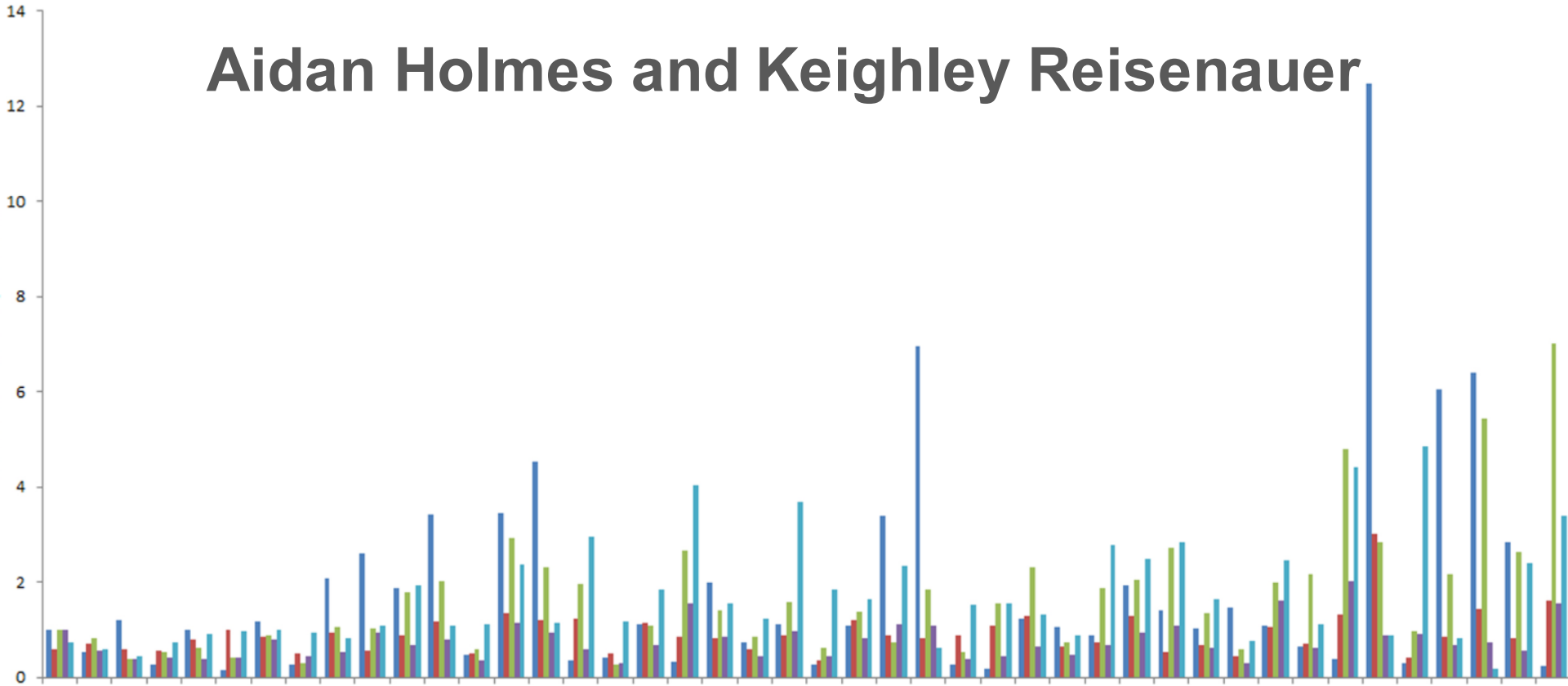
Quantitative Proteomics

Aidan Holmes and Keighley Reisenauer



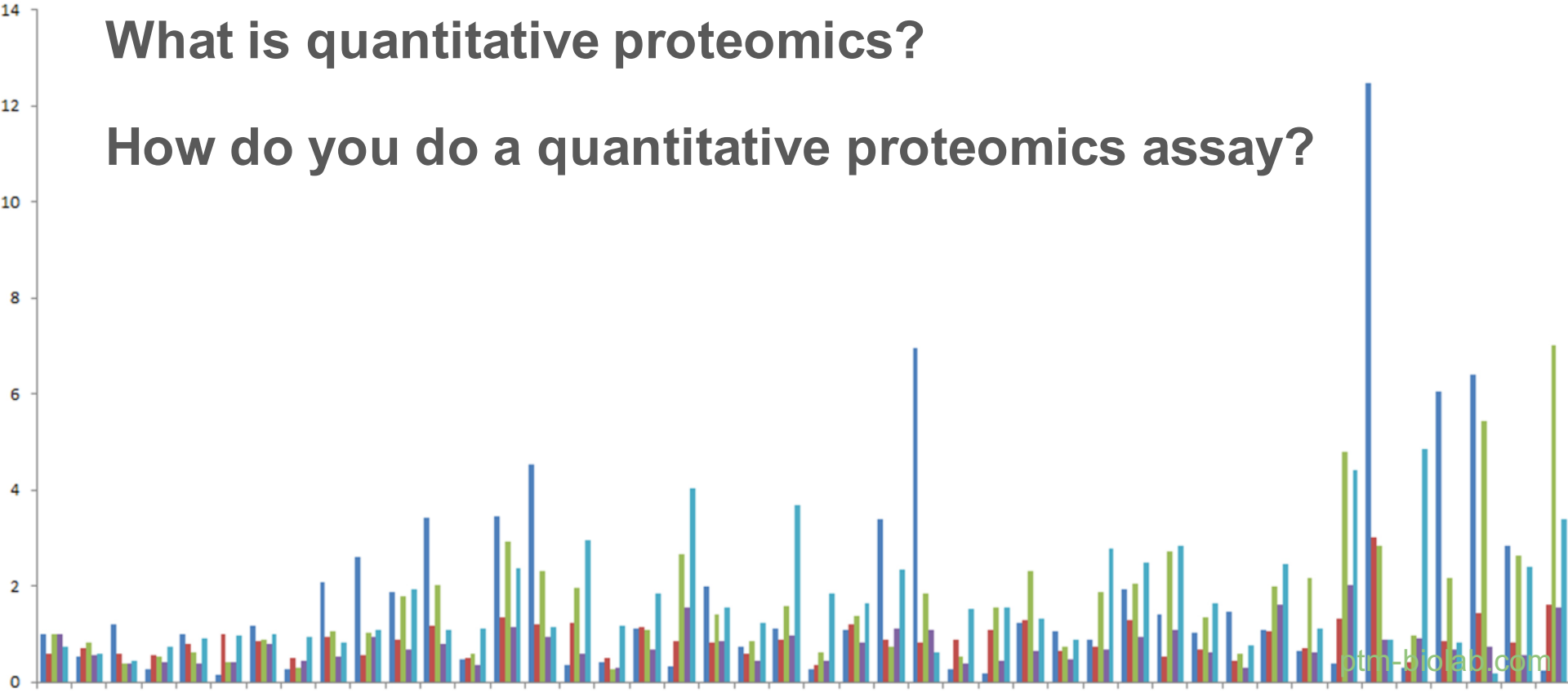
Overview

What is proteomics?

What is mass spectrometry?

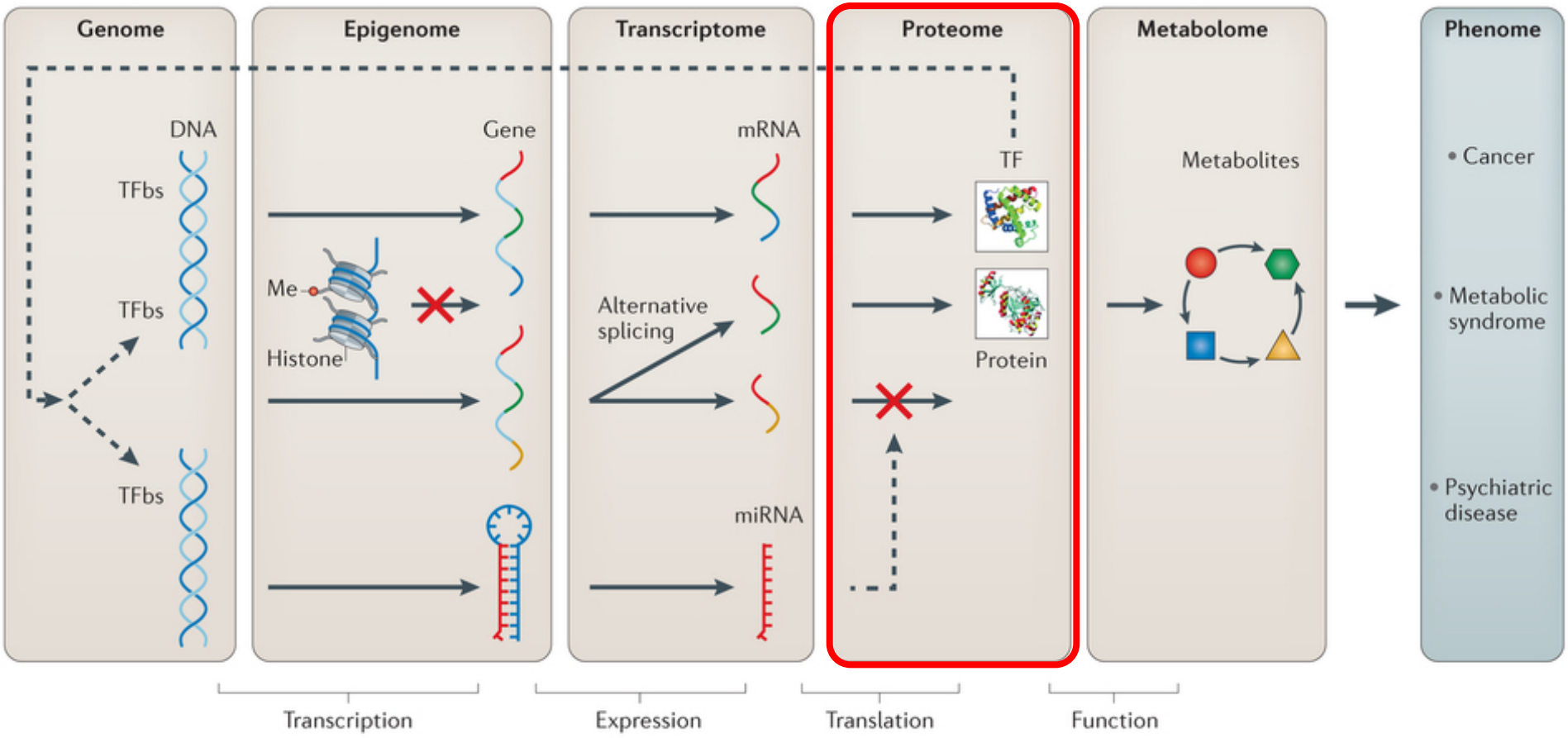
What is quantitative proteomics?

