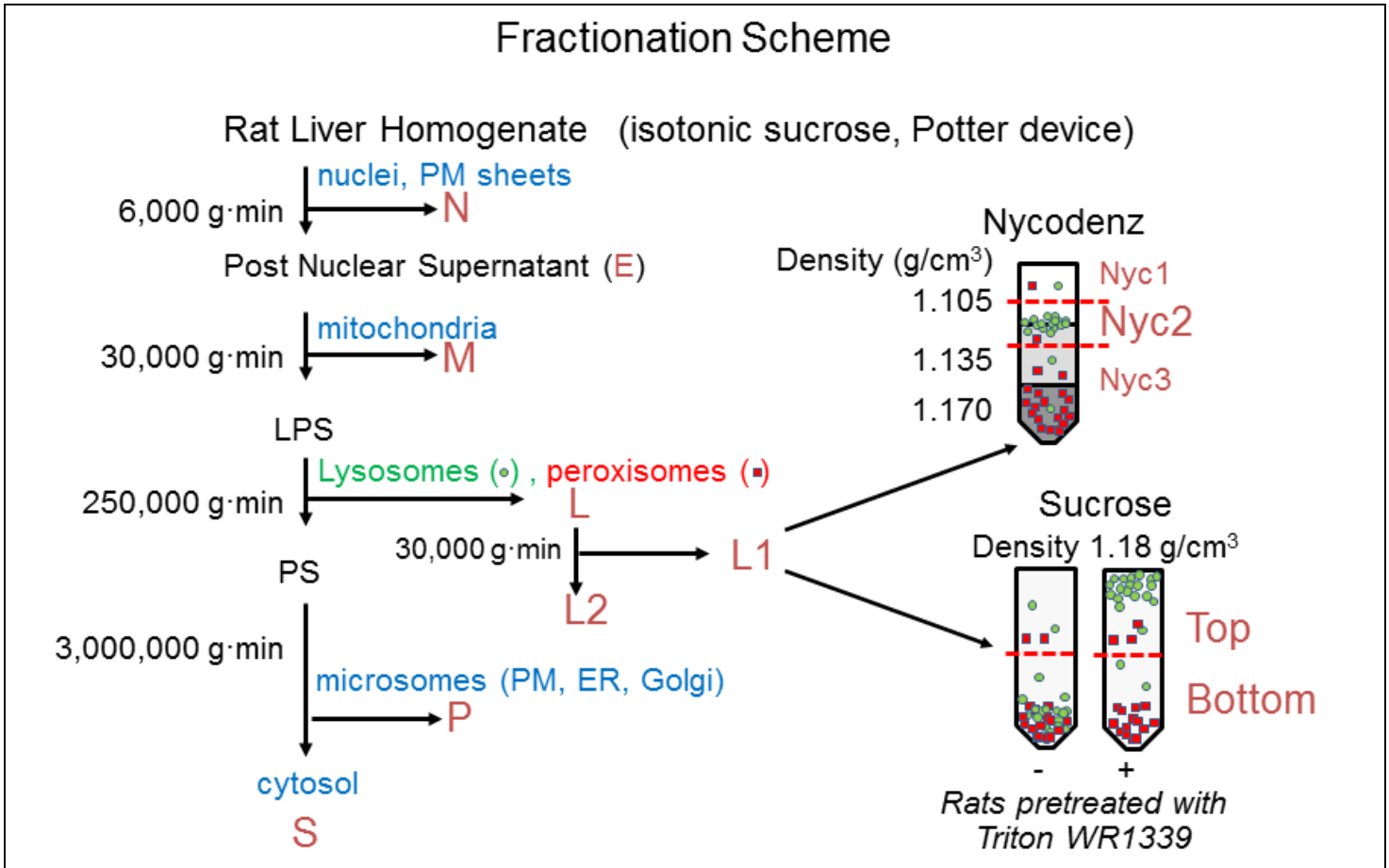
PROLOCATE contains information on protein subcellular location based on four analytical subcellular fractionation/quantitative mass spectrometry experiments conducted on rat liver.

Experimental design. In subcellular fractionation, different organelles are enriched or depleted in various centrifugation fractions. When analyzing an individual protein, its pattern of enrichment or depletion in different subcellular fractions can be used to determine what organelle(s) the protein is associated with. Balance sheet analysis provides confidence that the entire distribution of a protein is accounted for, and also allows estimates of the proportional distribution of proteins that are present in multiple compartments.

