Supplemental Figure 1: Over-expressed masses at threshold 0.65 for the CNS ROI in LeechE12a .

Supplemental Figure 2: Effects of normalization on data. a) Average peak intensity for spots before and after normalization. There is a spatial bias for the posterior region in the original data. Spatial bias is removed by normalization. b) Distribution of spot average peak intensities. The average intensity for each laser spots differ greatly. Lack of homogeneity in crystallization can cause this spatial bias. c) Resulting changes in spatial distributions. In both unnormalized images, we see the same bias as in a). At $m/z = 8563.04$, the intensities are more evenly distributed throughout the nervous system are distributed after normalization. At $m/z = 10357.1$, normalizing the data allowed us to see a signal in the head of the leech.

Supplemental Figure 3: Mask (top) and query (bottom) for the CNS in the LeechE12a . We can see in the query that a MALDI spot covers about a 10x10 pixels square on the histological image; thus, the query is a set of gray squares instead of the binary black and transparent pixels of the mask.

Supplemental Figure 4: MALDI images for CNS localization at different scores. a) At score $s = 0.62$, just below our cut-off of 0.65, we still see localization to CNS, but noise signicantly impairs the signal compared to the top-ranked images. b) At score $s = 0.51$ we expect no localization. Indeed, the intensities outside the nervous system almost perfectly balance out the intensities within the CNS. c) At score $s = 0.23$, we are at the other end of the spectrum. We detect a molecule which have inversed expression to the CNS; the molecule is highly expressed in the ventral region of the leech but shows distinct under-expression in the ganglia and the brain.

Supplemental Figure 5: Over-expressed masses in the leech CNS across different samples. We displayed all m/z values in LeechE12b with score $s \geq 0.65$, and all m/z values with score $s \geq 0.70$ in LeechE12a . We only showed m/z values with score $s \geq 0.65$ in LeechE12a if it corresponded to a hit in LeechE12b as the full table is quite extensive (Supplemental Table 1). Results for $m/z < 2200$ were discarded because spectra were still in the noise area.

Supplemental Figure 6: Simulation results decreasing ROI signal over the entire region. Intensities are decreased in the ROI spots by a certain percentage until the average intensity inside the ROI is the same as the average intensity outside the ROI. Score and signal decrease in a similar fashion. Similar intensities in the ROI and in the background lead to a score close to 0.5 as expected.

Supplemental Figure 7: Simulation results when degrading the signal in the ROI. We set a proportion of the ROI spots to have random non-ROI (or background) intensities (see Methods). It is then possible to balance the total ROI intensities by distributing the subtracted intensity to the remaining spots; that way, the total intensities in ROI and outside ROI remain the same throughout the simulation. Results are shown for two simulated runs: with and without balancing the ROI intensities. In both cases, the score linearly decreases as more ROI spots are set to background intensity. In the balancing case, the intensities of the remaining spots increase to compensate for the other spots; consequently, the score remains higher in the balanced case than in the unbalanced case, as expected. When all spots are are set to background intensities, the signal is lost and the probability score decreases to 0.5 as expected.

Supplemental Figure 8: ClustalW alignment of the HmIF4 protein sequence with those of three other known intermediate filaments in *Hirudo medicinalis*. The EST open reading frame aligned particularly well in the conserved rod domain, and has more variability outside of that domain. The peptide we identified is located in a variable region in the 5' end of the rod domain where the sequences are quite dissimilar, thus confirming the discovery of a novel protein.

Supplemental Figure 9: Annotated spectrum for a peptide from the histone H2B. Parent mass $\simeq$ 2500 Da was shown to have a CNS specific expression. *In situ* hybridization of the mRNA shows a preferential location in the CNS, but with a relatively weaker signal.

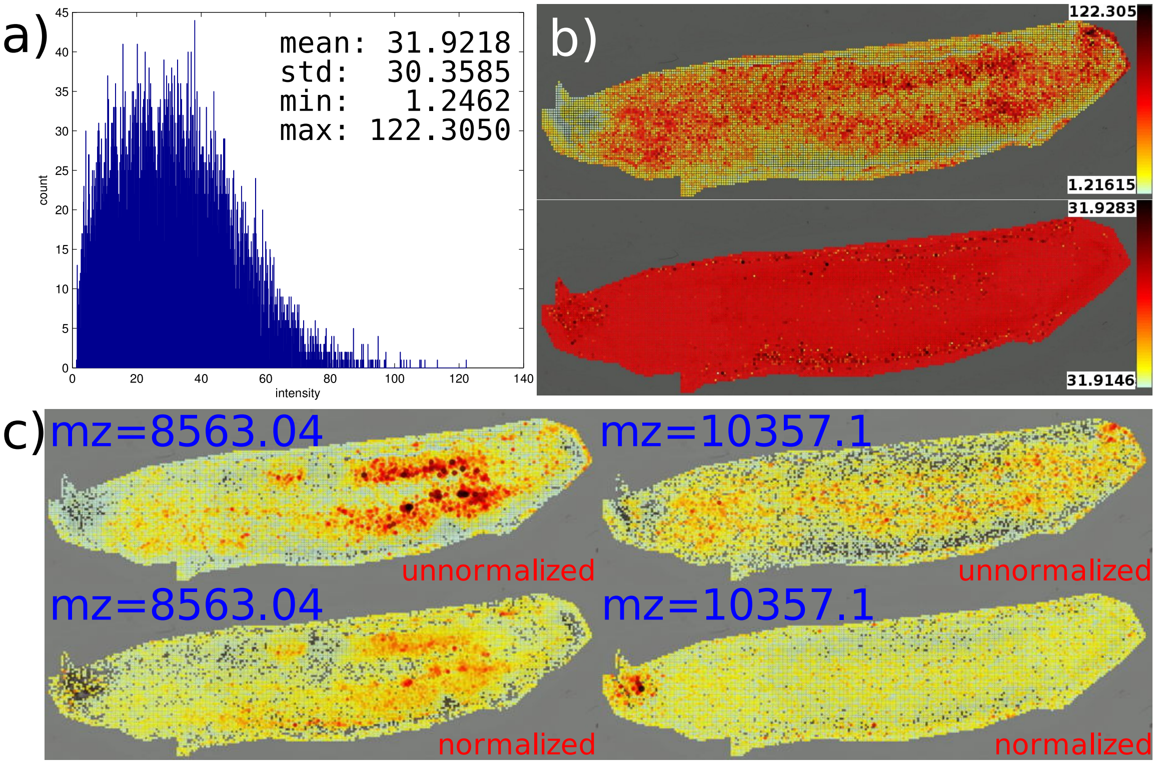
Supplemental Figure 10: Annotated spectrum for uncharacterized peptide. Parent mass $\simeq$ 3841 Da was shown to have a dorsal specific expression similar to another fragment 2 amino acids shorter showing in Figure 5.