How do you do a quantitative proteomics assay?




JEOPARDY!

Why is proteomics useful?



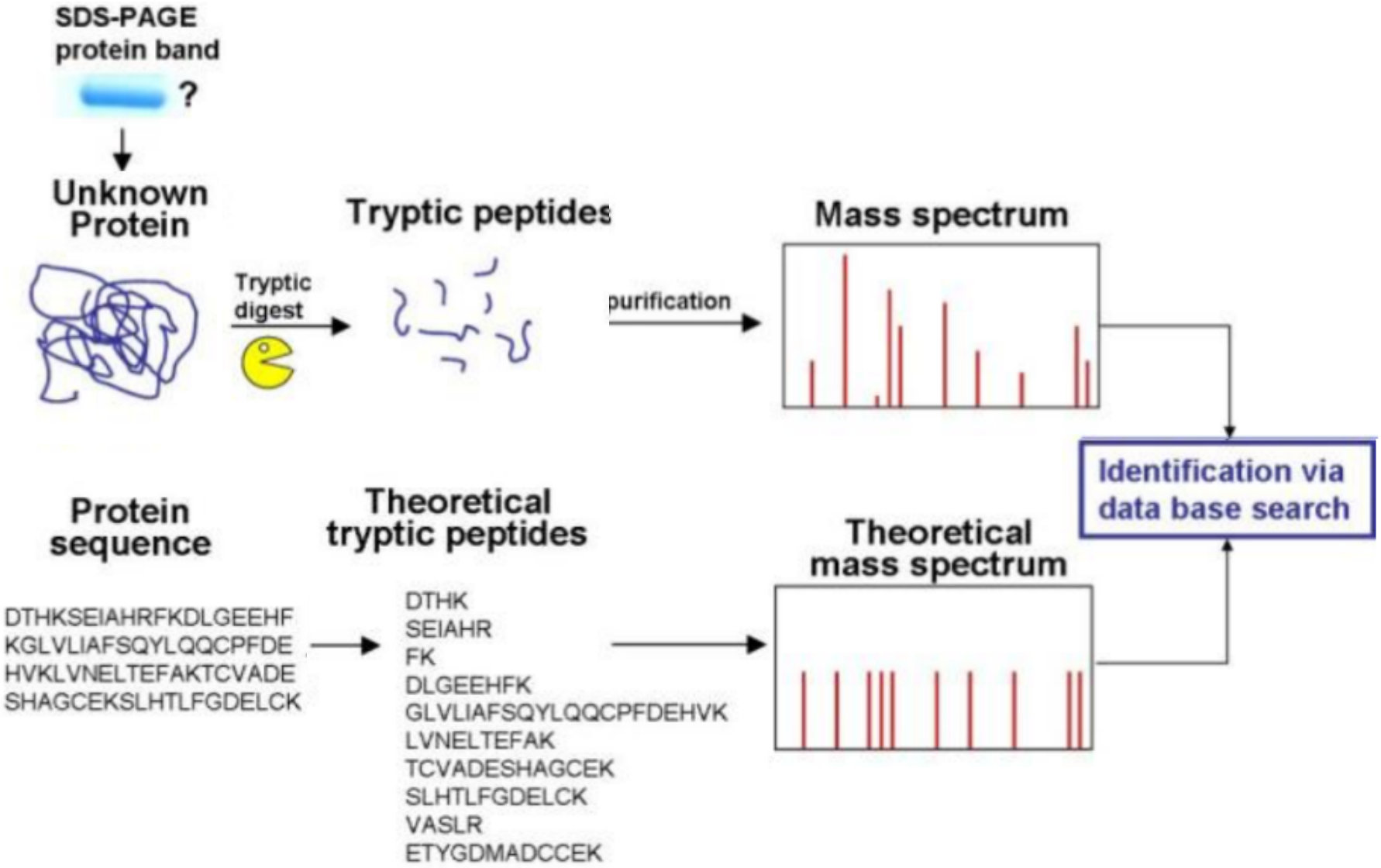
http://www.nature.com/nrg/journal/v16/n2/images/nrg3868-f1.jpg

How do we identify proteins?

LIKE THIS!

But, actually?

Refresh: What is mass spectrometry?



What is important to know about mass spec?

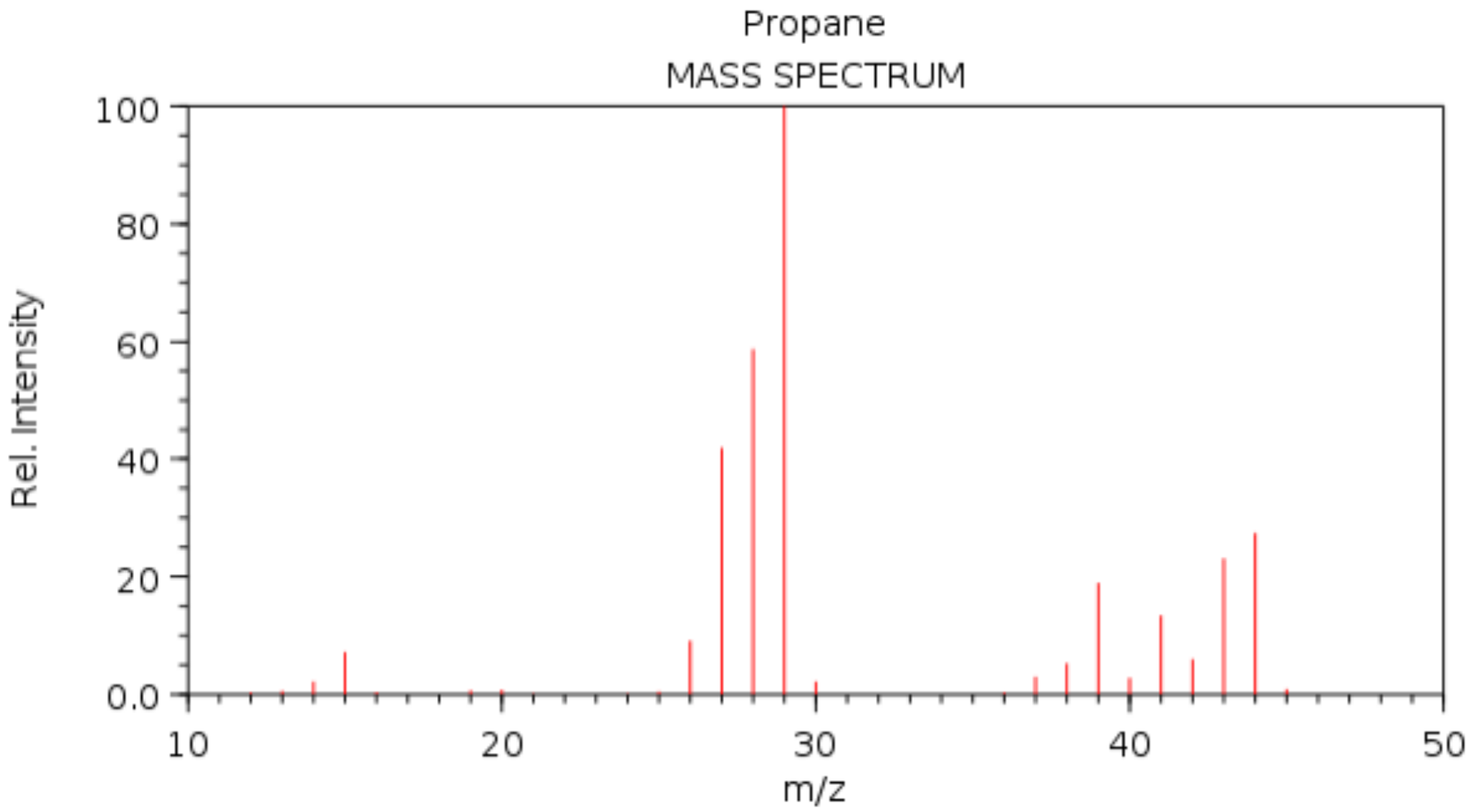
- m/*z* ratio

- Sequence coverage

protein set	Protein Header	Mass	Coverage	Pep Count	Different Peps	Occurrence	Rank
15,185	mousegp:MCHR9-086280|--|SWTR_MUS:APOA1_MOUSE: (Apo-AI) (ApoA-I).	30.6	8%	2	2	1	601
13,196	mousegp:MCHR7-071588|--|SWTR_MUS:APOE_MOUSE: Apolipoprotein E precursor (Apo-E).	35.8	13.20%	3	3	1	479
12,189	mousegp:MCHR7-063783|--|SWTR_MUS:LEG4_MOUSE: Galectin-4	36.3	34.70%	28	10	4	104
7	mousegp:MCHR2-061149|--|SWTR_MUS:S27A4_MOUSE: Long-chain fatty acid transport protein 4 (EC 6.2.1.-) (Fatty acid transport protein 4) (FATP-4)	72.3	2.60%	2	2	1	558

http://www.google.com/patents/WO2008138495A1?cl=en

What is a m/z ratio?

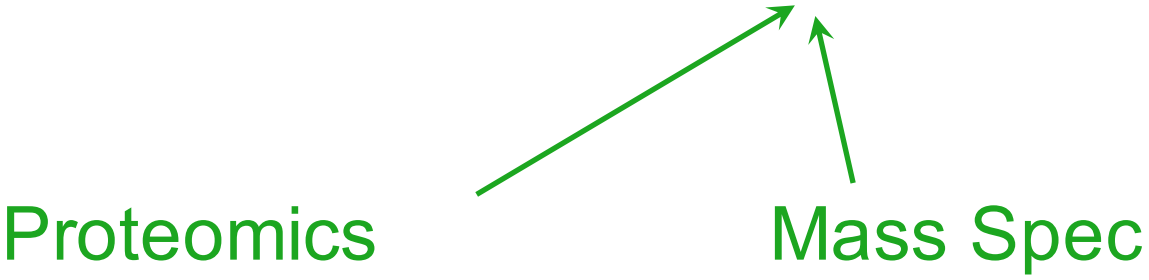


http://webbook.nist.gov/cgi/cbook.cgi?ID=C74986&Mask=200
https://commons.wikimedia.org/wiki/File:Propane-3D-balls-B.png