The fractionation schemes used in this study are shown below. The left panel shows the differential centrifugation scheme. Rat liver lysosomes and peroxisomes have very similar distributions by differential centrifugation. These can be resolved from each other based on their densities in Nycodenz (top right panel, fractionation scheme) or in sucrose when animals are administered Triton WR1339 to induce a lysosomal lipidosis (bottom right panel). Expts A and C are conducted on overnight fasted rats while Expts B and D are conducted on animals allowed to feed *ad libitum*. Indicated fractions were analyzed by mass spectrometry after isobaric labeling and extensive peptide prefractionation.



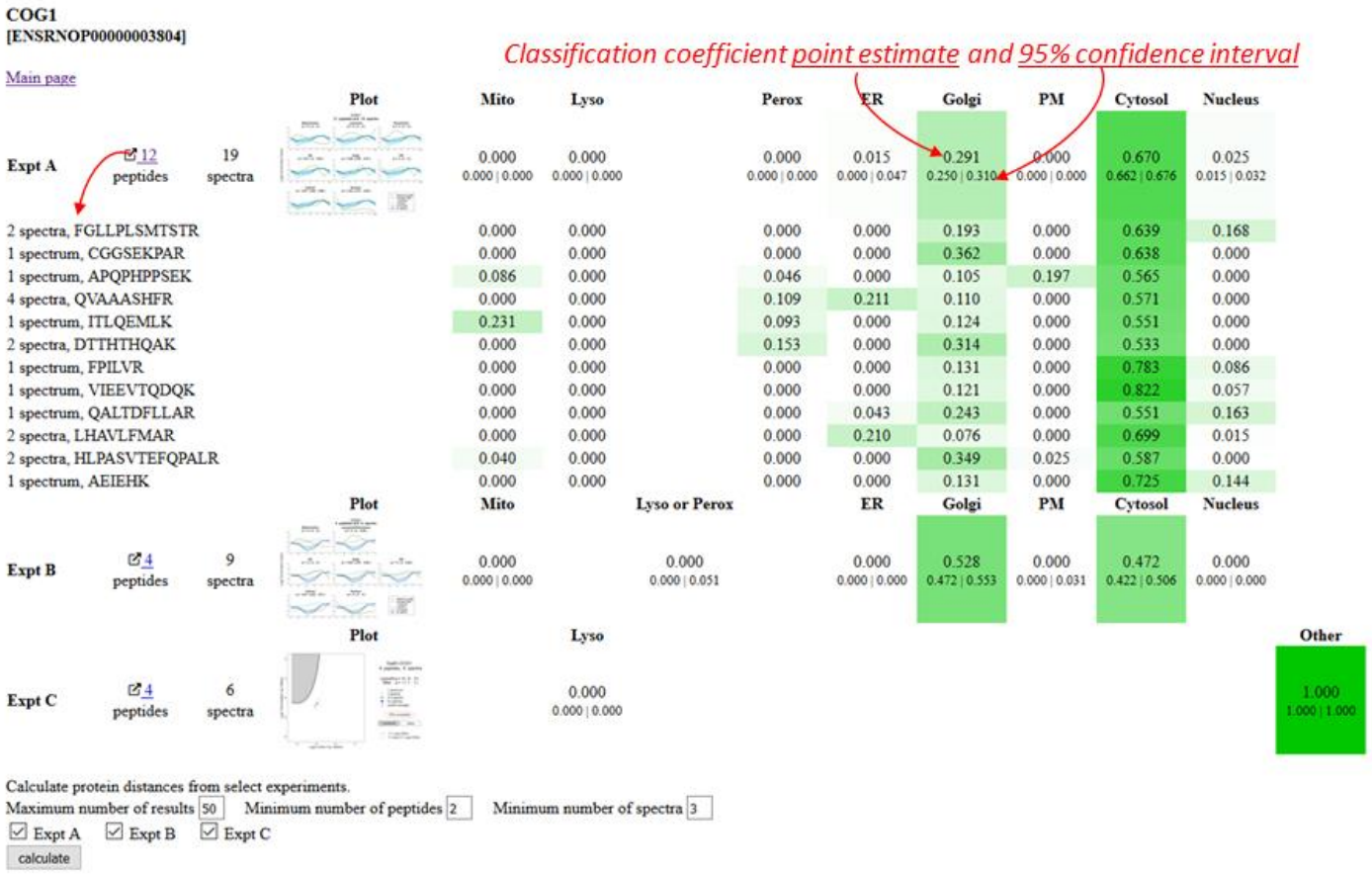
Experiments used to create the Prolocate database.						
Expt	Diet	Triton treatment	L1 fractionation	Fractions analyzed by MS	# unique proteins quantified ≥2 peptide & ≥3 spectra (all)	
A	Fast	-	Nycodenz	H, N, M, L1,L2, P, S, Nyc2	6073 (8071)	A∩B 2952 (4437) A∩C 4761 (6361) B∩D 1923 (3160) C∩D 2211 (3710) A∩B∩C 2852 (4200) A∪B∪C∪D 6920 (9871)
B	Fed	-	none	H, N, M, L1, L2, P, S	3021 (4707)	
C	Fast	+ and -	Sucrose	L1, Top & Bottom sucrose	5553 (7859)	
D	Fed	+ and -	Sucrose	L1, Top & Bottom sucrose	2245 (3968)	

