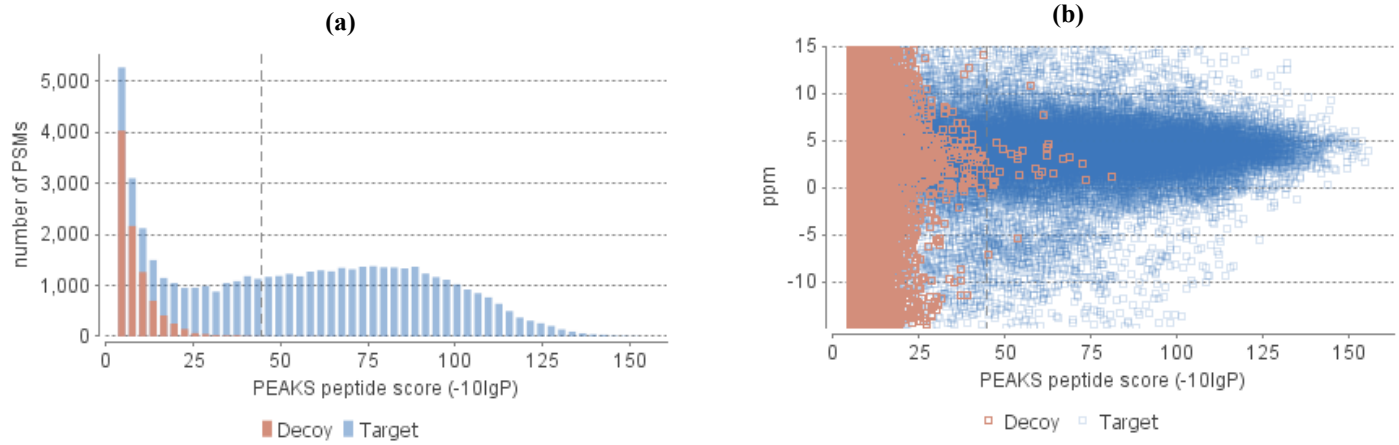

Figure 3. De novo result validation. Distribution of residue local confidence: **(a)** Residues in de novo sequences validated by confident database peptide assignment; **(b)** Residues in "de novo only" sequences.

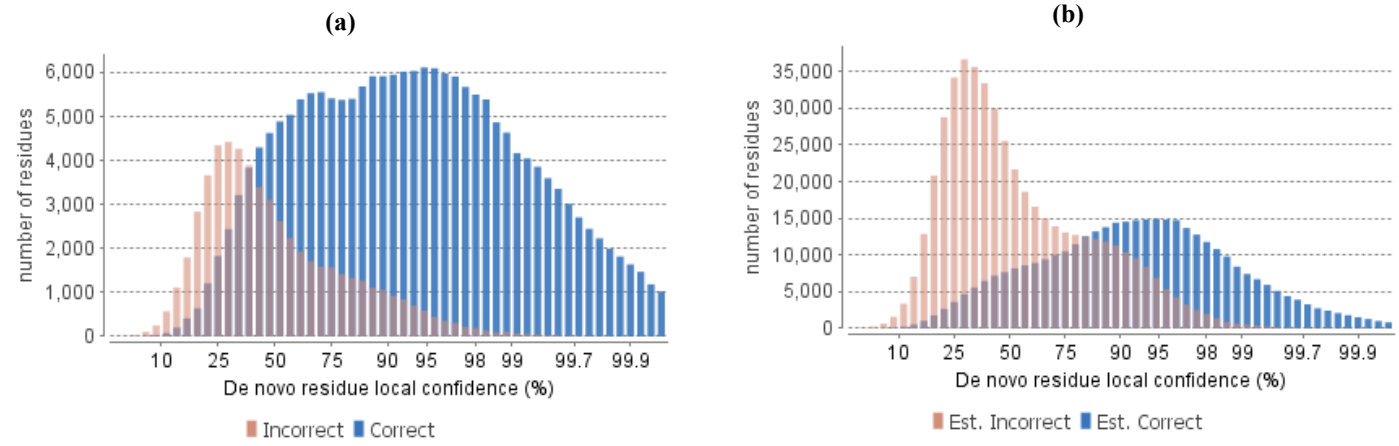

Table 1. Statistics of data.

# of MS scans	133047
# of MS/MS scans	780106

Table 2. Result filtration parameters.

Peptide -10lgP	≥44.7
Peptide Ascore	≥0
Protein -10lgP	≥20
Proteins unique peptides	≥0
De novo ALC Score	≥50%

Table 3. Statistics of filtered result.

Peptide-Spectrum Matches	29960
Peptide sequences	887
Protein groups	393
Proteins	393
Proteins (#Unique Peptides)	152 (>2); 44 (=2); 20 (=1);
FDR (Peptide-Spectrum Matches)	0.1%
FDR (Peptide Sequences)	0.9%
FDR (Protein)	80.3%
De Novo Only Spectra	62036

Table 4. PTM profile.

