Exercise 4:Protein Quantification & Identification in Mass Spectrometry

Jeremy Wong

CH-419 – Protein Mass Spectrometry and Proteomics

Lectures: Chapters 4 and 5

May 18, 2022

Question 1

Why is MS-based proteomics not inherently quantitative?

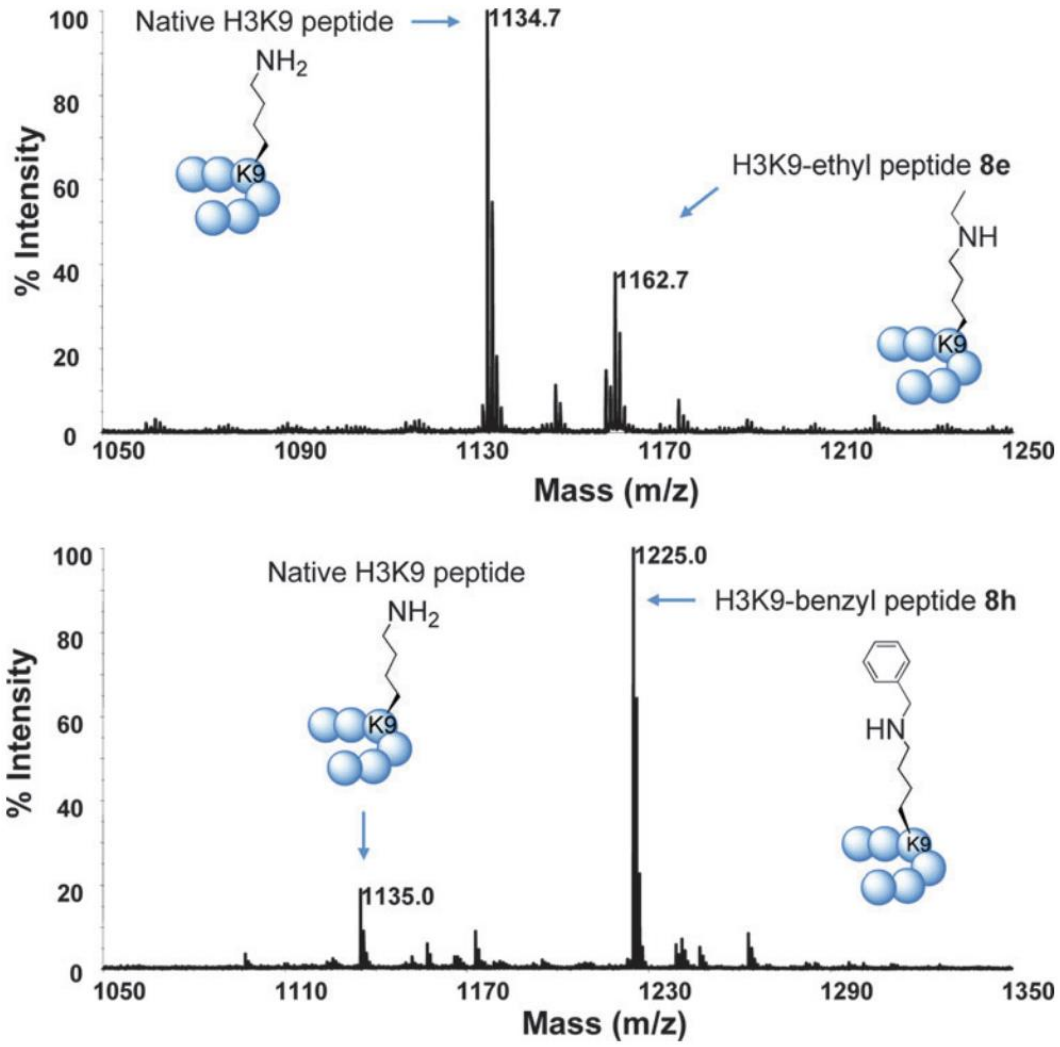
□ There are losses of peptides during analysis and differences in the ionization efficiency of peptides

□ The intensity of a peak in a mass spectrum is not a good indicator of the amount of the analyte in the sample, although differences in peak intensity of the same analyte between multiple samples accurately reflect relative differences in its abundance

□ Detection efficiencies for ions with different m/z values are unequal

□ The relationship between the amount of analyte present and measured signal intensity is complex and incompletely understood

