Use CEBS to explore proteomics data.

Objective: Identify a study in CEBS that generated proteomics data; use the CEBS tools to browse the data, study the differentially expressed proteins, and download gel images, MS spectra and peptide identification data.

Detailed work flow:
Go to the CEBS home page (www.cebs.niehs.nih.gov/prototype/)
Select the "Display All Studies" and then use the data type icons on the right to identify the investigation and study with associated proteomics data (APAP_Oral_F344_Proteomics). The blue "PR" icon in the "Available Data" column gives this information; the legend for the icons is at the bottom of the page.

Use the links under the "Design Information: column to explore the Timeline, Experimental Groups ("Design") and Details of the Study.

Reach the treatment protocols by selecting Timeline, then select protocol link next to "Treatment". These show that two doses each of two different compounds, acetaminophen and the "non toxic" isomer 3'-hydroxyacetanilide were applied to the study animals. These treatments are also reflected in the experimental groups seen in "Design" (to get here use the back button twice then click on "Design" in the "Detailed Information" column).

Tick the check box next to the APAP_Oral_F344_Proteomics Study, then scroll down to the bottom of the page and select Browse Proteomics Data. You will be brought to the Proteomics browser. From this page you can see that two tissues and several subcellular fractions were sampled. Choose to show or hide the subject list (using the show / hide subject data link) to confirm which subjects are included in each analysis. Recall that APAP = acetaminophen, AMAP = "inactive isomer".

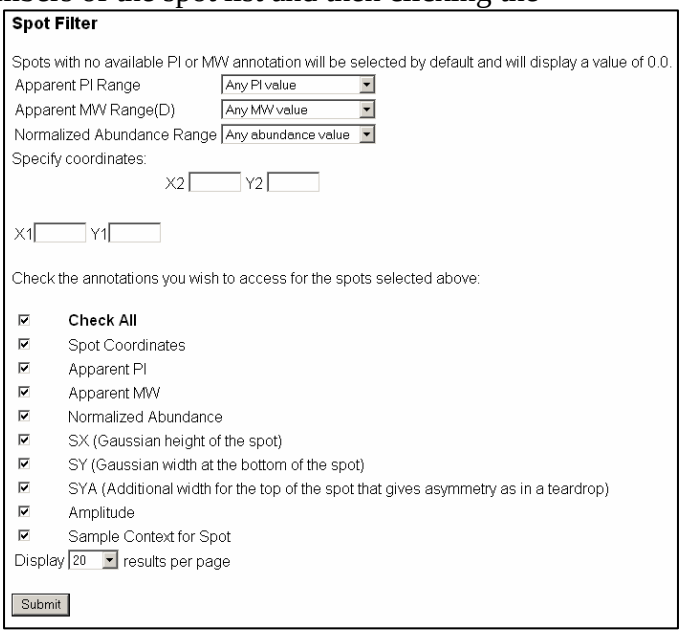
View Images:
Select nuclear fraction, TED-007, and then select "View Image List" (bottom of page). This returns a list of gel images associated with the Study. Select one at random; download and open on your workstation to see an image of the original 2D gel.

View Spot Lists and peptide data:
Return to the list of experiments using the back button and re-select experiment TED-007. Select the "View Spot Lists". This brings up a list of spots that can be analyzed further.

The Excel, XML or CSV export functions at the bottom of the page create a copy of the entire list which can be saved to the user's desktop.

Alternatively, selecting one or more of the members of the spot list and then clicking the Submit button takes the user to a page from which the spot data can be filtered further and the peptide data viewed.

An image of this page is provided at the right. Use the filtering functions on this page to select parameters of interest, and then click the Submit button to retrieve the spots that match the filter criteria, and return the desired data.



After the Submit button is clicked CEBS displays the following type of information for spots meeting the filtering criteria:

All:	Ref Spot #	Ref Gel	Gel Name	Spot #	Sample	Stressor	Dose	Treatment Time	Normalized Abundance	PI	MW	X	Y	SX	SY	SYA	Amplitude
☐	5	F344LVN001	IM8N573J_F 1	AMAP 1500mg 6h:24::Liver OMICS	AMAP;Oral	1500mg/kg	6 h		47756171.264	5.28	69384.01	733.510	634.674	9.375	7.718	0.922	61605.680
☐	30	F344LVN001	IM8N573J_F 3	AMAP 1500mg 6h:24::Liver OMICS	AMAP;Oral	1500mg/kg	6 h		24079510.611	5.94	54861.14	1398.850	789.122	5.533	5.796	0.413	71708.828
☐	55	F344LVN001	IM8N573J_F 4	AMAP 1500mg 6h:24::Liver OMICS	AMAP;Oral	1500mg/kg	6 h		103509046.626	6.87	53877.34	1696.304	802.921	14.712	10.120	0.148	68271.594

There are two export options on this page also. The Export functions (Excel, XML, CSV) return a file containing the all spot information and associated data seen in the display.