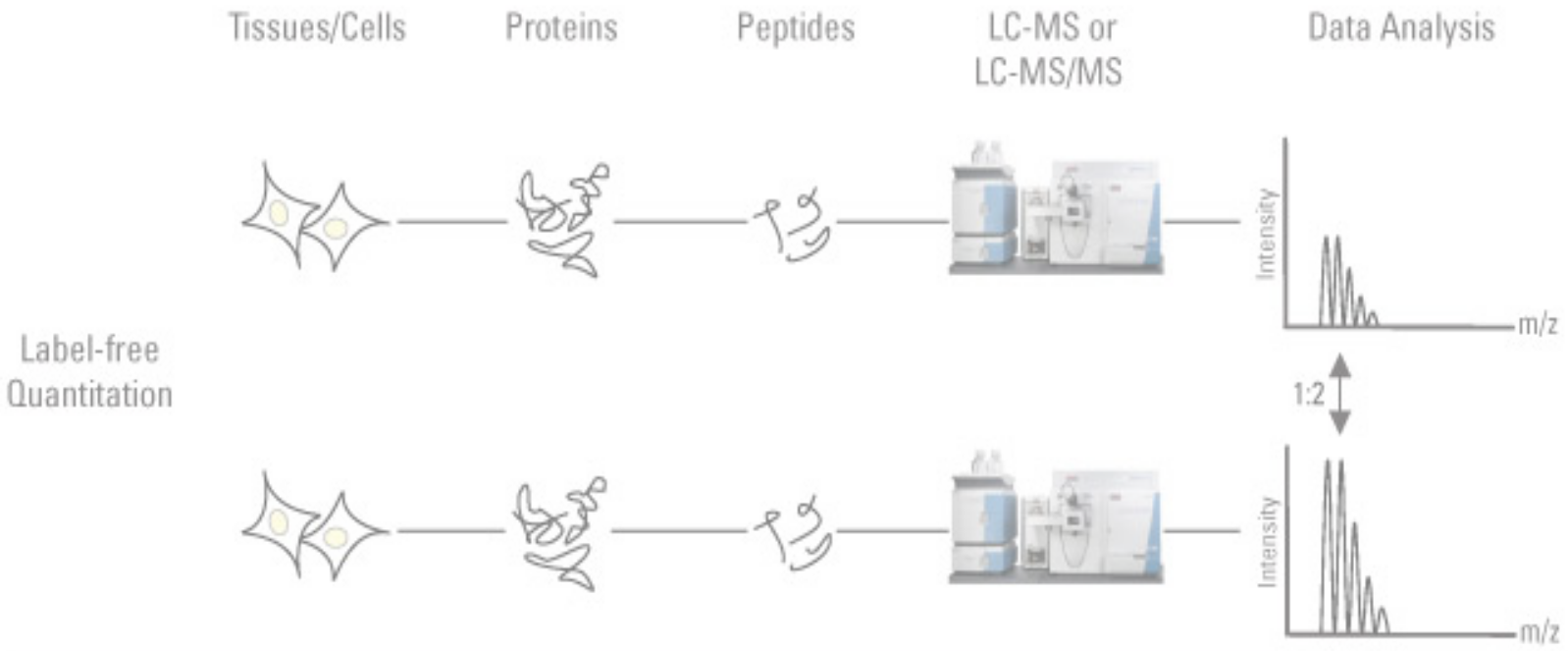
What is sequence coverage?

MWSHPQFEKISGANGAMAKLTSAVPVLTARDVAGAVEFWTDRLGFS
RDFVEDDFAGVVRDDVTLFISAVQDQVVPDNTLAWVWVRGLDELYAE
WSEVVSTNFRDASGPAMTEIGEQPWGREFALRDPAGNCVHFVAEEQ
DENLYFQGGSHHHHHHHHHHHHAGVPAPASQLTKVLAGLRHTFVVA
DATLPDCPLVYASEGFYAMTGYGPDEVLGHNCRFLQGEGTDPKEVQ
KIRDAIKKGEACSVRLLNYRKDGTPFWNLLTVTPIKTPDGRVSKFVGV
QVDVTSKTEGKALADNSGVPLLVKYDHRLRDNVARTIVDDVTIAVEKA
EGVEPGQASAVAAAAPLGAKGPRGTAPKSFPRVALDLATTVERIQQN
FCISDPTLPDCPIVFASDAFLELTGYSREEVLGRNCRFLQGAGTDRGT
VDQIRAAIKEGSELTVRILNYTKAGKAFWNMFTLAPMRDQDGHARFFV
GVQVDVTAQSTSPDKAPVWNKTPEEEVAKAKMGAEAASLISSALQG
MAAPTTANPWAAISGVIMRRKPHKADDKAYQALLQLQERDGKMKLM
HFRRVKQLGAGDVGLVDLVQLQGSELKFAMKTLDKFEMQERNKVAR
VLTESAILAAVDHPFLATLYCTIQTDTHLHFVMEYCDGGELYGLLNSQ
PKKRLKEEHVRFYASEVLTALQYLHLLGYVYRDLKPENILLHHTGHVL
LTDFDLSYSKGSTTPRIEKIGGAGAAGGSAPKSPKKSSSKSGGSSSG
SALQLENYLLLAEPSARANSFVGTEEYLAPEVINAAGHGPAVDWWSL
GILIFELLYGTTPFRGARRDETFENIIKSPLKFPSKPAVSEECRDLIEKL
LVKDVGARLGSRTGANEIKSHPWFKGINWALLRHQQPPYVPRRASK
AAGGSSTGGAAFDNY

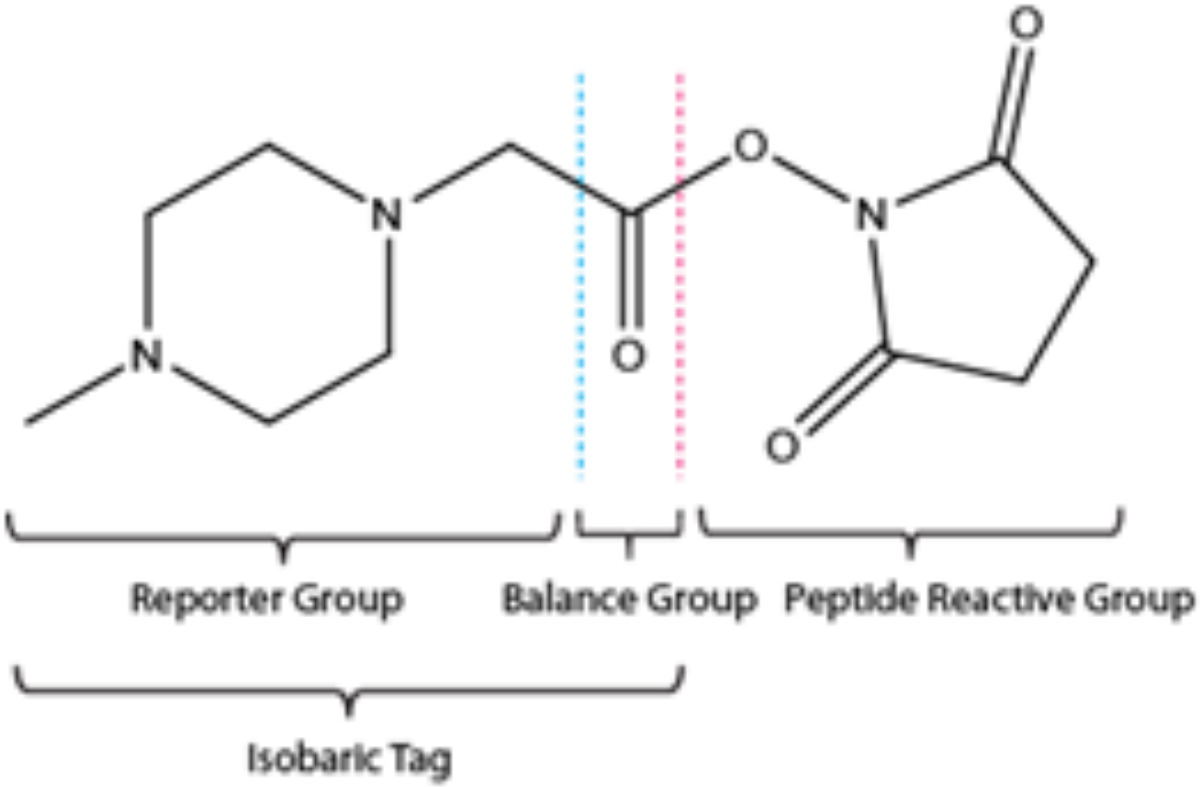
What is quantitative proteomics?



What types of quantitative proteomics are there?



What is iTRAQ?

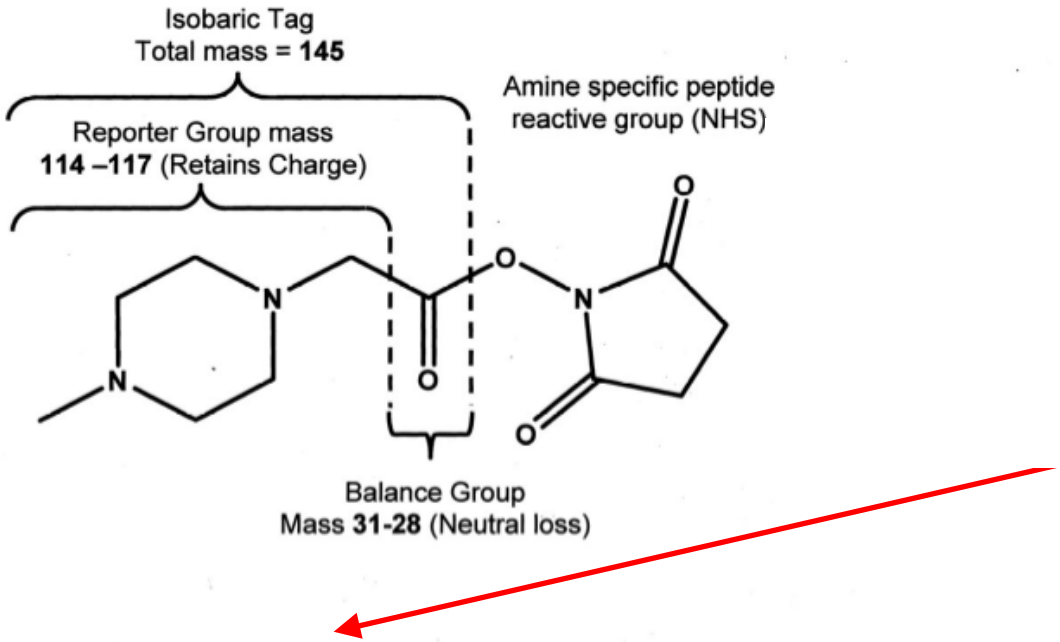


iTRAQ

Isobaric **T**ags for **R**elative and **A**bsolute **Q**uantification

https://www.broadinstitute.org/files/shared/proteomics/images/

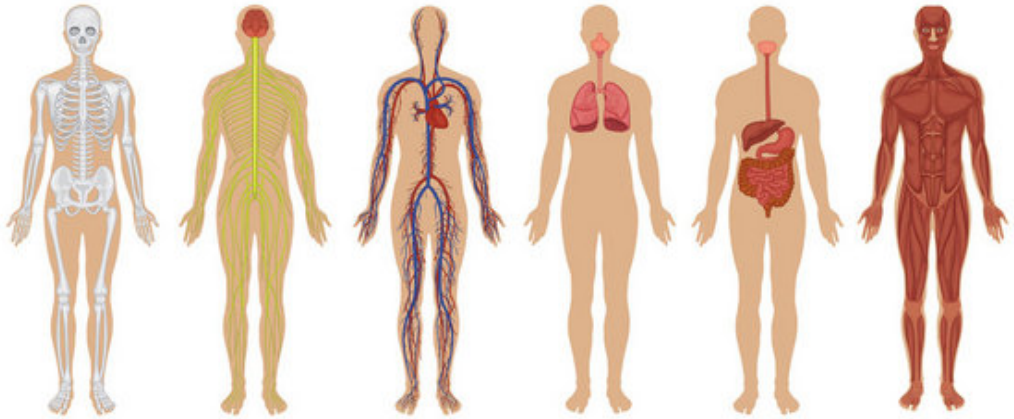
How does iTRAQ work?