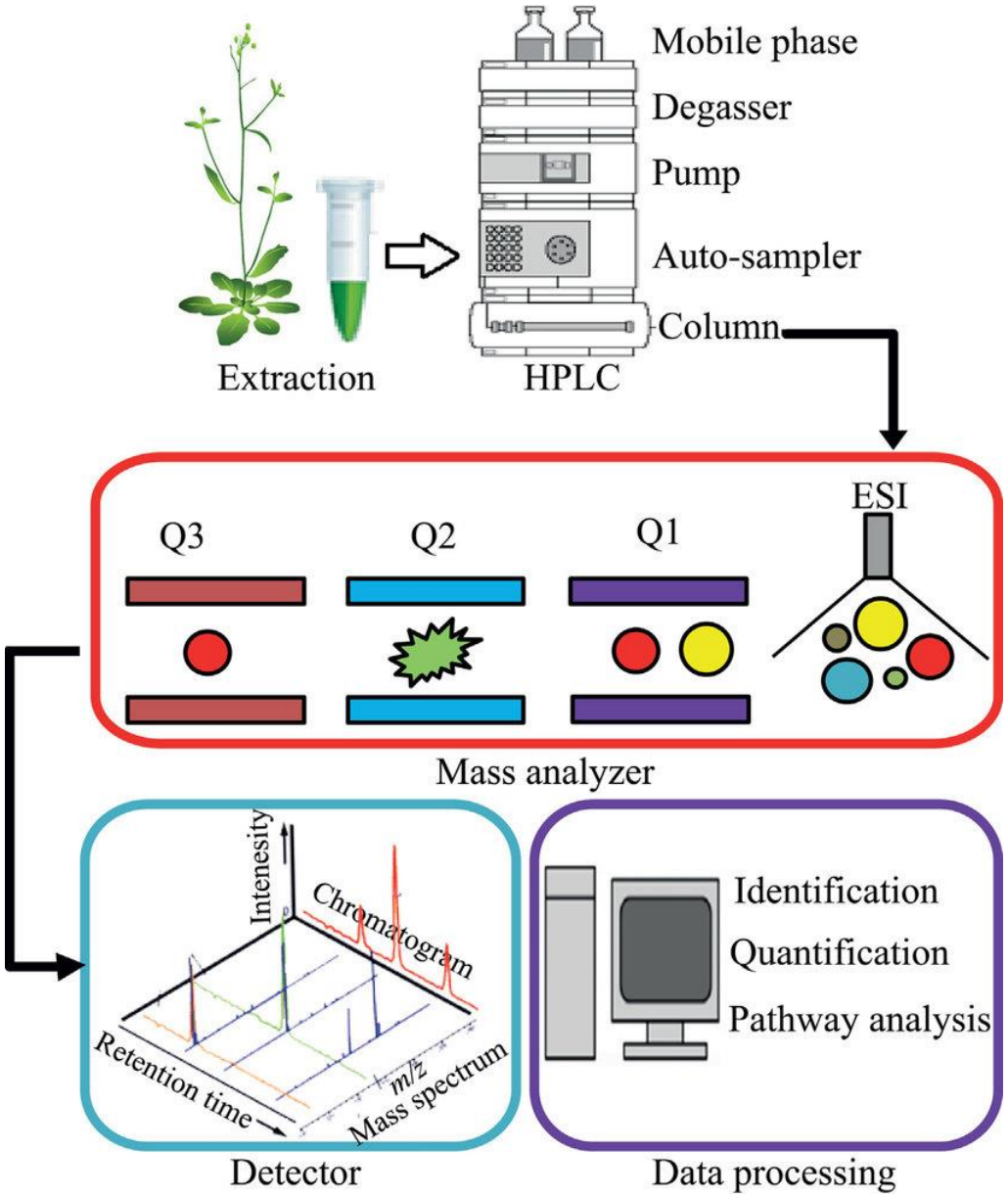
Different molecules = different ionization efficiency



- Peptides are no different. Ionization efficiency across different peptides are different
- For example: in the diagram to the right, two different peptides were injected into the mass spectrometer at equal molar ratios
- But: one peak is clearly more intense than the other even if the molar ratio is known to be the same. **This is due to difference in ionization efficiency**
- This also mean that **we cannot simply compare peak intensity across different species** to quantify samples. However, **same peaks across different samples can reflect relative differences in abundance**