Supplemental Figure 11: Distribution of the MS1 raw data peaks for 2 experiments. Intensity of top peaks (162.2 and 1526.5) for annotated spectra are indicated by red line.

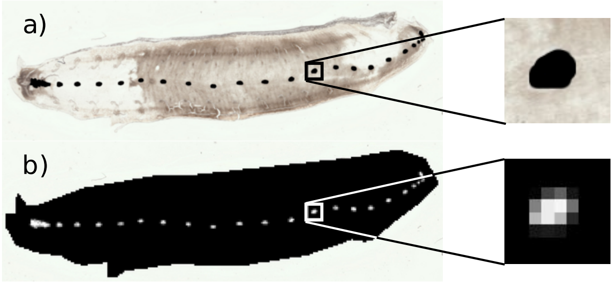
Supplemental Figure 12: Identification of several other peptides. Peptides were identified using Inspect with an FDR cut-off of 0.01. Biological annotations were achieved by doing a Blast search of the protein sequence against NCBI nr and keeping the top hit if the E-value was less than or equal to $1e$-5.

m/z-range	peak	score
2470.68-2481.44	2474.34	0.80
2505.32-2508.31	2506.7	0.69
2521.69-2527.93	2524.69	0.78
3090.6-3096.75	3093.93	0.71
3157.97-3165.99	3160.3	0.71
3295.94-3303.86	3299.64	0.76
3485.31-3496.45	3491.01	0.80
3496.45-3531.63	3511.43	0.90
3536.01-3542.31	3540.11	0.71
3651.01-3651.57	3651.29	0.65
3652.4-3652.4	3652.4	0.65
4211.19-4214.18	4211.79	0.67
4371.25-4383.43	4378.86	0.69
4523.78-4529.97	4526.88	0.69
4539.58-4547.64	4542.06	0.70
5268.46-5278.82	5273.47	0.76
5279.83-5280.16	5279.83	0.66
5410.4-5429.38	5417.85	0.81
5430.06-5432.43	5431.41	0.68
5434.81-5436.5	5435.48	0.68
5560.06-5606.15	5574.15	0.90
6024.92-6024.92	6024.92	0.65
6025.63-6026.7	6026.7	0.66
6186.44-6186.8	6186.44	0.65
6187.53-6200.57	6196.95	0.68
6212.91-6212.91	6212.91	0.65
6216.54-6220.89	6217.99	0.66
6223.8-6223.8	6223.8	0.65
6593.09-6593.46	6593.09	0.65
6594.58-6595.33	6595.33	0.66
8205.81-8205.81	8205.81	0.65
8207.89-8216.66	8213.32	0.66
8217.49-8218.33	8218.33	0.65
8220-8220	8220	0.65
8223.76-8223.76	8223.76	0.65
8399.29-8400.14	8399.72	0.66
8400.98-8447.91	8428.45	0.88
8448.76-8449.61	8449.18	0.66
9032.26-9041.02	9038.39	0.74
9224.17-9249.86	9240.11	0.70
9250.74-9257.83	9253.4	0.67
9779.68-9779.68	9779.68	0.65
9780.59-9781.05	9780.59	0.65
10868.4-10869.8	10869.4	0.67