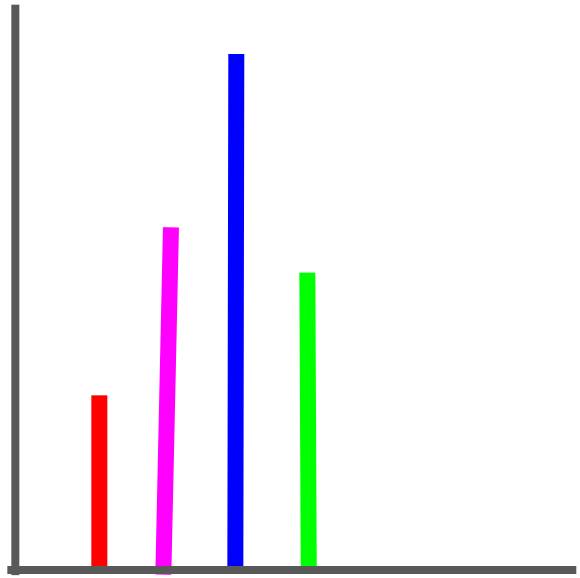


Ross et al., Mol Cell Prot (2004), 3, 1154-1169.

What is an example iTRAQ workflow?

Is iTRAQ a good choice?





http://www.hindawi.com/journals/ijpro/2013/581862/
https://www.broadinstitute.org/scientific-community/science/platforms/proteomics/itraq
http://www.livescience.com/37009-human-body.html
http://www.rrdonnelley.com/retail/solutions/

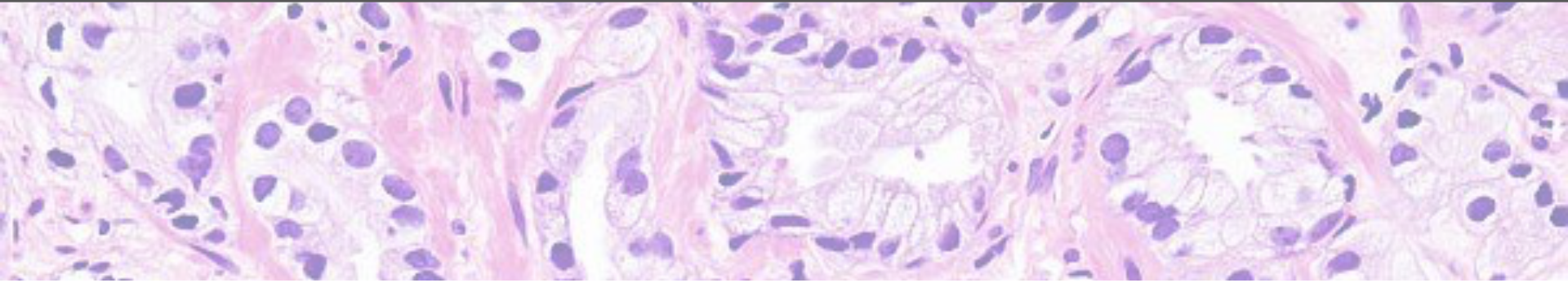
How could you use quantitative proteomics in your project?

Conclusions/ Summary

What is proteomics? Refresher

What is quantitative proteomics?

What is iTRAQ protein labeling?

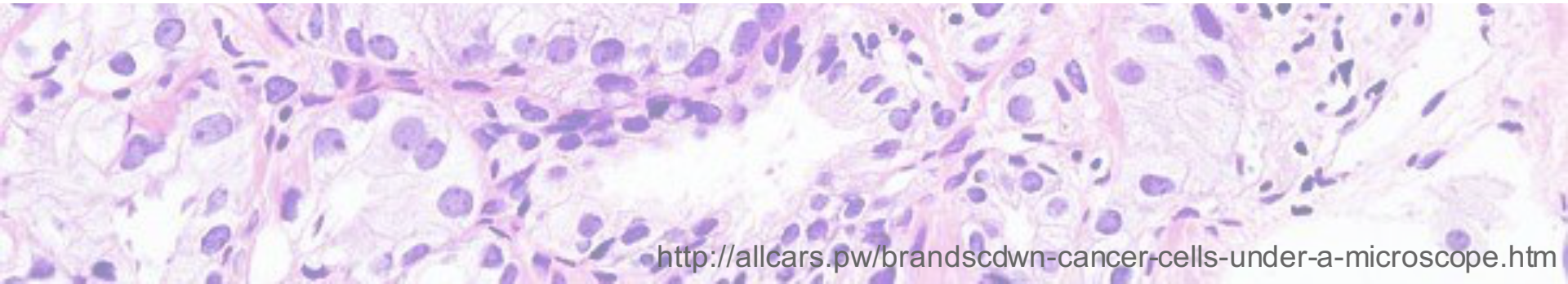


Tumor Biol. (2015) 36:9829–9837
DOI 10.1007/s13277-015-3758-7

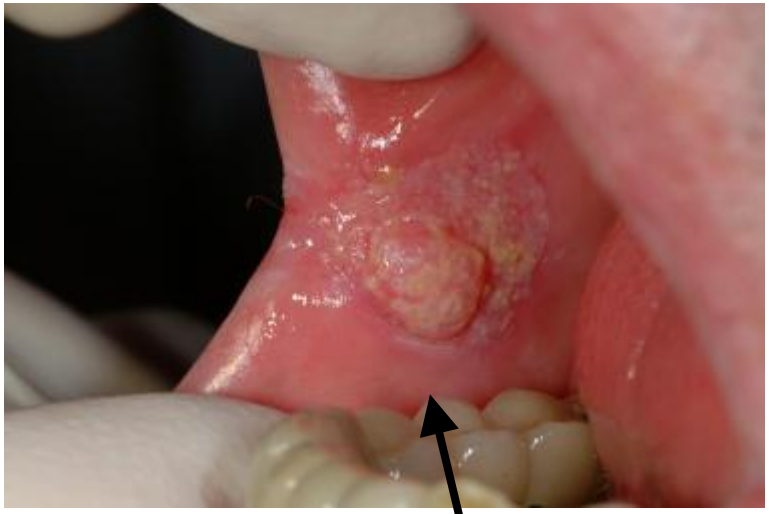
RESEARCH ARTICLE

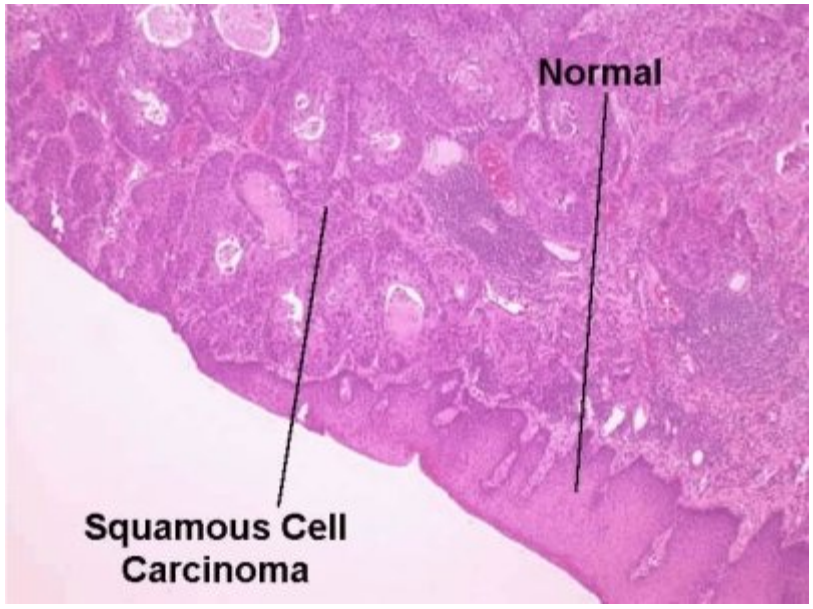
Identification of RAB2A and PRDX1 as the potential biomarkers for oral squamous cell carcinoma using mass spectrometry-based comparative proteomic approach

Kaushik Kumar Dey[1] · Ipsita Pal[1] · Rashmi Bharti[1] · Goutam Dey[1] · B. N. Prashanth Kumar[1] · Shashi Rajput[1] · Aditya Parekh[1] · Sheetal Parida[1] · Priyanka Halder[1] · Indranil Kulavi[2] · Mahitosh Mandal[1]



What is oral squamous cell carcinoma?



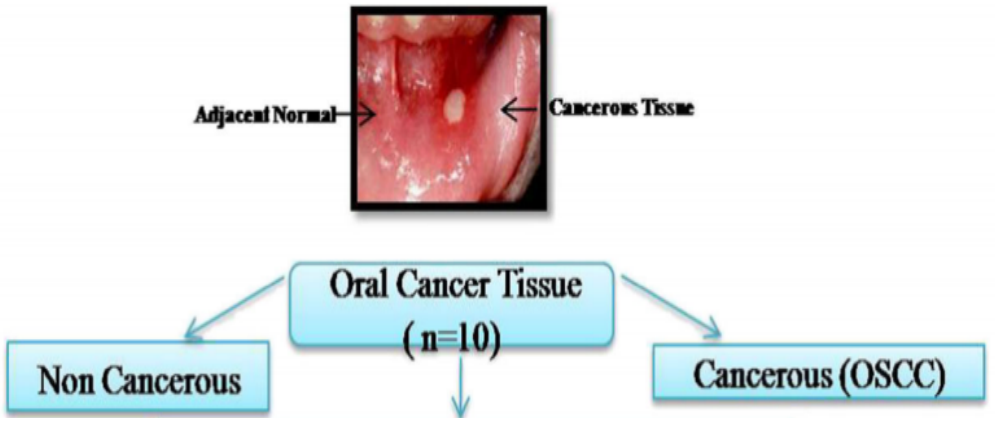


Kinda gross

Emedicine.medscape.com

http://www.dentist-charlotte-north-carolina-nc.com/images/oral-mouth-cancer-in-detail-9-squamous-cell-carcinoma-slide1.jpg

Figure 1: What is the experimental approach?