Chakraborty, D. et al. Chem. Commun., 2012, 48, 1514–1516

Loss of peptide in liquid chromatography

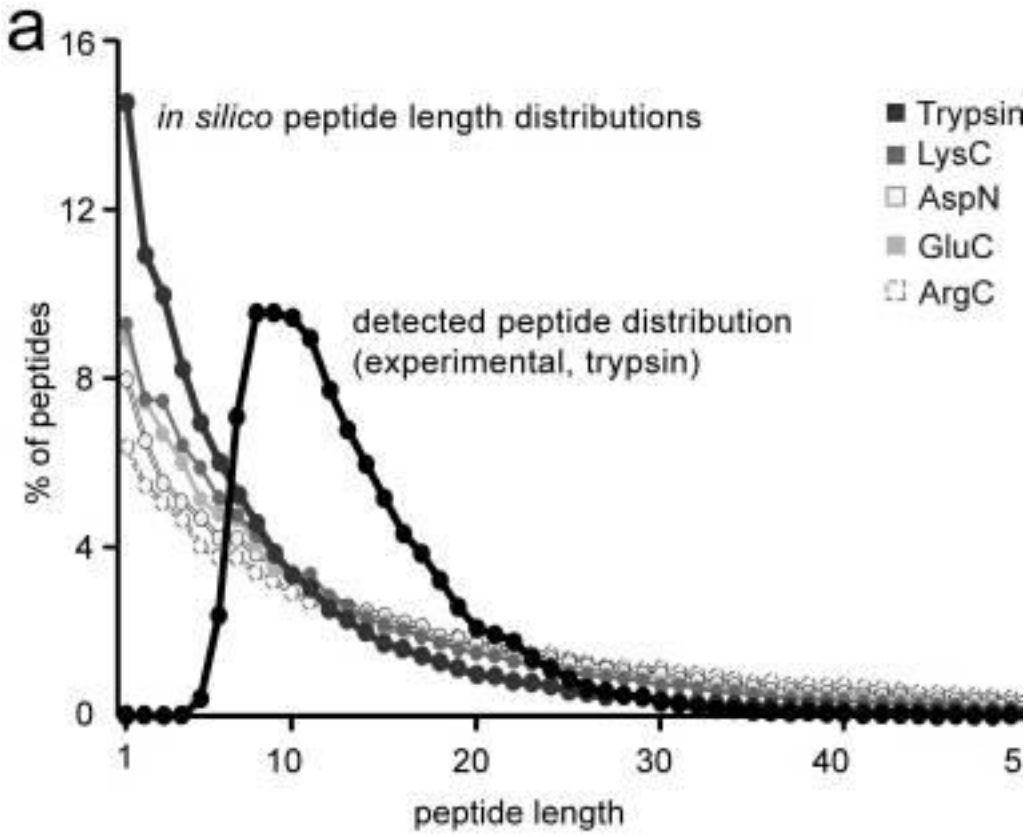


- When you separate a sample with a column, you loose some of the sample. The extent to the loss is again dependent on different molecules
- In DDA, we also only select the top 10 most abundant ions and fragment those ions only. We do not analyze everything

Detection efficiencies for ions with different m/z values are unequal

- MS is better at detecting m/z of ions within a given mass to charge range than others

- Answer: all of the above!