* Isobaric-labeled tryptic peptides were prefractionated by SCX and/or alkaline reverse phase HPLC prior to LC-MS/MS using an Orbitrap Velos. Over 250 LC-MS runs were conducted per experiment. After database searching and reporter ion quantification, spectra were included if they mapped to fully tryptic peptides with no missed cleavages and no adventitious modifications except monooxidation at methionine, 100% labeling efficiency, and acceptable (2/3 to 3/2) recovery (∑fractions/starting material), yielding 82,662 unique peptides.

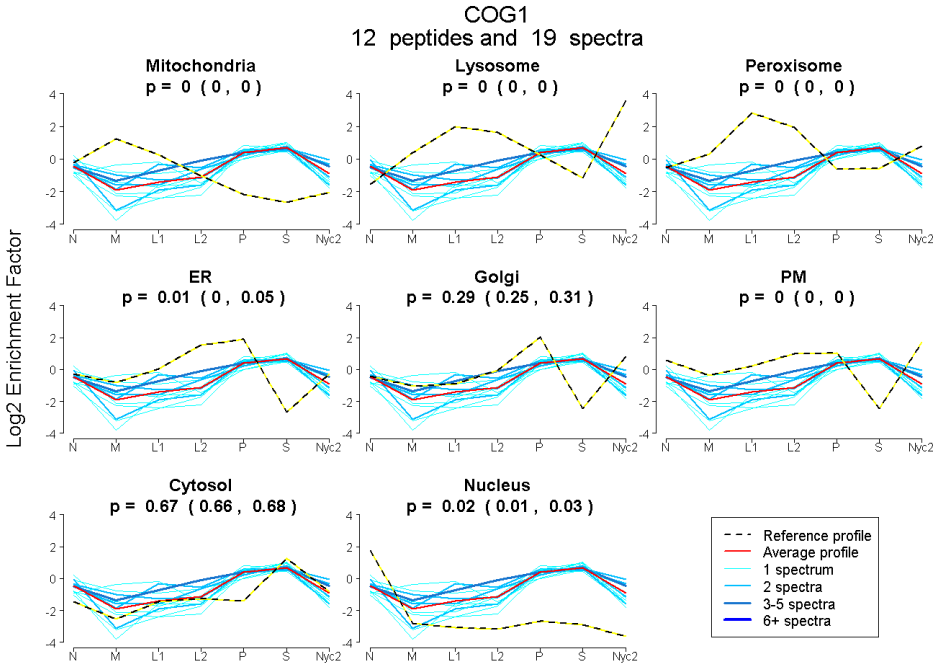
PROTEIN-CENTRIC INFORMATION.

Find your protein of interest by looking up its name or accession number or by performing a BLAST search from the main page. Click to retrieve a summary of all experiments for that protein. The sum of the classification coefficients in a given experiment is one, and each provides an estimate of proportional residence in the indicated compartment or, for single compartmental proteins, the confidence of organelle assignment. Classification coefficients for the individual peptides assigned to the protein can be seen by clicking on the link.