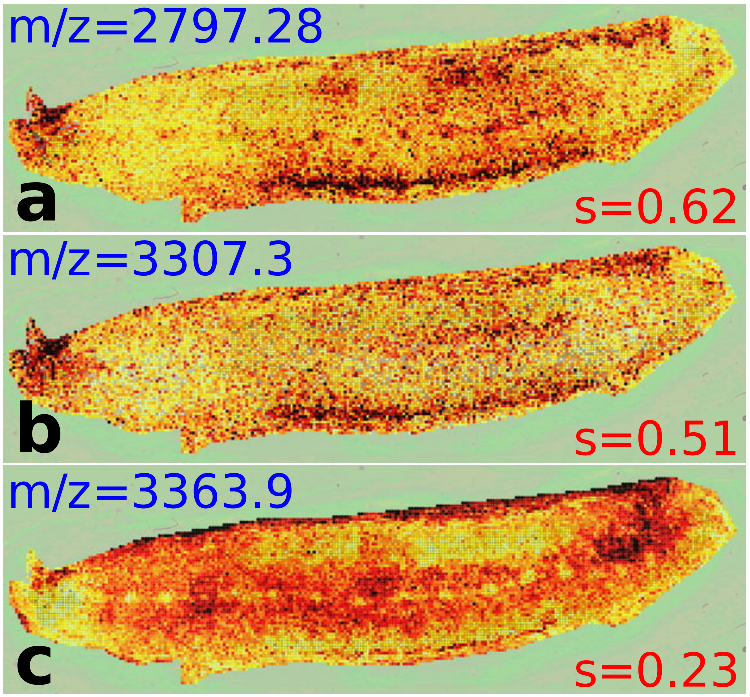
Supplemental Figure 1



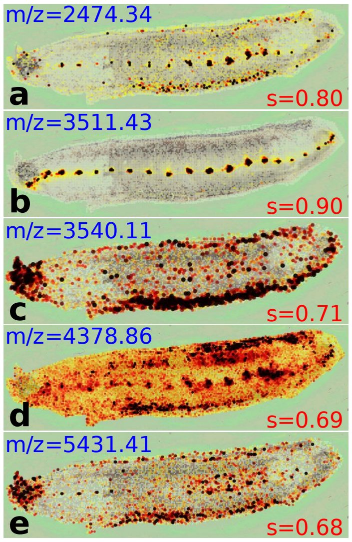
Supplemental Figure 2



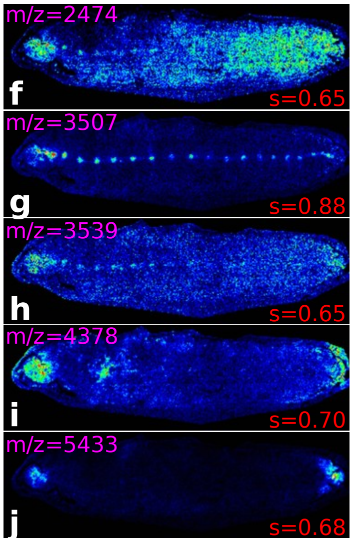
Supplemental Figure 3



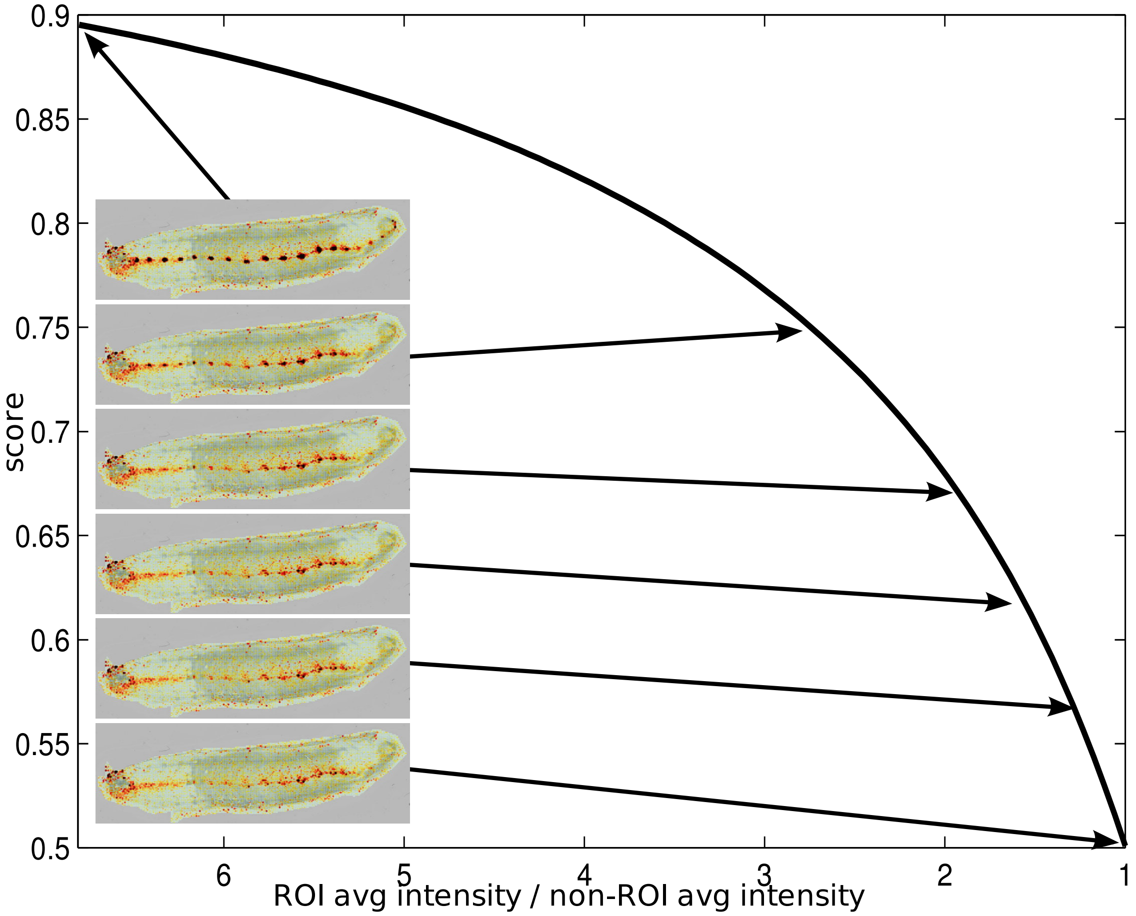
Supplemental Figure 4



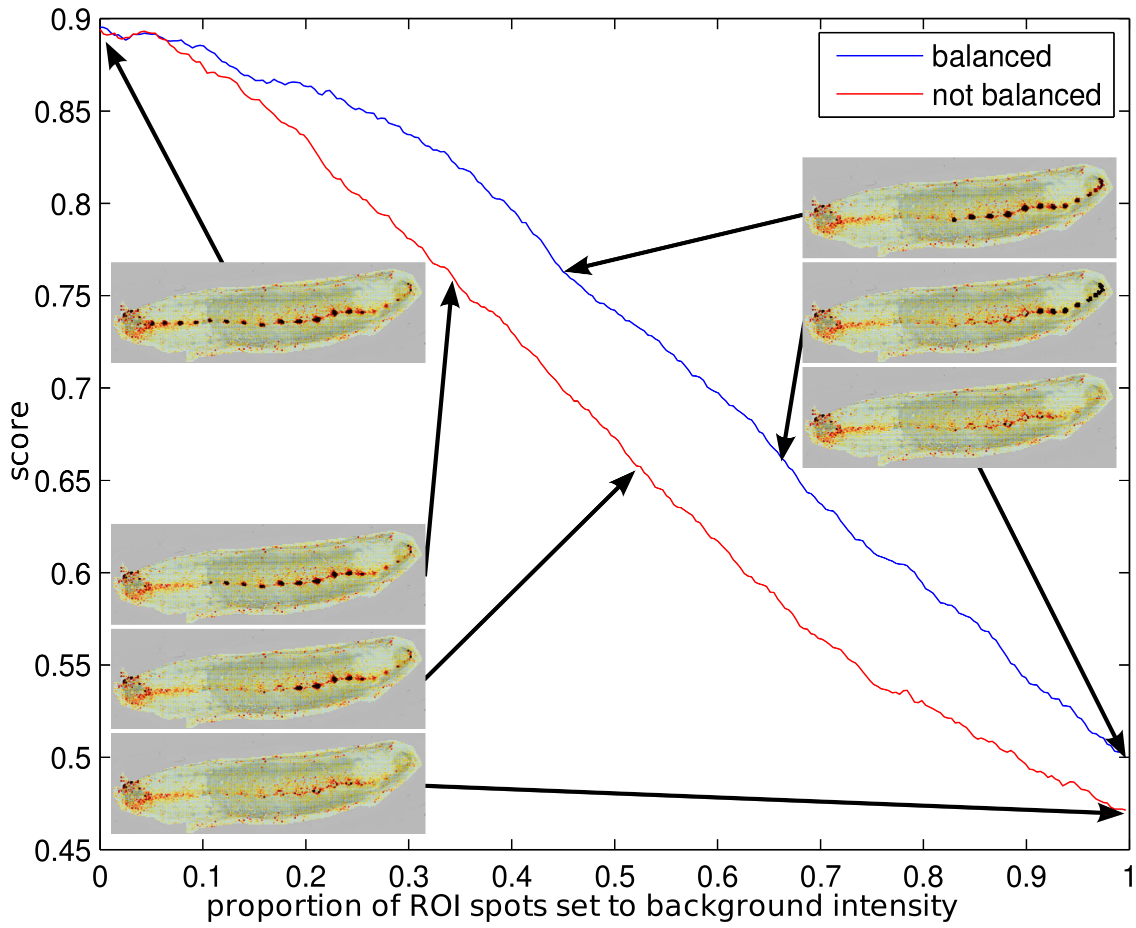
LeechE12a		LeechE12b	
m/z peak	score	m/z peak	score
2474.34	0.797	2474.34	0.653
2479.6	0.727		
2524.69	0.78		
3093.93	0.705		
3160.3	0.708		
3299.64	0.76		
3491.01	0.797		
3511.43	0.896	3507.89	0.882
3523.98	0.713		
3540.11	0.71	3539.84	0.654
4378.86	0.691	4377.95	0.697
4542.06	0.702		
5273.47	0.758		
5417.85	0.814		
5426.66	0.708		
5431.41	0.675	5433.79	0.681
5574.15	0.899		
5586.87	0.805		
5601.33	0.717		
8428.45	0.899		
9038.39	0.736		
9240.11	0.701		