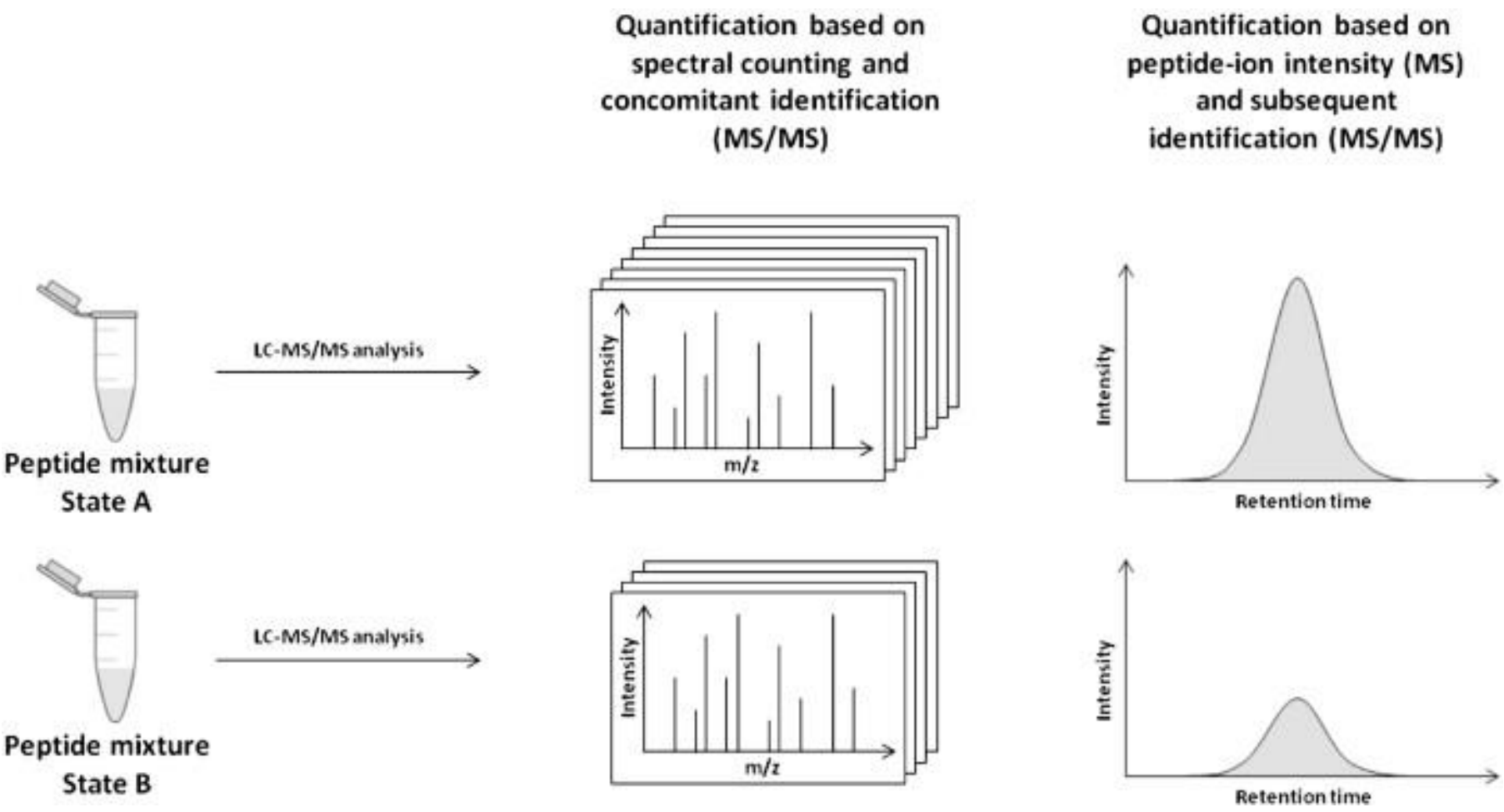
Swaney, D. L. *et al. J. Proteome Res* (2010)

Question 2

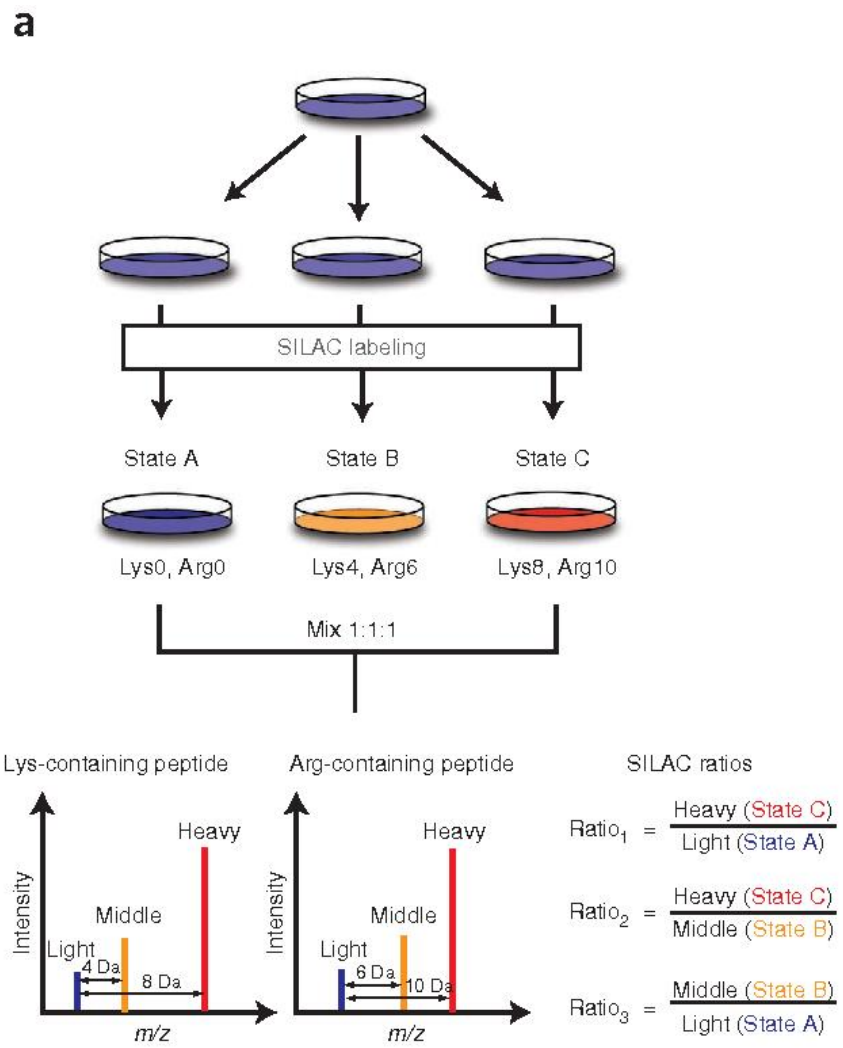
What technique(s) can provide comparison of each individual peptide between experiments/samples?