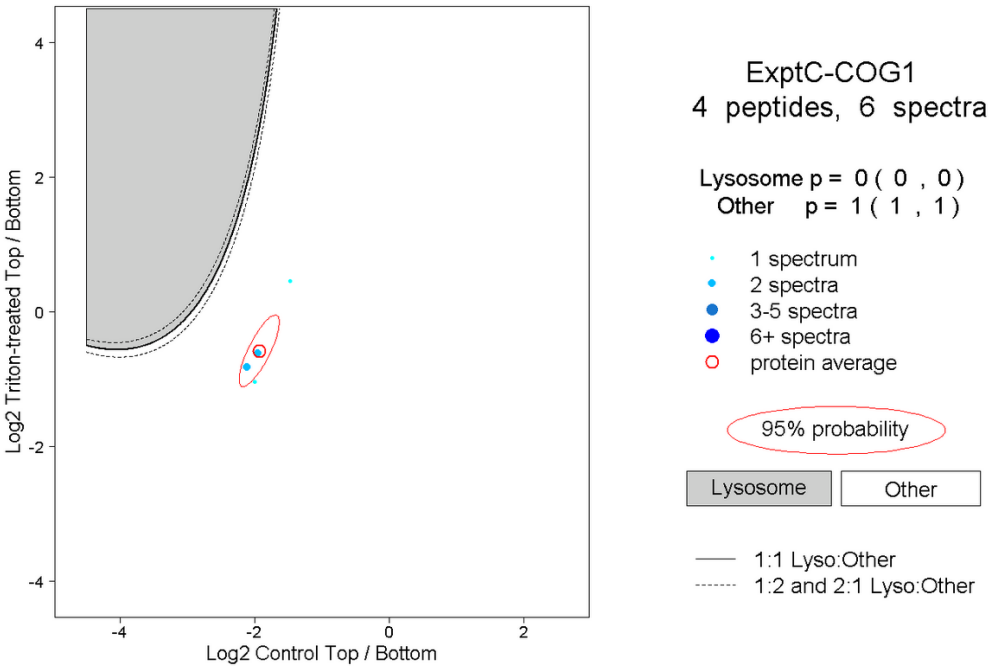
Note that for Expts C and D, the density-based separation is conducted on a fraction (L1) pre-enriched in lysosomes and peroxisomes. Thus, for multi-compartmental proteins with lysosomal and non-lysosomal, non-peroxisomal locations, Expts C and D will overestimate the lysosomal proportion of the total population.



For Expts A and B, the plots on the protein page show an average profile for each protein (red line) and its composite peptides (blue lines). Profiles are fit to a linear combination of reference organelle marker protein profiles (dashed black and yellow line). Note that for Expt. B, lysosomes and peroxisomes cannot be distinguished from the fractions analyzed and thus there is a combined Lyso or Perox classifier.



For Expts C and D, plots show the average distribution of a protein (red circle, point estimate; red ellipse, 95% probability) and component peptides (blue circles) in the sucrose density centrifugation analysis. Marker proteins are used to define regions of the graph where lysosomal proteins (gray area) and other types of proteins (no shading) distribute. This information is used to calculate classification coefficients and confidence limits for residence in the lysosome or other compartment.



Distance measurements. The distance calculator can be used to find proteins based on similarity in fractionation behavior. This is particularly useful for identifying candidate components of protein complexes that are associated with multiple organelles. Increasing the number of minimum peptides and spectra will decrease the number of proteins analyzed but may provide more accurate information.

COG1
[ENSRNOP00000003804]

Main page

...

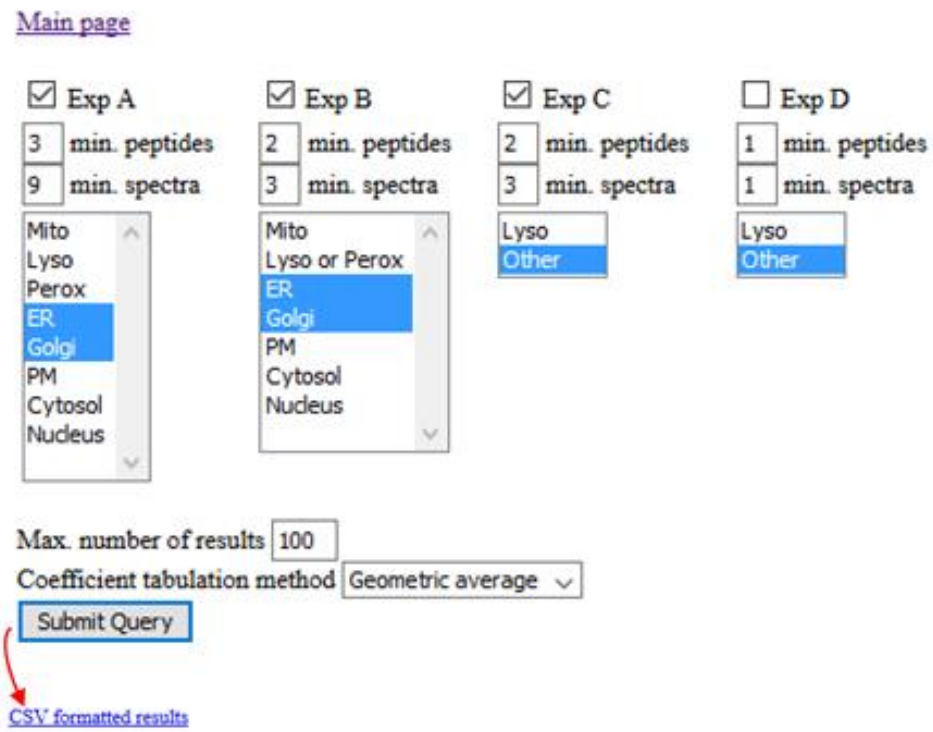
Calculate protein distances from select experiments.
Maximum number of results 10 Minimum number of peptides 2 Minimum number of spectra 3
☑ Expt A ☑ Expt B ☑ Expt C
calculate