Supplemental Figure 5



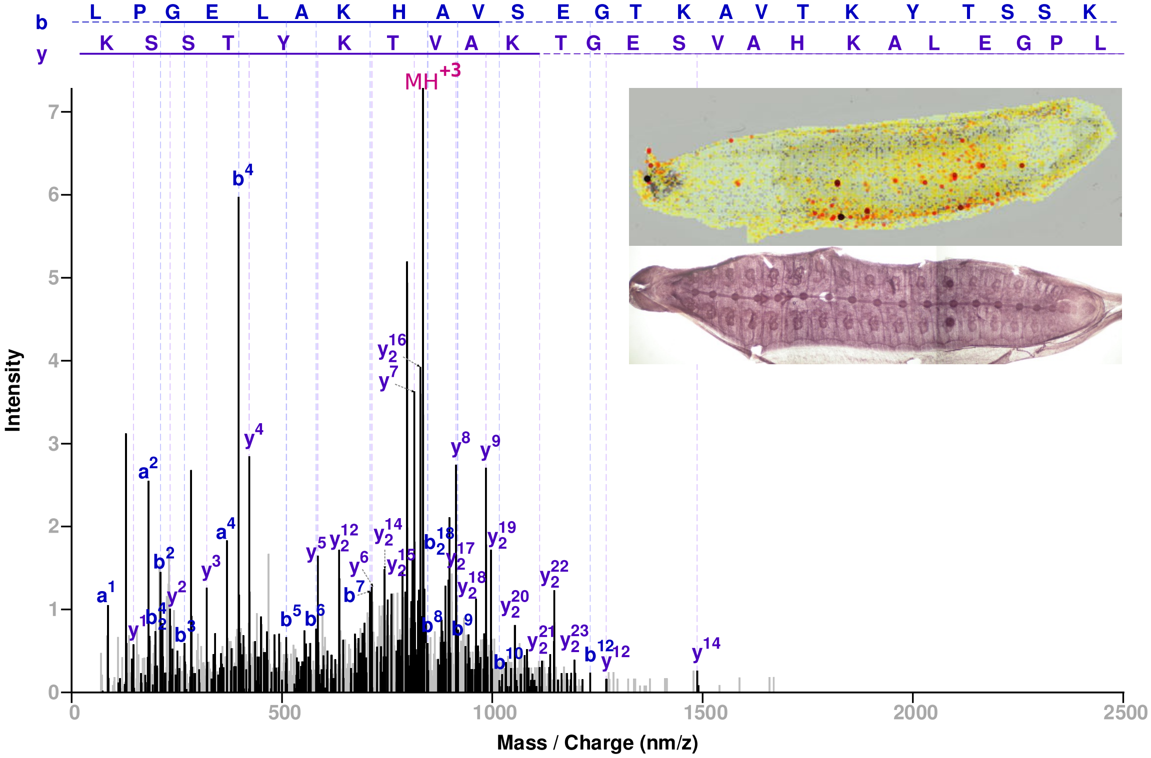
Supplemental Figure 6



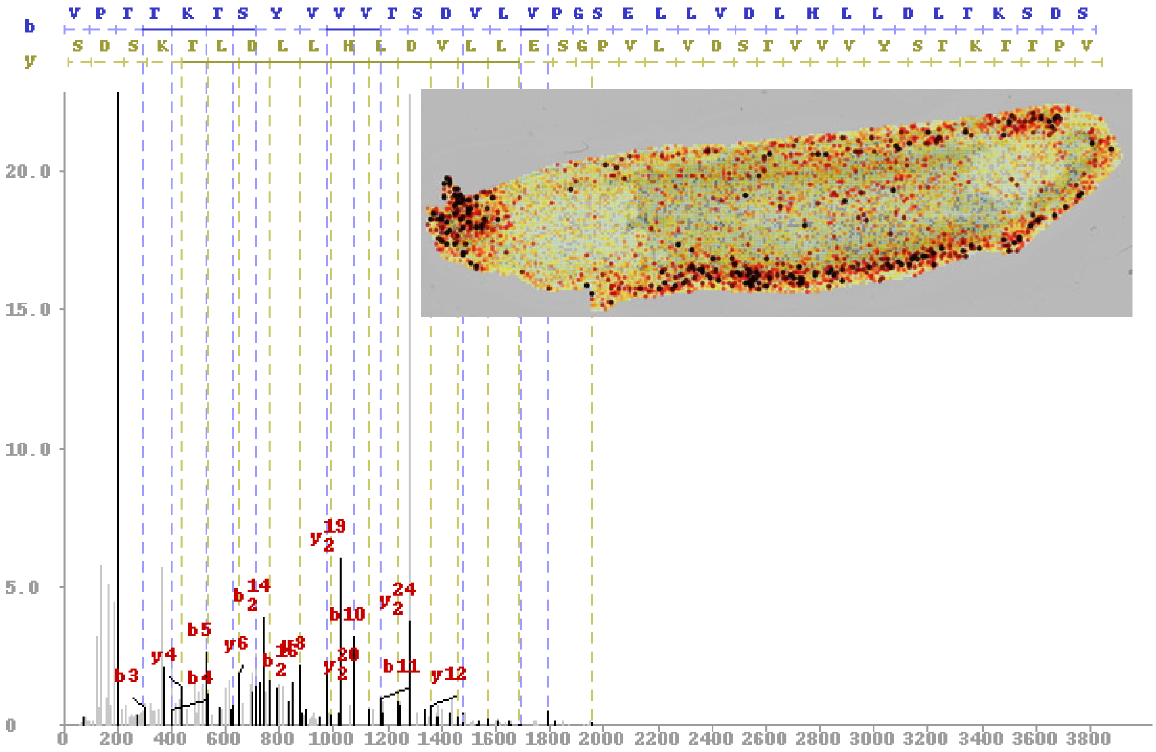
Supplemental Figure 7

```
CLUSTAL 2.0.12 multiple sequence alignment


leech_EST       --TFTEMERTRTVRNTTYTTSSGVDGNVDGGASTFRES--SYGAGGPITG 46
gliarin         ---MAEEVITKTTRRVYKTEVSGGEGGSSILATTYR----------PSVT 37
filarin         ------MESGNFVAEYERTIRSN------------------------IQ 19
macrolin        ARQIEPLGMDTRIDKSSYNVSQSGDKVVSTKSMVDLDTGSTYSTYKATVV 50
                              .   .  ..

leech_EST       GRTLVISRTIGSTGSPGMGSIGLGGTRTMERSVRTSSQYASGGPMPNYSV 96
gliarin         PRNVIIHRSIQAPLGSSMSSS----TIRREKTIQYGNAYAISP--SSYAP 81
filarin         PRNILIQRATPGAFTNVSSRS--------VVTRSVGNVYGAGG------- 54
macrolin        PHQLIIQRTLTGGLSSGGGSLS------RSTADRFRSVMVPGVH------ 88
                 : ::* *:  .      .            :    .    .
                               ▷rod domain
leech_EST       ITATGVSGIKESRDQEKKDMQDLNERFANYIDKVRNLEAQNRKLAEDLSR 146
gliarin         LASSGVSSVKNSRERERKKDMQDLNERFASYIEKVRFLEAQNKRLTDELDK 131
filarin         VAGGAATSVTDQRKNEKREMQDLNERFAGYIEKVRFLEAQNKKLADELDA 104
macrolin        LATKEVDSARFTRDREKKDMQDLNLRLTRYIETVRFLEAQNKQLDNEIKT 138
                ::   . .    *..**::***** *:: **:.** *****::* :::.

leech_EST       LKEKWGKDTVQVKAMFQVDLDECRHQLDEAEKEKARLEIRLASLEEELED 196
gliarin         LKSRWGKDTTQIKAMFQVELDEARRLLDDGEKEKARLEIKIASLEEINEE 181
filarin         LKSRWGKQTTVIKTMYQTDLDQVKRLLDDCEKETQRLQIQVASMEEKVDD 154
macrolin        LKAKWGKETSQVRAMFEADLEEARRIKDDLEKDTAKLEIRISSVIEALDV 188
                ** :***:*  :::*::.:*:: ::  *: **:. :*:*:::*: *  :

leech_EST       LRQELALAAQQLSENQAFISKNNQLLIDYESEIQGLRKRIEQLENEKERD 246
gliarin         LAVKLNEALQTNEEQRQKIDRQNQQLSDYEGEISLLRRRVEGLEADKDKD 231
filarin         LRRKLDEANAAVDESRDKLEKQIQQIAEIQSEVHLLRLRSDLLDGDKRYN 204
macrolin        EKRRNATSEKTIIEYREKIENQNRQLVDLQANNDLLQRRLELLEGDRDRD 238
                   .   :     * :  :..: : : : :.:   *: * : *: ::  :

leech_EST       KKNIAQLKELLAKARQDLDNETLEHIHAENRCQTLQEEIDFLKSIHEQEM 296
gliarin         RKTIATLNAALNTARANLDDETLRHIDAENRRQTLEEELEFLKSVHEQEL 281
filarin         KAILSKLQENLNRARTDFDAQAVEHDAAEAKRLALEEELAFIKELHEQEL 254
macrolin        KKLVGELKEAVTRYRTDLDSQTLLYVDADNRRQSLEEELDFLKQVHEQEM 288
                :  :. *:  :   * ::* ::: :  *: :  :*:**: *:*.:****:

leech_EST       KELAALAYRDTAPE-RDYWKNEMAQALREIQEMYDDKFDSIRTELETHYT 345
gliarin         KELAALAYRDTTTENRDFWKNEMGNALREIQEMYDEKLDLMRTEIESSYT 331
filarin         RELAAKAYFDSTASNREYWKSEMSMELKKLQEHYGEKIDELQNEMSLNYS 304