Name	ΔMass	Position	#PSM	-10lgP	Area	AScore

3. Experiment Control

Figure 4. Precursor mass error of peptide-spectrum matches (PSM) in filtered result. **(a)** Distribution of precursor mass error in ppm; **(b)** Scatterplot of precursor m/z versus precursor mass error in ppm.

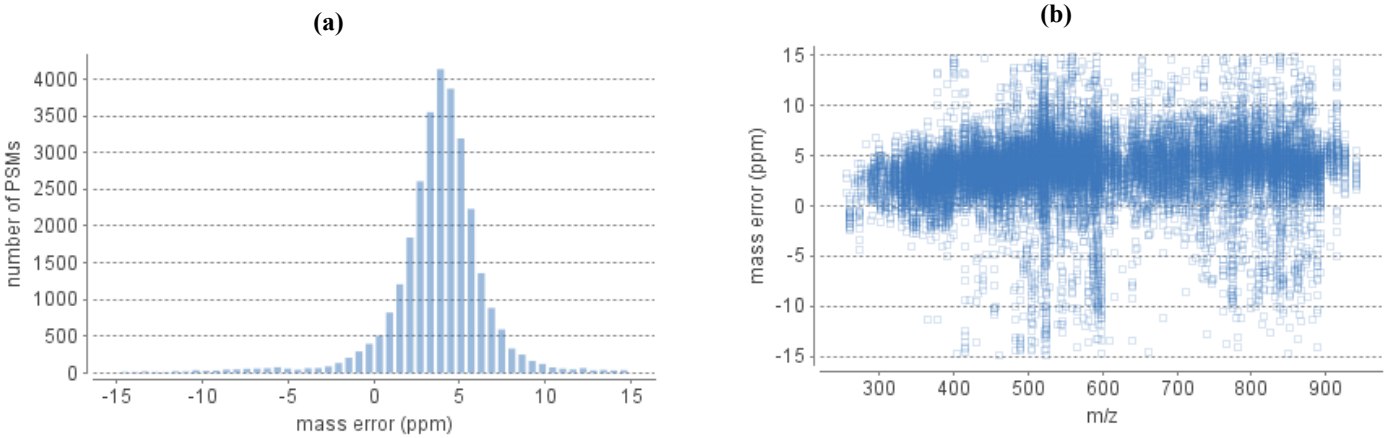

Table 5. Number of identified peptides in each sample by the number of missed cleavages

Missed Cleavages	0	1	2	3	4+
PC1_00_1	0	0	0	0	61
PC1_00_2	0	0	0	0	68
PC1_00_3	0	0	0	0	10
PC1_00_4	0	0	0	0	16
PC1_30_1	0	0	0	0	21
PC1_30_2	0	0	0	0	25
PC1_30_3	0	0	0	0	17
PC1_30_4	0	0	0	0	75
PC1_60_1	0	0	0	0	37
PC1_60_2	0	0	0	0	111

...

4. Other Information

Table 6. Search parameters.

Search Engine Name: PEAKS
Parent Mass Error Tolerance: 15.0 ppm
Fragment Mass Error Tolerance: 0.01 Da
Precursor Mass Search Type: monoisotopic
Enzyme: None
Max Missed Cleavages: 100
Non-specific Cleavage: both
Max Variable PTM Per Peptide: 3
Database: TDP_237library_07202017
Taxon: All
Searched Entry: 228
FDR Estimation: Enabled
Different data refine parameters are used for this search:

Table 7. Instrument parameters.

Fractions: MY_20180821_PC1_0_1.raw, MY_20180821_PC1_0_2.raw, MY_20180821_PC1_0_3.raw, MY_20180821_PC1_0_4.raw, MY_20180821_PC1_30_1.raw, MY_20180821_PC1_30_2.raw, MY_20180821_PC1_30_3.raw, MY_20180821_PC1_30_4.raw, MY_20180821_PC1_60_1.raw, MY_20180821_PC1_60_2.raw, MY_20180821_PC1_60_3.raw, MY_20180821_PC1_60_4.raw, MY_20180821_PC2_0_1.raw, MY_20180821_PC2_0_2.raw, MY_20180821_PC2_0_3.raw, MY_20180821_PC2_0_4.raw, MY_20180821_PC2_30_1.raw, MY_20180821_PC2_30_2.raw, MY_20180821_PC2_30_3.raw, MY_20180821_PC2_30_4.raw, MY_20180821_PC2_60_1.raw, MY_20180821_PC2_60_2.raw, MY_20180821_PC2_60_3.raw, MY_20180821_PC2_60_4.raw
Ion Source: ESI(nano-spray)
Fragmentation Mode: high energy CID (y and b ions)
MS Scan Mode: FT-ICR/Orbitrap
MS/MS Scan Mode: FT-ICR/Orbitrap