CSV formatted results

Rank	Gene	Distance	Exp:peptides,spectra		
1	COG5	0.969	A:13,22	B:4,6	C:3,5
2	COG7	1.663	A:11,16	B:3,5	C:6,13
3	COG8	1.794	A:12,20	B:4,5	C:2,4
4	UBE4A	2.020	A:16,36	B:5,8	C:6,12
5	COG4	2.216	A:9,25	B:3,3	C:12,19
6	ARFGEF2	2.381	A:40,84	B:4,7	C:27,52
7	NPLOC4	2.435	A:15,32	B:4,9	C:14,30
8	GCC2	2.559	A:17,27	B:4,5	C:8,11
9	STRAP	2.626	A:11,31	B:4,9	C:8,17
10	UFD1L	2.671	A:10,17	B:2,4	C:5,13

COMPARTMENT-CENTRIC INFORMATION.

To search for candidate compartment proteins, click on the coefficient search button on the main page. Specify what experiments to include, inclusion criteria (minimum required number of peptides and spectra), and coefficients. Note that if multiple coefficients are specified in a given experiment, the calculator uses the sum of those coefficients.



CSV formatted results

Rank	Gene id	Geometric average	Point estimate, # peptides, # spectra		
			Exp A	Exp B	Exp C
1	UGT2B	1.000	ER 1.000 + Golgi 0.000=1.000, 17, 391	ER 1.000 + Golgi 0.000=1.000, 17, 308	Other 1.000, 18, 842
2	UGT2B10	1.000	ER 1.000 + Golgi 0.000=1.000, 8, 63	ER 1.000 + Golgi 0.000=1.000, 6, 27	Other 1.000, 8, 93
3	UGT2B37	1.000	ER 1.000 + Golgi 0.000=1.000, 14, 197	ER 1.000 + Golgi 0.000=1.000, 12, 133	Other 1.000, 13, 306
4	DNAJC25	0.998	ER 0.995 + Golgi 0.000=0.995, 8, 17	ER 1.000 + Golgi 0.000=1.000, 2, 3	Other 1.000, 8, 19
5	OS9	0.998	ER 0.831 + Golgi 0.162=0.993, 12, 46	ER 0.827 + Golgi 0.173=1.000, 3, 7	Other 1.000, 10, 35
6	ENSRNOG00000020817	0.997	ER 0.806 + Golgi 0.185=0.991, 8, 42	ER 0.877 + Golgi 0.123=1.000, 8, 43	Other 1.000, 9, 67
7	ENSRNOG00000045652	0.997	ER 0.999 + Golgi 0.000=0.999, 3, 72	ER 0.993 + Golgi 0.000=0.993, 3, 21	Other 1.000, 3, 219
8	PIGN	0.996	ER 0.891 + Golgi 0.096=0.987, 8, 26	ER 0.860 + Golgi 0.140=1.000, 6, 9	Other 1.000, 9, 41
9	PON2	0.996	ER 0.989 + Golgi 0.000=0.989, 7, 49	ER 1.000 + Golgi 0.000=1.000, 6, 13	Other 1.000, 9, 40
10	PTPLAD1	0.995	ER 0.877 + Golgi 0.120=0.997, 10, 43	ER 0.989 + Golgi 0.000=0.989, 4, 15	Other 1.000, 10, 52
11	SEC61B	0.995	ER 1.000 + Golgi 0.000=1.000, 4, 33	ER 0.984 + Golgi 0.000=0.984, 4, 14	Other 1.000, 2, 23
12	CANX	0.994	ER 0.989 + Golgi 0.000=0.989, 21, 273	ER 0.916 + Golgi 0.078=0.994, 16, 108	Other 1.000, 22, 373
13	CYP2C11	0.994	ER 0.657 + Golgi 0.325=0.982, 25, 606	ER 0.836 + Golgi 0.164=1.000, 22, 319	Other 1.000, 25, 715
14	CYP4A2	0.993	ER 0.904 + Golgi 0.076=0.980, 26, 422	ER 0.884 + Golgi 0.116=1.000, 20, 166	Other 1.000, 25, 519
15	DNAJC3	0.993	ER 0.980 + Golgi 0.000=0.980, 23, 129	ER 1.000 + Golgi 0.000=1.000, 18, 66	Other 1.000, 23, 265