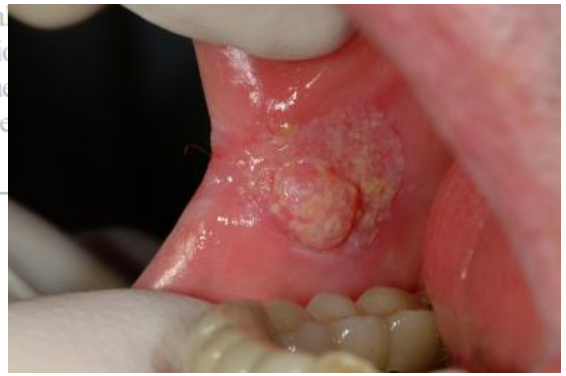


Characteristic	No. of patients (%)
Number of samples	10
Age (years)	
Mean	64.8±6.2
Range	51-73
Gender	
Male	6 (60%)
Female	4 (40.0)
Risk Factor	
Smoking	4 (40%)
Tobacco Chewing	6(60%)
Tumor Site	
Tongue	6 (60%)
Floor of the mouth	4(40%)

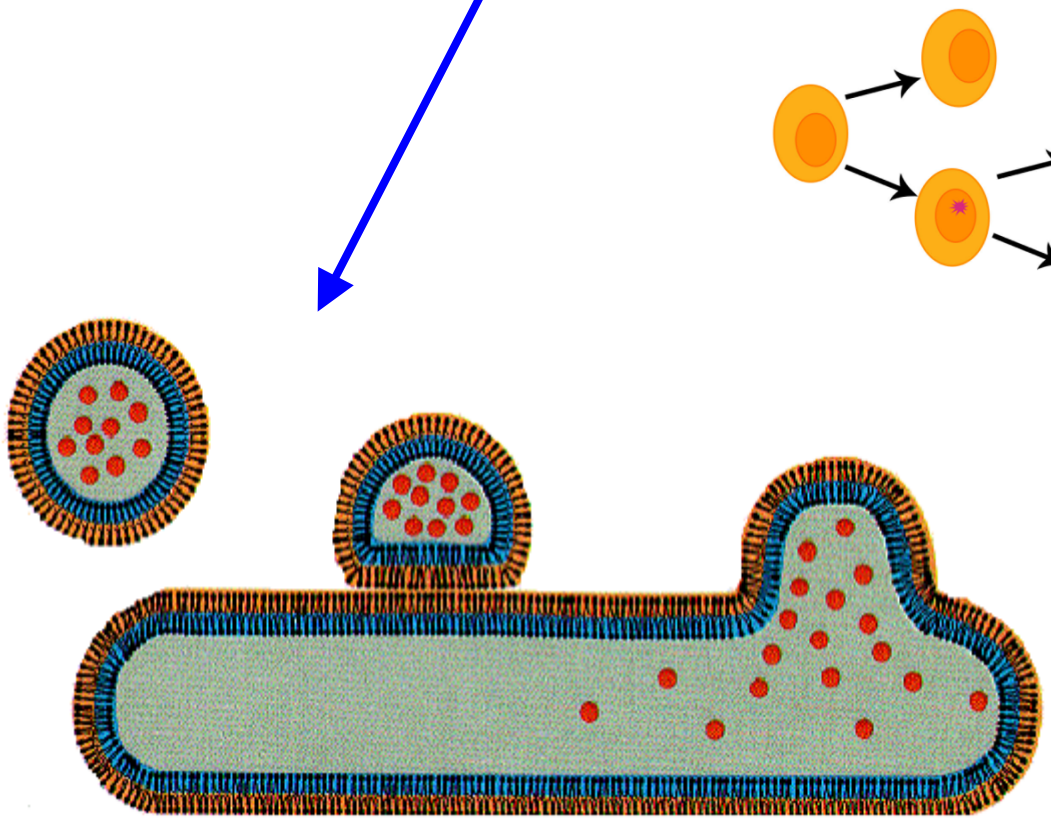
Table 2: What are the novel molecules in this cancer?

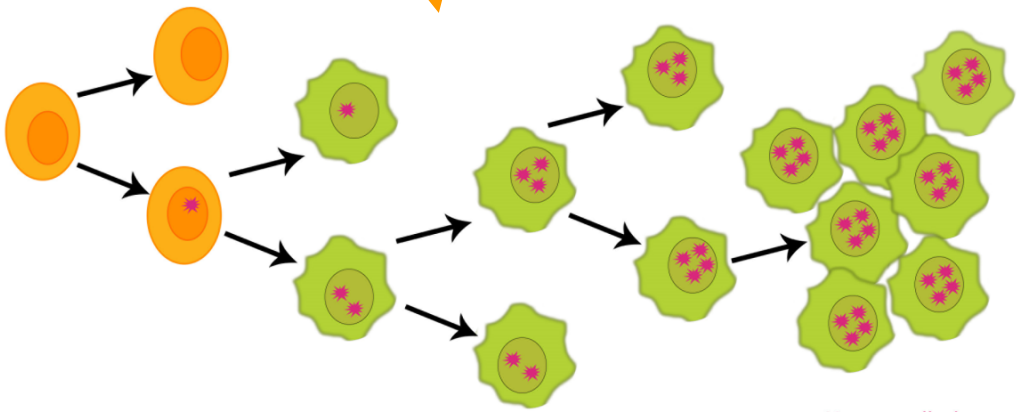
Table 2 A partial list of novel molecules found to be highly expressed in the OSCC tissue proteome

Gene symbol	Protein	Localization	Biological function	Fold change
RAB2A	Ras-related protein Rab-2A isoform a	Cytoplasm, nucleus, endoplasmic reticulum	It is a small molecular weight membranous protein that is involved in vesicular fusion and organelle biogenesis further irregularity in RAB function leads to cancer. In OSCC tissue sample, it is highly overexpressed	4.6
PRDX1	Peroxiredoxin-1	Nucleus, cytoplasm	It is an antioxidant which has a serious role in cancer progression and plays a significant role in the growth and development of several human cancers and influences in various cellular processes including cell survival, proliferation, and apoptosis	2.2
CDSN	Corneodesmosin	Cytoplasm, nucleus	It is expressed in epithelial cells in the infiltrative aggressive relapsing tumor of basal cell carcinoma	2.1
SCRN1	Secernin-1 isoform c	Cytoplasm, nucleus	Secerinin-1 is a cytosolic protein that excites exocytosis in mast cells, but the mechanisms of its function in exocytosis remain unexplored	2.1
RAB14	Ras-related protein Rab-14	Cytoplasm, nucleus	It is a low molecular mass GTPases that are involved in intracellular membrane trafficking. It is shown to be overexpressed in human non-small cell lung carcinoma which has severe role in tumorigenesis	1.9
LUM	Lumican	Cytoplasm	Lumican belongs to the family of small interstitia has a significant role in regulating cell behavi repair, and cancerous growth. In stromal tissu growth and invasion of various types of canc aggressiveness during tumor progression	



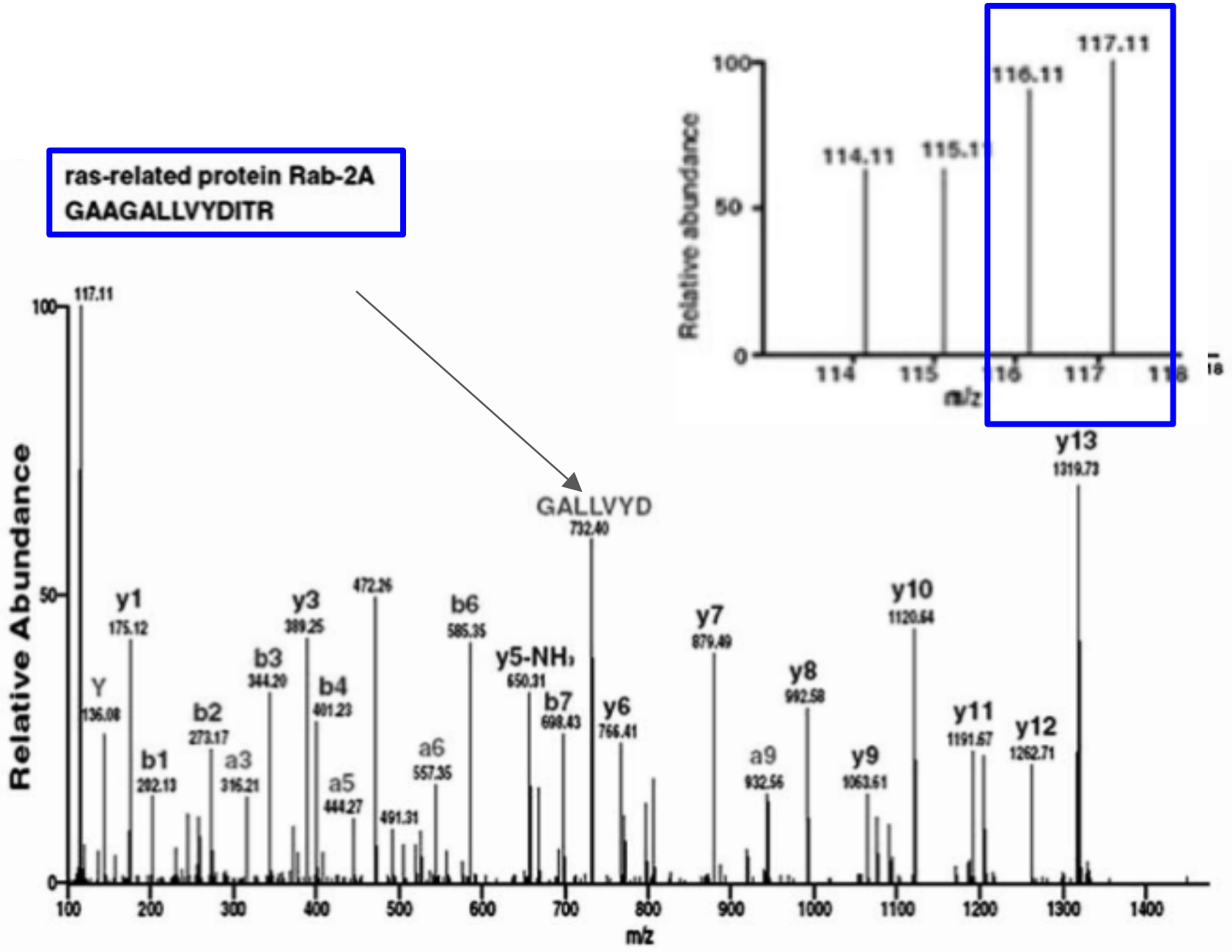
Why do we care about RAB2A and PRDX1?