macrolin        KELNILILKDYSIVNRQYWKTEMERALKEIQDLYDDELDSMRDETETFYQ 338
                :**      * :   *::**.**   *:::*: *.:::* :: * .  *

leech_EST       TKVQEYKTGVARAVTETQRIKEDTTKLRQDLTDLRDKLNELTSQNTSLTR 395
gliarin         LKLQEFRTGATKQNLESTHTKEETKRLKIQVTDLRDKLSDLEGKNLQLVR 381
filarin         MQVQSLN--LPRPAIVST-TKEESVRIRMQTNEIRTKVLEFEGRNDMMLR 351
macrolin        LKIQEIRNSSQRSALEVDQAKDTAKKHKSNVIELRDRVTILEGQNTNLQN 388
                 ::*. .    :        *: : : : :  ::* ::  : .:*  : .

leech_EST       EFEALRRKKEDQERDLEYQRTQLHAEASTLRFELERILSDLQRIMDTKLG 445
gliarin         ELENLRRSKEELEREFEHENGELKAEIARLRAELESIIQELQNIMDTKLG 431
filarin         EIEEMRRDMEEREIEVMKDIDATKSEIITMKAEMDAITKELEALLDAKLS 401
macrolin        ELDFFKLESEHRERDLEVENDTLRLEACKYKAELESLWIEIDKIRSAKDG 438
                *:: :: . *. * :.  :    : *    : *:: :  ::: : .:* .
                     rod domain  ◁
leech_EST       LELEIAAYRKLLEGEENRLKQISMQSGGSGFAYGGGAGAGGSGGAGG--A 493
gliarin         LELEIAAYRKLLEGEESRVGLKQLVEMYSG---GGGGGVAISGGLSGGLG 478
filarin         LELEIATYKKLLEGEADGEGLRQVVDNMFD------SYASATASAAAAYA 445
macrolin        LELEIAAYRKLLEAEEGRFGMEKIIEKLRL------DKHDCSQHVKYRVT 482
                ******:*:****.* .     .: .               :

leech_EST       GGAGGAGSYGYSS----YGSASGFASGGGNAATSVVGQMSAKTTYHKSAL 539
gliarin         GGSSYNESYSYSSSNAASASAGGLAGGAFSSGVLSKGEISARTTFQKSAK 528
filarin         DGIYGEGYNGSAAG---FSSSTSLSRKITGSSSLVGG----RTSYQRNAR 488
macrolin        ENYSGD------------SQSTINQSVKKETTS-------KTSVQKSSK 512
                 .                 *   :                 :*: ::.:

leech_EST       GPISISNCTPDGKLVELENTGTREENIEGWKIVRVVDGKDQPEFKLDGRF 589
gliarin         GTTSIAETSPDGKYVLIENSGRKTESLGGWRLNRVVDDVEVVNFVFPSDL 578
filarin         GPVSISECSPDGKFIVLENTGKRREELGGFTLKRKVDGKDVPVYTFKADA 538
macrolin        GPVAIAECSMDGKFIVLENTGRKDEQLGGYKIRRNINGLDKVEFNFDRNF 562
                *. :*:: : *** : :**:* : *.: *: : * ::. :   : :

leech_EST       SSLKRGQKIGIYARGAKPRTAGARDIEANFESWGIGAQATTKLINPEGEE 639
gliarin         K-LNAGEKFKVWGAGQKPINASSNDVEANVDNFGIGSNILTRLYNTLGEE 627
filarin         G-VDPHMKIKLWVKGAKPSNATASDIEVDIIDWGTGQNITTQLFNSSGQD 587
macrolin        V-LRAGAKIKIWANKLRPLSAFSSDLEADFPSWGVGERIVTGLINQSGEE 611
                  :    *: ::    :* .* : *:*.:. .:* * .  * * *  *::

leech_EST       KATHTQKTVYTS-------------------------- 651
gliarin         RATHVQKTVYG--------------------------- 638
filarin         RATHIQKTLA---------------------------- 597
macrolin        RASYMQIVPRTVLEPCPRPALEGPTDHPLRHFFLDPQI 649
                :*:: * .
```