Alternatively, selecting one or more spot files and then the "View Protein IDs" link permits the user to access the mapping results. To do this select the spot(s) of interest and then click the "Download files" link at the bottom of the page. The files selected will be packaged into a .zip file and are available for download. After unzipping the files can be viewed to see the map results. An example is below:

PeptideMap - Map Results

Version 1.0.0
©1997-1999
ProteoMetrics

Name: gi|30580620|sp|P70196|TRA6_MOUSE TNF receptor associated factor 6 [MASS=60083]

Number of Measured Peptides : 43
Number of Matched Peptides : 6
Max. Coverage of Sequence : 16%

Measured Mass	Ave/ Mono	Computed Mass	Error (Da)	Residues Start	Residues End	Missed Cut	Peptide Sequence
882.531	M	882.467	0.064	97	104	1	SIRDAGHK
1332.540	M	1332.566	-0.026	286	297	1	PCDAASPSRGCR (2)+C2H3ON@C
1489.619	M	1489.698	-0.079	368	379	1	FGMHLKSQEEER
1796.800	M	1796.840	-0.040	172	185	1	CQVNTHIIEDCPRR (2)+C2H3ON@C
2253.100	M	2253.092	0.008	349	367	1	VAEMEAQQCNGIYIWKIGK (1)+C2H3ON@C (1)+O@M
2314.179	M	2314.193	-0.014	478	497	1	GFGYVTFMHLEALRQGTFIK

Unmatched Average Masses:
3347.580

Unmatched Monoisotopic Masses:
698.611 759.492 773.517 797.567 814.572 841.612 855.591 859.538 867.594 869.614
893.564 898.541 962.494 980.489 1025.579 1125.500 1154.400 1168.520 1189.489
1208.530 1262.319 1277.619 1300.599 1316.540 1327.660 1345.569 1611.739 1633.670
1647.729 1790.770 2002.020 2131.060 2224.100 2238.100 2245.159 2449.669

If you downloaded a file with search software = "MASCOT" the file will come in a .dat format which can be viewed using a text editor. Files with search software = "PROFOUND" come in .html format, similar to the example above.

Returning to the "Protein Spot Data" page (two back clicks), it is possible to retrieve the spectra associated with a given mass spec run, although specialized software is required to open the spectral files. The spectra can be downloaded from the "Protein Spot Data" page via the "View Spectra List MS" or "View Spectra List MS/MS" links at the bottom of the page.

Viewing proteins with altered levels in control and treated animals:

At this point in the analysis one can choose to view the expression data from one or more spots in relation to the control data. Alternatively one can navigate back to the original page ("CEBS Proteomics browser") and select this function for the entire experiment. This can most easily be done by returning to the CEBS home page using the menu at the left or the blue CEBS home link in the workflow summary at the top of the page (directly under the header CHEMICAL EFFECTS IN BIOLOGICAL SYSTEMS).

After selecting "All Proteomics" from the sidebar menu, re-select the TED-007 study. Identify which of the groups to investigate; check all the stressors and times, and then select "View Treatment to Control Expression Data". This returns a page where thresholds can be set, and the query restricted to identified proteins or left at all proteins.

Leaving the search unbounded returns 20,950 items.

All: ☐	Ref Spot #	RefGel	Average PI	PI(+/- SD)	Average MW	MW(+/- SD)	Treatment Group	Treatment Time	Comparator Group	Experimental Group Average Abundance	Ratio of Experimental to Average Control	Expression Fold Change	p-value
☐	1	F344LVN001	5.24	0.013	117316.66	2830.643	AMAP_HD_24H	24HR	CONTROL_24H	17594636.00	1.569	1.569	n/a
☐	1	F344LVN001	5.24	0.013	117316.66	2830.643	AMAP_HD_48H	48HR	CONTROL_48H	7500759.00	0.637	-1.569	n/a
☐	1	F344LVN001	5.24	0.013	117316.66	2830.643	AMAP_HD_6H	6HR	CONTROL_6H	8375249.00	0.700	-1.429	n/a
☐	1	F344LVN001	5.24	0.013	117316.66	2830.643	AMAP_LD_24H	24HR	CONTROL_24H	13128371.00	1.171	1.171	n/a
☐	1	F344LVN001	5.24	0.013	117316.66	2830.643	AMAP_LD_48H	48HR	CONTROL_48H	11180492.00	0.950	-1.052	n/a
☐	1	F344LVN001	5.24	0.013	117316.66	2830.643	AMAP_LD_6H	6HR	CONTROL_6H	7212941.00	0.603	-1.659	n/a
☐	1	F344LVN001	5.24	0.013	117316.66	2830.643	APAP_HD_24H	24HR	CONTROL_24H	1641957.10	0.146	-6.829	n/a

Limiting the query to proteins with IDs ("Expression Analysis Data (Confirmed Protein ID only)") returns 5163 items.

All: ☐	Ref Spot #	RefGel	Protein	Accession	Confidence Score	% Coverage	Average PI	Average MW	Treatment Group	Experimental Group Average Abundance	Ratio of Experimental to Average Control	Expression Fold Change	p-value	Links to Annotations
☐	1524	F344LVN001	NUCLEOPHOSMIN (NPM) (NUCLEOLAR PHOSPHOPROTEIN B23) (NUMATRIN) (NUCLEOLAR PROTEIN NO38)	sp|O61937	39.0	n/a	n/a	n/a	APAP_HD_24H	415.06	0.001	-1957.487	n/a	GO BioCarta KEGG
☐	2021	F344LVN001	KERATIN, TYPE II CYTOSKELETAL 8 (CYTOKERATIN 8) (CYTOKERATIN ENDO A)	sp|O10758	41.0	n/a	5.86	36388.66	AMAP_HD_6H	1217.11	0.004	-235.479	n/a	GO BioCarta KEGG

Note the links to KEGG, BioCarta and GO at the far right; unfortunately the proteins with the greatest change in this Study are not in BioCarta or KEGG at this time.