http://jonlieffmd.com/blog/vesicles-transport-information
http://yorkshirecancerresearch.org.uk/wp-content/uploads/2014/07/DNA-mutation-causes-cancer-diagram.png

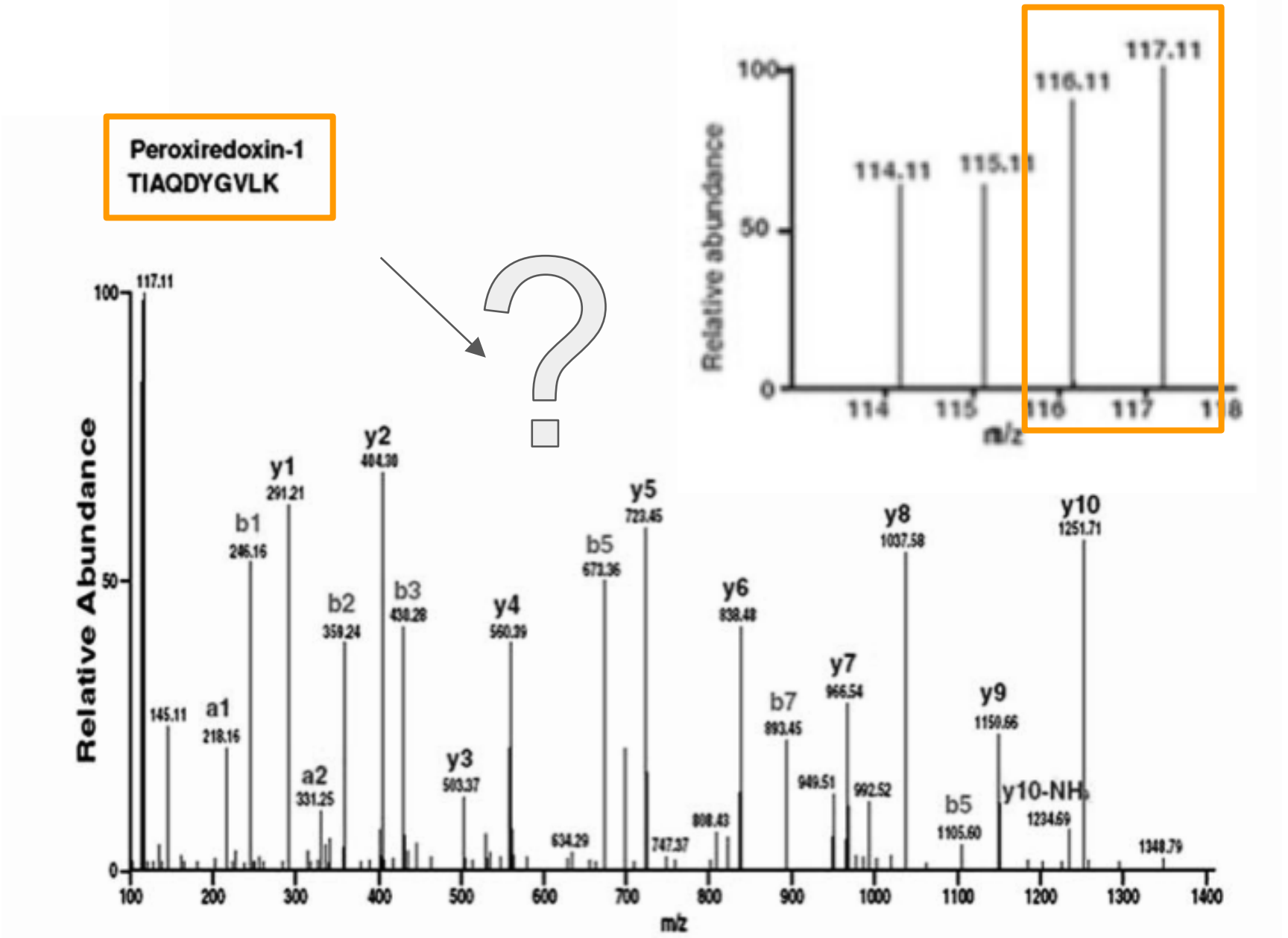
Figure 2A: What is the RAB2A peptide MS spectra?



Which peak corresponds to RAB2A fragment?

4506365	RAB2A	ras-related pr	32.08	4	6	6	7	212	23.5	6.54													
											TASNVEEAF	2	1	4506365	N-Term	0.949	1.059	1.080	4.40	0	2	1782.948 891.9777 0.286276 30.76845833
											IQEGVFDIN	2	1	4506365	N-Term	0.850	1.113	1.121	3.60	0	3	2049.084 683.6998 -0.30903832 55899
											GAAGALLV	4	1	4506365	N-Term	1.860	0.797	1.064	4.16	1	3	1619.943 540.6525 4.940307 28.579895
											GAAGALLV	4	1	4506365	N-Term	1.110	1.115	0.990	5.29	0	2	1463.834 732.4207 0.170609 34.94948333
											LQIWDTAG	4	1	4506365	N-Term	1.059	1.024	0.881	4.22	0	2	1694.861 847.9342 -0.38979332 34504333
											SCLLLQFTD	2	1	4506365	N-Term	14.448	0.857	1.160	3.40	1	3	1657.901 553.3052 0.302795 22.91716
[illegible]																							
											LADDLGFEF	1	1	4759000	N-Term	0.780	0.762	0.822	3.34	0	3	1947.996 650.0035 0.908403 51.92689333
											YADDSFTPA	2	1	4759000	N-Term	0.993	1.083	0.902	5.15	0	2	2368.206 1184.606 4.819371 52.248455
19923260	RAB4A	ras-related pr	11.93	1	2	2	2	218	24.4	6.07													
											SCLLHQFIE	1	1	19923260	N-Term	0.298	1.052	1.062	3.33	0	3	1551.847 517.9539 13.39495 27.39915333
											FKDDSNHTI	1	1	19923260	N-Term	1.062	0.940	1.524	3.66	1	4	2213.150 554.0430 -2.72509216.45045
19923262	RAB5A	ras-related pr	16.28	5	2	2	2	215	23.6	8.15													
											AVDFQEAQI	1	1	19923262	N-Term	1.162	2.533	0.921	3.71	0	3	2968.430 990.1482 5.950159 48.36561667
											FEIWDTAG	5	1	19923262	N-Term	0.569	1.031	0.950	3.29	0	2	1495.728 748.3677 -1.01302927.53206833
34403020	RAB6A	ras-related pr	8.00	3	1	1	1	175	19.9	5.07													
											AKELNVMF	1	3	1	34403020	N-Term	0.827	0.989	1.086	5.18	1	3	2013.143 671.7193 0.992231 30.28343
34147513	RAB7A	ras-related pr	33.82	1	6	6	14	207	23.5	6.70													
											DPENFPFVV	3	1	1	34147513	N-Term	0.914	0.967	0.916	4.33	0	2	1763.958 882.4827 0.606693 41.31111667
											NNIPYFETS	2	1	1	34147513	N-Term	3.162	0.825	1.038	3.73	0	2	1571.830 786.4186 -0.51128022.332435

Figure 2B: What is the PRDX1 peptide MS spectra?



Which peak corresponds to PRDX1 fragment?

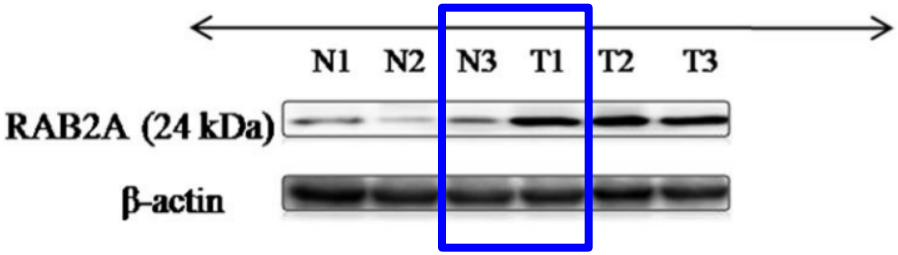
4505591	PRDX1	peroxiredoxin	48.74	1	10	13	34	199	22.1	8.13

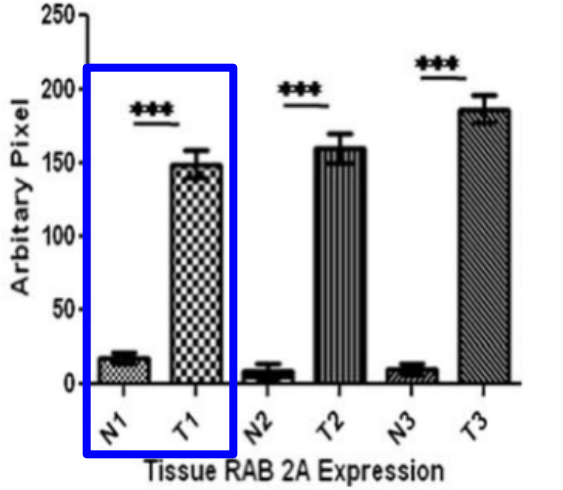
TIAQDYGVI	2	1	1	4505591	N-Term(i	0.852	1.030	0.920	4.76	0	2	1395.808;698.4078;-0.15768;25.64602
LVQAFQFTI	2	1	1	4505591	N-Term(i	0.538	0.997	0.971	4.62	0	2	1484.836;742.9220;1.119510;33.41626
ATAVMPDG	2	1	1	4505591	N-Term(i	0.563	0.961	0.957	4.08	0	2	1452.779;726.8932;-2.240604;18.19584333
GLFIIDDK	2	2	2	5453549;	N-Term(iTRAQ4plex); K8(iTRAQ4plex)				3.81	0	2	1208.711;604.8596 -0.80717;29.72694
DISLSDYK	2	1	1	4505591	N-Term(i	0.514	0.949	0.981	3.64	0	2	1228.663;614.8355 -2.02768;22.51929
ATAVMPDG	1	1	1	4505591	N-Term(i	0.497	0.915	0.993	3.08	0	2	1452.779;726.8932;-2.240604;18.19584333
ADEGISFR	1	1	1	4505591	N-Term(i	0.995	0.906	1.117	3.02	0	2	1038.535;519.7714;1.933327;17.43708167
GLFIIDDKGI	5	2	2	5453549;	N-Term(iTRAQ4plex); K8(iTRAQ4plex)				4.79	1	2	1648.014;824.5111;6.840275;30.97155333
TIAQDYGVI	5	1	1	4505591	N-Term(i	6.179	0.774	1.085	5.71	1	3	2271.235;757.75 5.745529;31.21478833
QGGLGPMN	3	1	1	4505591	N-Term(i	1.392	1.043	0.958	6.60	1	2	2067.174 1034.090;5.670816 24.23098333
ATAVMPDG	1	1	1	4505591	N-Term(i	2.046	0.812	0.985	5.37	1	2	2534.333;1267.670;6.395524;24.24746333
QITVNDLPV	5	2	2	32189392	N-Term(iTRAQ4plex)				4.22	0	2	1355.784;678.3960;6.235014;23.08444333
KQGGLGPM	2	1	1	4505591	N-Term(i	1.334	1.099	0.837	4.98	1	2	2183.256;1092.131;-0.76532;30.25304333
IGHPAPNFK	1	1	1	4505591	N-Term(i	1.062	1.314	1.046	3.35	0	3	1268.732;423.5822 -2.60692;13.20725667

TIAQDYGVI

MH	m/z	deltaM(pp	RT
1395.808	698.4078	-0.15768(	25.64602

Figure 3: How were the iTRAQ results validated?