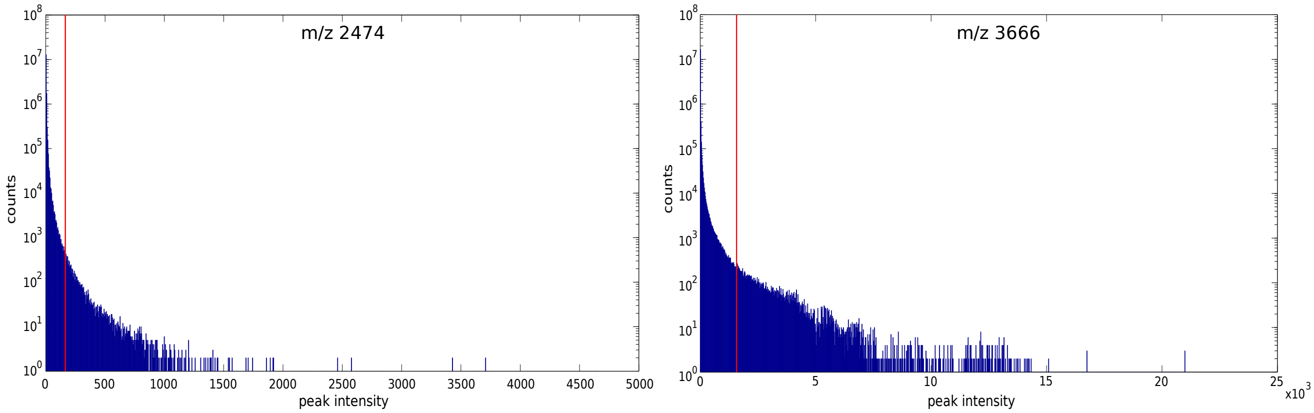
Supplemental Figure 8



Supplemental Figure 9

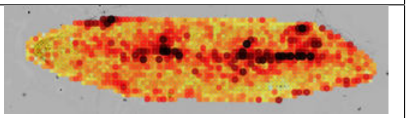
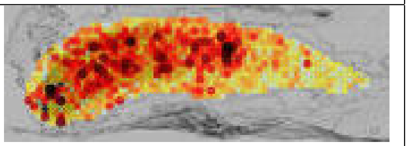
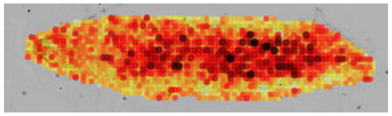
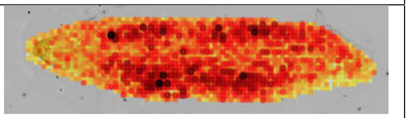
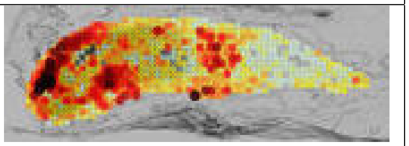
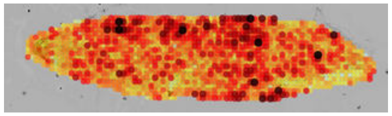
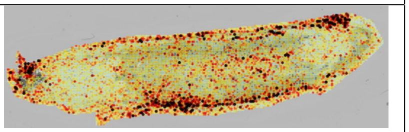
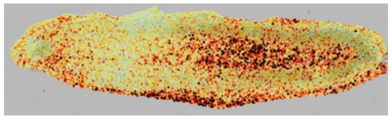
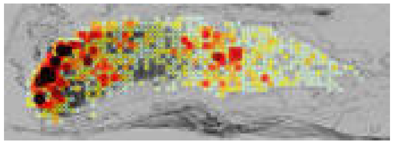


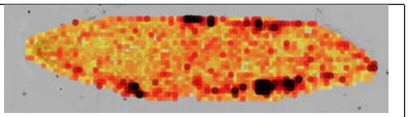
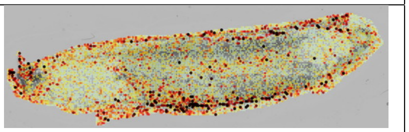
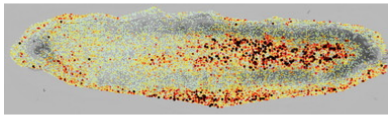
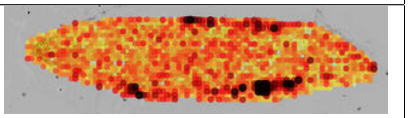
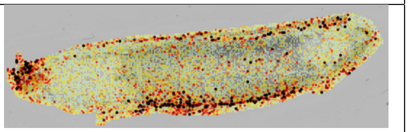
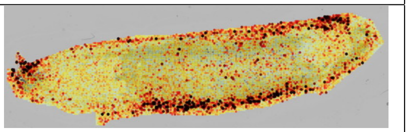
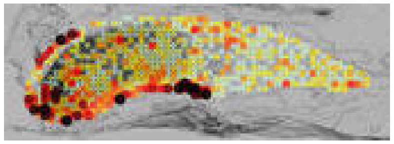
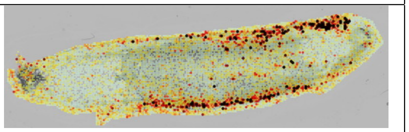
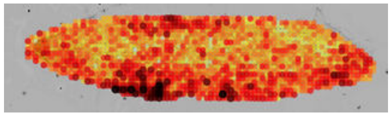
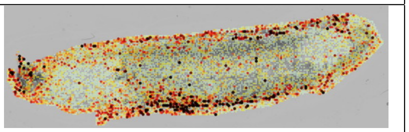
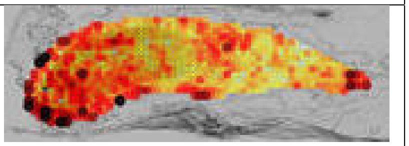
Supplemental Figure 10

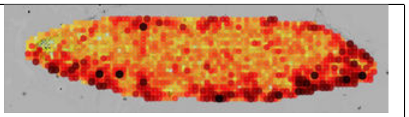
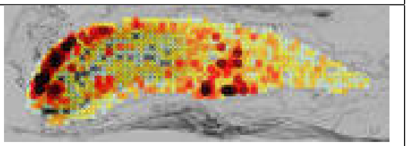
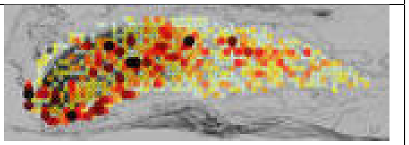
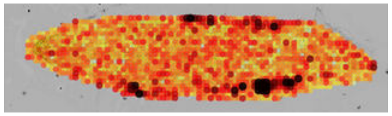
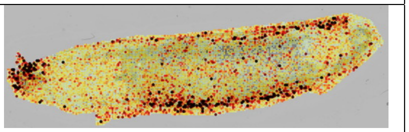
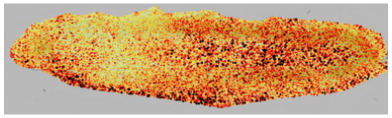
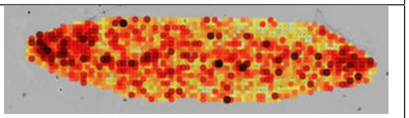
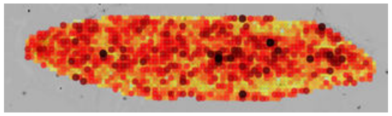
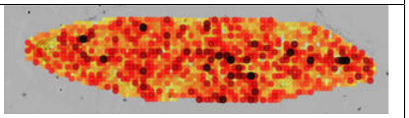
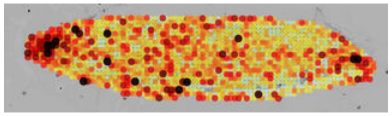
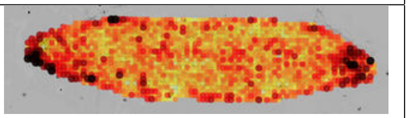


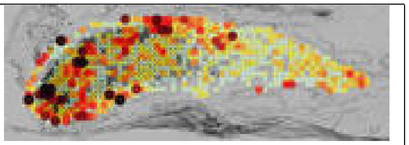
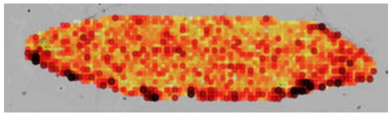
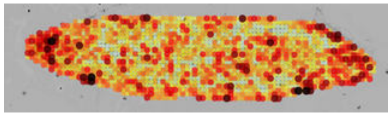
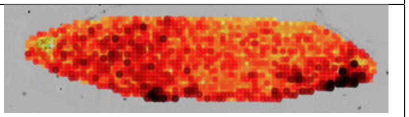
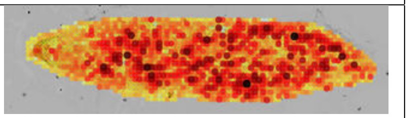
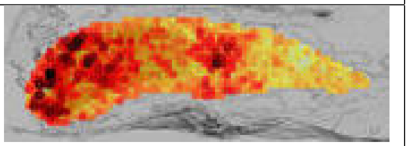
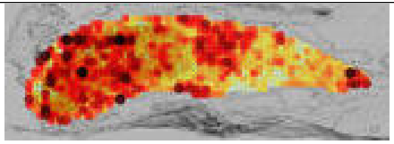
Supplemental Figure 11