Western Blots

Figure 4 A-B: Where do RAB2A and PRDX1 localize?

(A) RAB2A

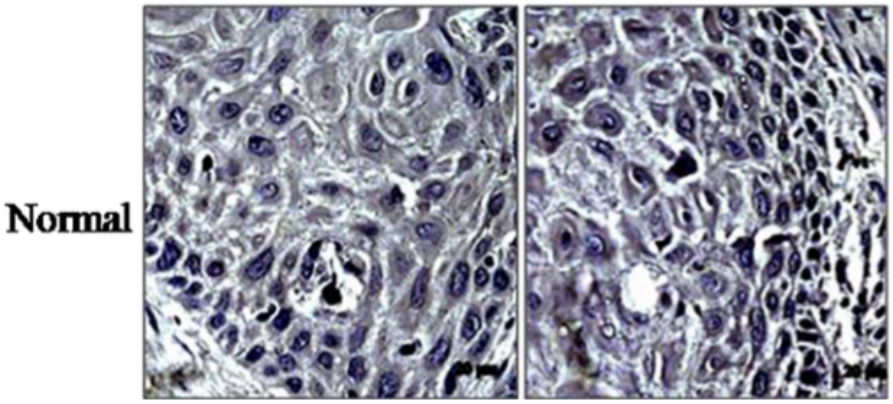
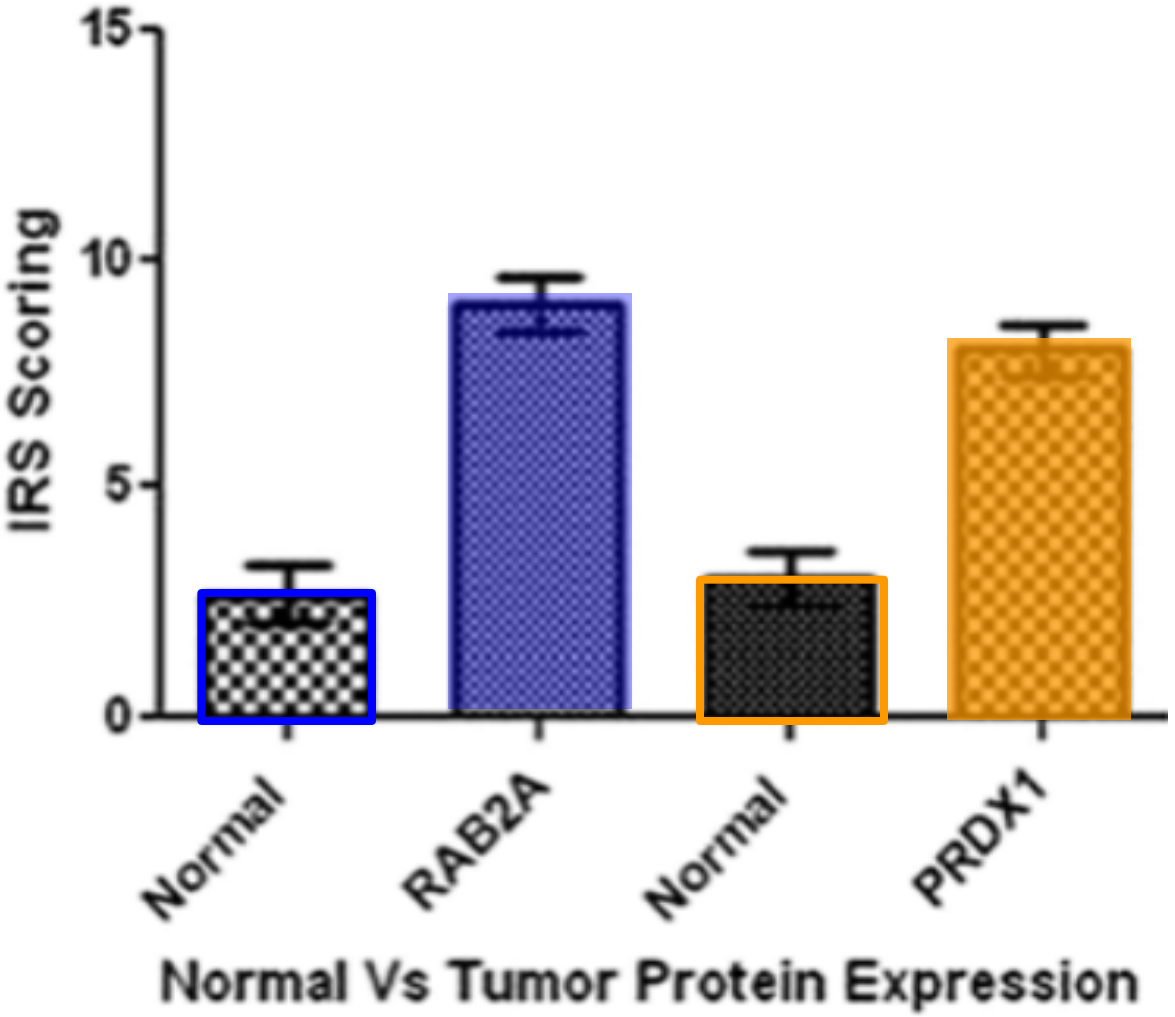
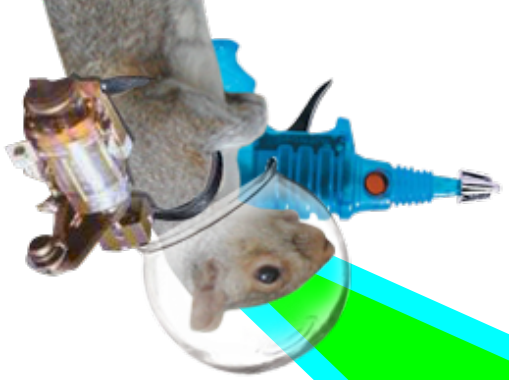


Figure 4C: How did they compare protein expression?