Supplemental Figure 12

parent mass	annotation	signal	blast top hit ($E \leq 1e\text{-}5$)	image
1353.81	G.ASGGIGQPLSLLLK.L	CNS	`gi|75040807|sp|Q5NVR2.1|MDHM_PONAB` RecName: Full=Malate dehydrogenase, mitochondrial	N/A
1993.94	T.SYIEDFDVSTLPEHQLT.G	CNS	N/A	
2107.00	L.SLWANNSEKINFQ-17LDGNSS.R	ventral	`gi|22095553|sp|Q9R0M0.2|CELR2_MOUSE` RecName: Full=Cadherin EGF LAG seven-pass G-type receptor 2; AltName: Full=Flamingo homolog;	
1829.91	K.SDQVHGLFSVNVDREK.C	ventral	`gi|117949389|sp|Q6YHK3.2|CD109_HUMAN` RecName: Full=CD109 antigen; AltName: Full=150 kDa TGF-beta-1-binding protein; AltName: Full=C3 and PZP-like alpha-2-macroglobulin domain-containing protein 7;	
1940.00	L.FFTQFTPLFFKPGLSY.V	lateral	N/A	
1633.84	L.FFSQFTPAFFKPGL.T	dorsal	`gi|117949389|sp|Q6YHK3.2|CD109_HUMAN` RecName: Full=CD109 antigen; AltName: Full=150 kDa TGF-beta-1-binding protein	 
2598.29	R.HVADIRADDFSKEGQGVIGLQAGTN.Q	dorsal	`gi|584954|sp|Q08093.1|CNN2_MOUSE` RecName: Full=Calponin-2; AltName: Full=Calponin H2, smooth muscle;	  

1916.89	R.NVAEVPNVADENDFPSLS.I	dorsal	`gi|52783213|sp|Q9CY58.2|PAIRB_MOUSE` RecName: Full=Plasminogen activator inhibitor 1 RNA-binding protein; AltName: Full=PAI1 RNA-binding protein 1;	
2456.10	D.ASNQDTTTCIYISSRYENYLN.A	dorsal	`gi|88909565|sp|Q64449.2|MRC2_MOUSE` RecName: Full=C-type mannose receptor 2; AltName: Full=Lectin lambda; AltName: Full=Macrophage mannose receptor 2;	 
1918.04	N.ILSVGIGSAPYMSPQLTAI.A	dorsal	`gi|189082905|sp|A2AX52.2|CO6A4_MOUSE` RecName: Full=Collagen alpha-4(VI) chain;	
2977.39	N.GRKLEEDDEVPDLVENFDEASKTEVN.M	dorsal	`gi|66774043|sp|Q64152.3|BTF3_MOUSE` RecName: Full=Transcription factor BTF3; AltName: Full=RNA polymerase B transcription factor 3	
2842.66	M.TKNHLNASMIIKTFLKIPGLITLYL.L	dorsal	N/A	 
2467.33	R.PNLAPVTRDVLVTFVDQSHVLF.F	dorsal	N/A	
1869.07	S.YVGQVLLKPVDEVLPLS.S	dorsal	N/A	
2447.36	K.PNLVVPVTRDVLVTFVDQSHVL.F	dorsal	N/A	
1270.67	R.GEEVVLEVHLF.N	dorsal	`gi|259016204|sp|Q8IZJ3.2|CPMD8_HUMAN` RecName: Full=C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8	

1887.98	*.AKFEDKESGKVVTAHLE.E	dorsal	N/A	
1154.63	T.KRKYVAFDQ.K	dorsal	N/A	
1918.01	L.QLWFTCLTFHGIIQ+1PI.N	dorsal	N/A	 
2949.65	N.LRFVRRRSSLRQHYLLRKFDW+16T.C	dorsal	N/A	 
1735.87	L.FFSQFTPVFFKPGL.S	not specific	`gi|81879137|sp|Q8R422.1|CD109_MOUSE` RecName: Full=CD109 antigen; AltName: Full=GPI-anchored alpha-2 macroglobulin-related protein	
1661.87	E.FTPVFFKPGLSYVG.Q			
1967.99	L.FFSEFTPQFFKPGLLY.V	not specific	`gi|259016204|sp|Q8IZJ3.2|CPMD8_HUMAN` RecName: Full=C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8	
1957.96	L.VQNVPEFNEAVPSSRPTS.V	not specific		
1673.78	R.SFAPMSLKDASKDNY.F	not specific	N/A	

1318.67	R.GEEFVLEVHLF.N	not specific	`gi|117949389|sp|Q6YHK3.2|CD109_HUMAN` RecName: Full=CD109 antigen; AltName: Full=150 kDa TGF-beta-1-binding protein	
1842.90	L.VQNVPEFNEAAPSSRPT.S			
1824.17	E.VVLKPKKFEVVTVPLK.T			
1814.90	R.CQSFLKSILSHCQAHL.V	not specific	`gi|55583881|sp|Q68Y21.1|GRID2_DANRE` RecName: Full=Glutamate receptor delta-2 subunit; Short=GluR delta-2 subunit;	
2073.02	G.YDERQQSSSAVETKGFKL.R	not specific	`gi|215273864|sp|Q08174.2|PCDH1_HUMAN` RecName: Full=Protocadherin-1; AltName: Full=Cadherin-like protein 1; AltName: Full=Protocadherin-42; Short=PC42;	
1232.64	C.ASQAGMTAIGAVR.H	not specific	`gi|6225157|sp|Q15417.1|CNN3_HUMAN` RecName: Full=Calponin-3; AltName: Full=Calponin, acidic isoform	
1284.58	N.QQ-17PVHMGMPYN.N	not specific	CHECK `gi|55584092|sp|Q08473.3|SQD_DROME` RecName: Full=RNA-binding protein squid; AltName: Full=Heterogeneous nuclear ribonucleoprotein 40;	
1168.53	S.VTSSSSASIGEGS.R		`gi|205371790|sp|Q9P2Q2.3|FRM4A_HUMAN` RecName: Full=FERM domain-containing protein 4A	N/A
1082.66	Q.LLGLGINIGVN.I		`gi|34921426|sp|O96790.2|DPGN_DIPMA` RecName: Full=Serine protease inhibitor dipetalogastin; (E = 5.7e-04)	N/A
824.32	D.DGPFCATN.G		`gi|1346041|sp|P47931.1|FST_MOUSE` RecName: Full=Follistatin; Short=FS; AltName: Full=Activin-binding protein;	N/A