Immune Reactivity Scoring

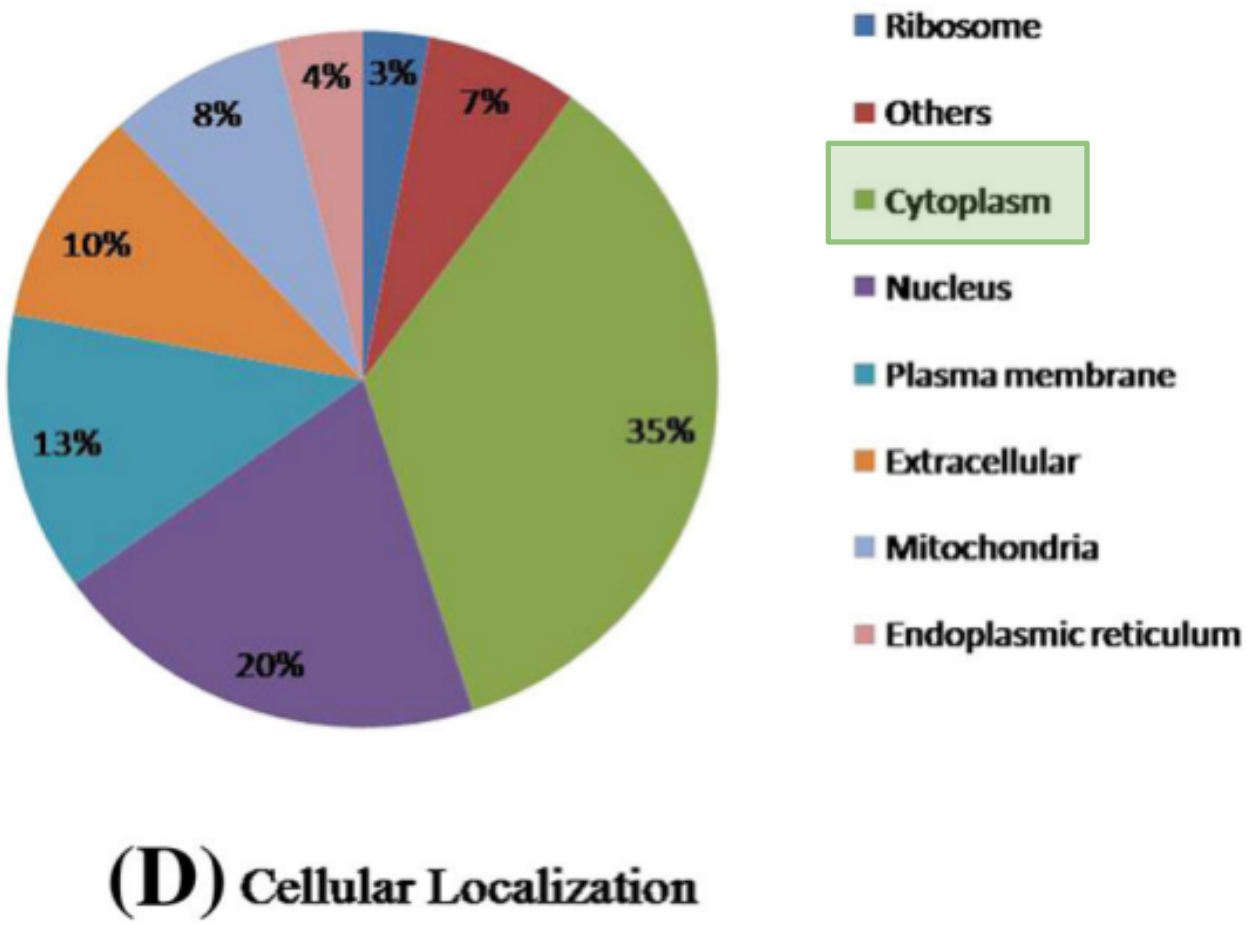




QUIZ TIME

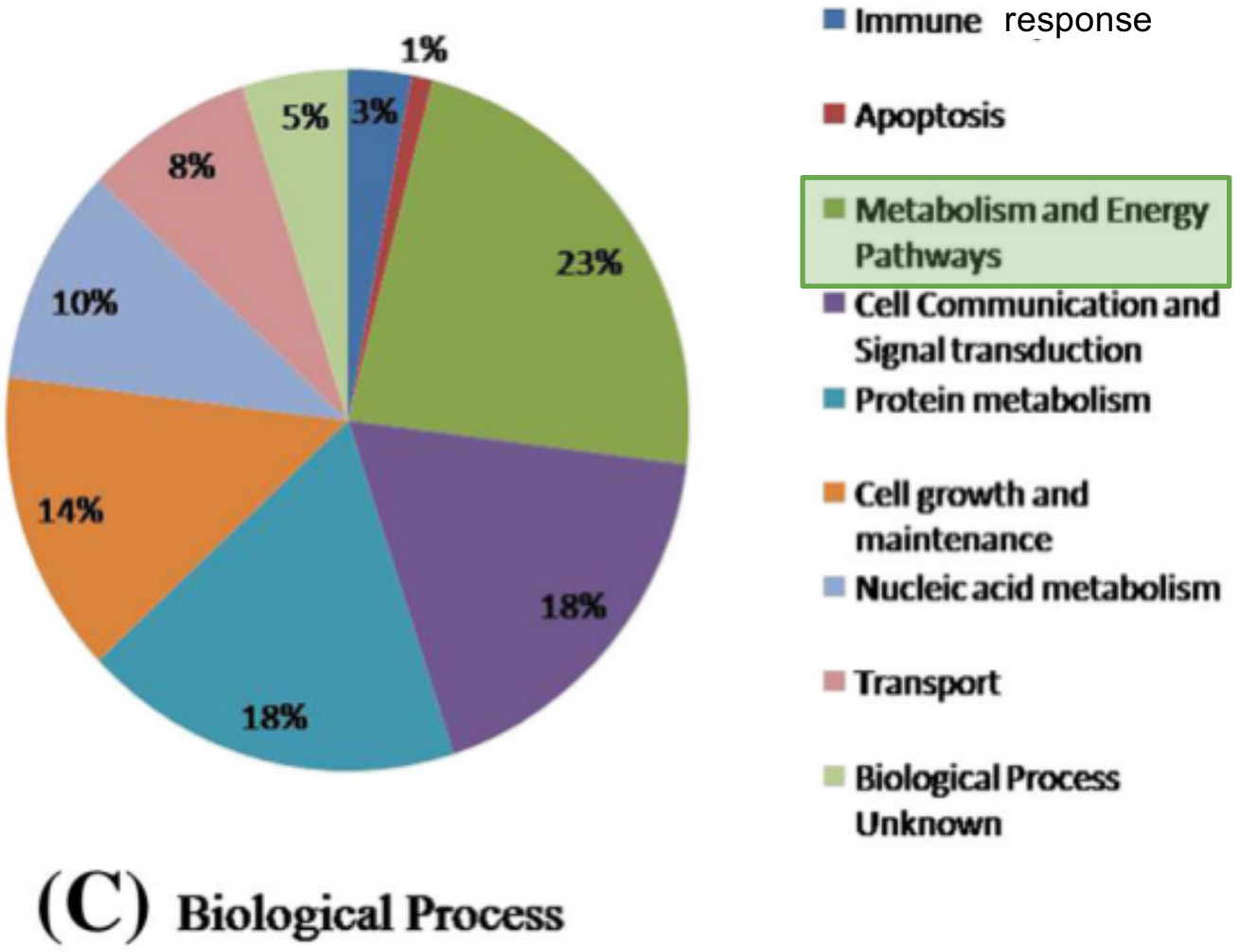
What are the 3 gene ontologies?
Bonus: give an example!

Figure 2D: How did they sort their proteins based on cellular localization?



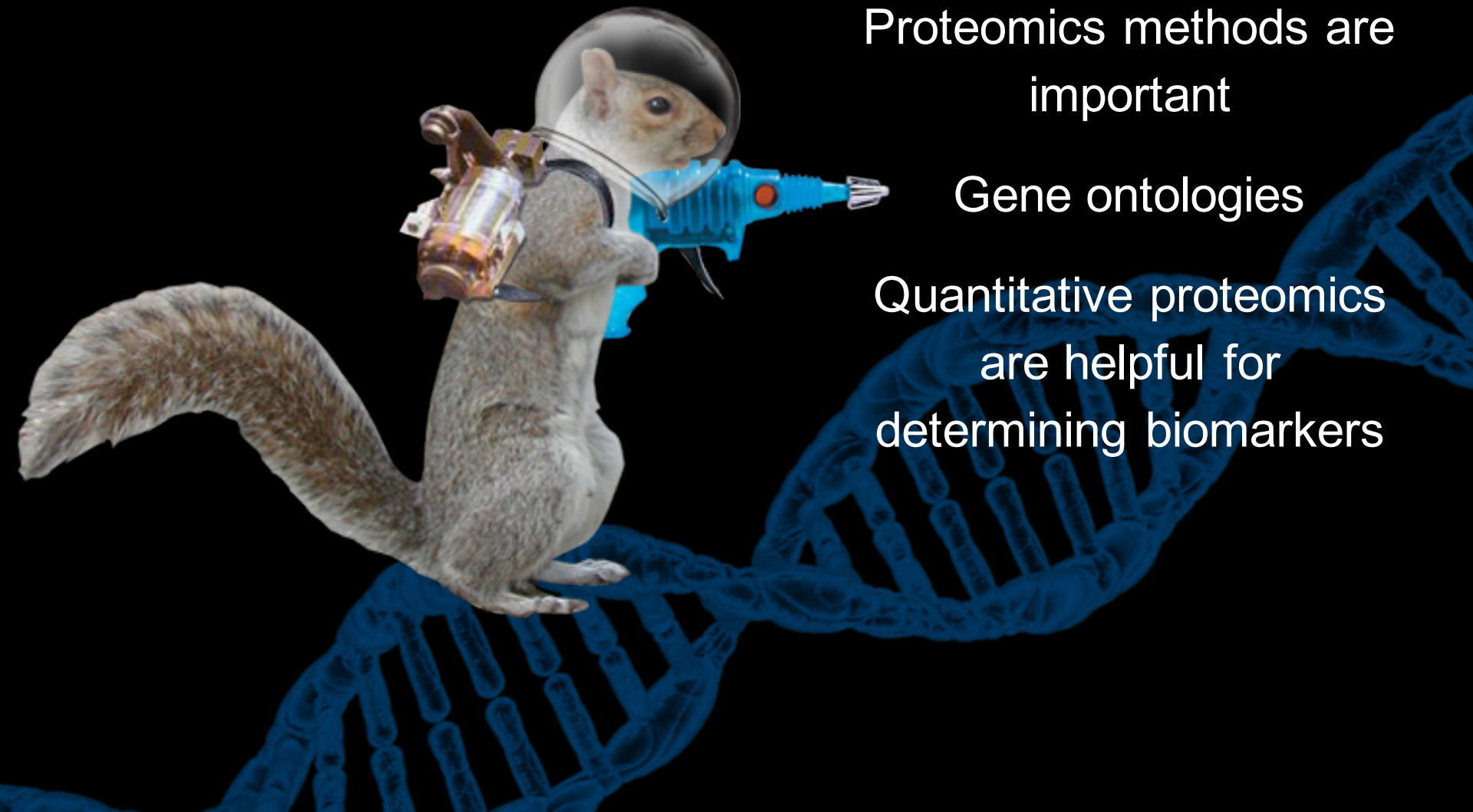
Human Protein Reference Database (http://www.hprd.org)

Figure 2C: How did they sort their proteins based on biological processes?



Human Protein Reference Database (http://www.hprd.org)

What are the main takeaways from the paper?



Proteomics methods are important

Gene ontologies

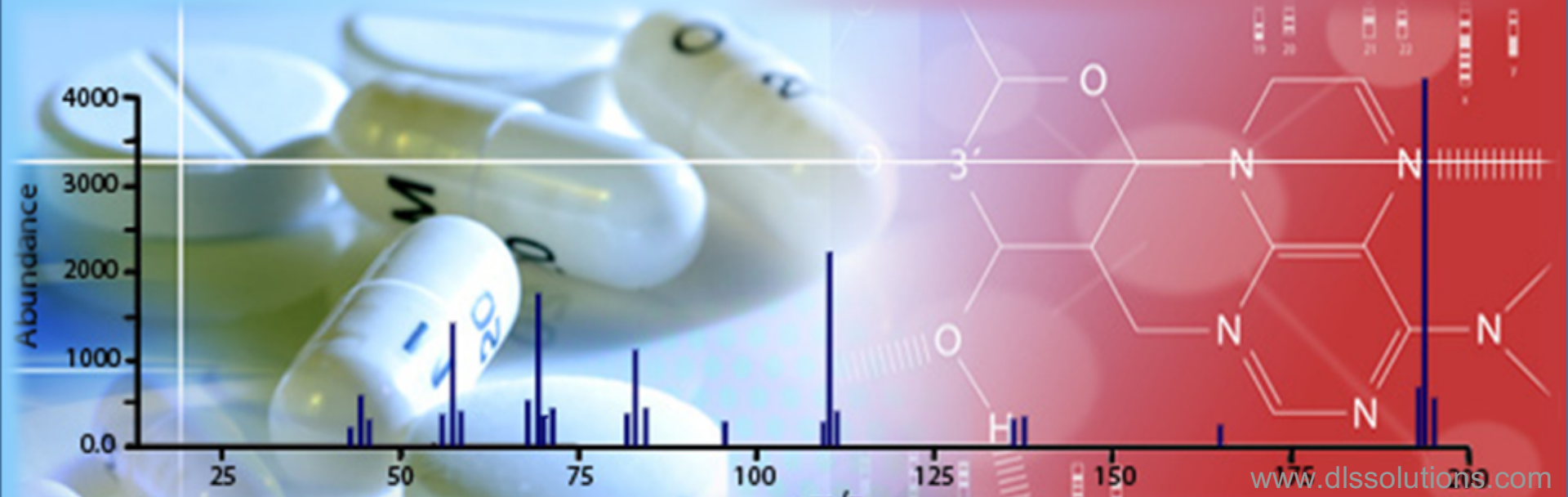
Quantitative proteomics are helpful for determining biomarkers

How are biomarkers useful in cancer research?

How does gene ontology help us move forward with quantitative proteomic data?

Why is biomarker discovery a helpful tool in cancer diagnosis?

Future research →
- Confirm biomarkers in vivo
- Determine specificity and sensitivity
- Improve biomarker strategies



Questions