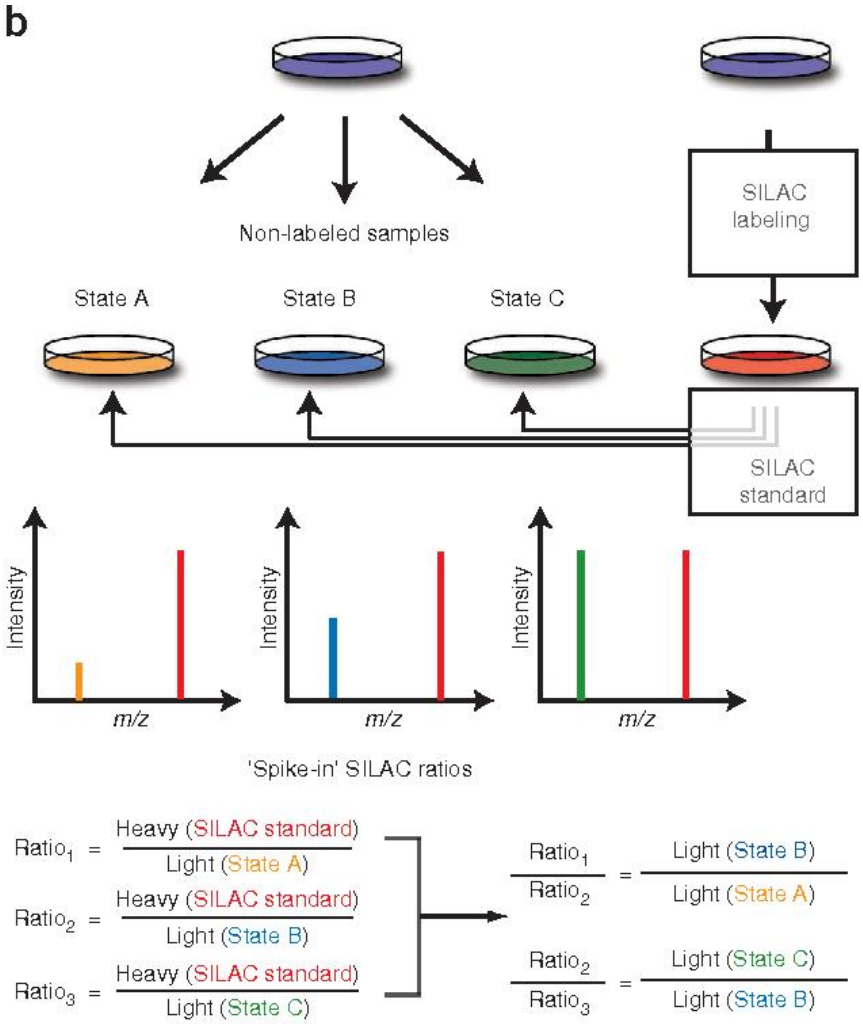
□ Label-free techniques
□ SILAC
□ Isobaric labelling
□ MS gives directly the concentrations of analytes, which can be compared between the samples

Label free quantification (LFQ)

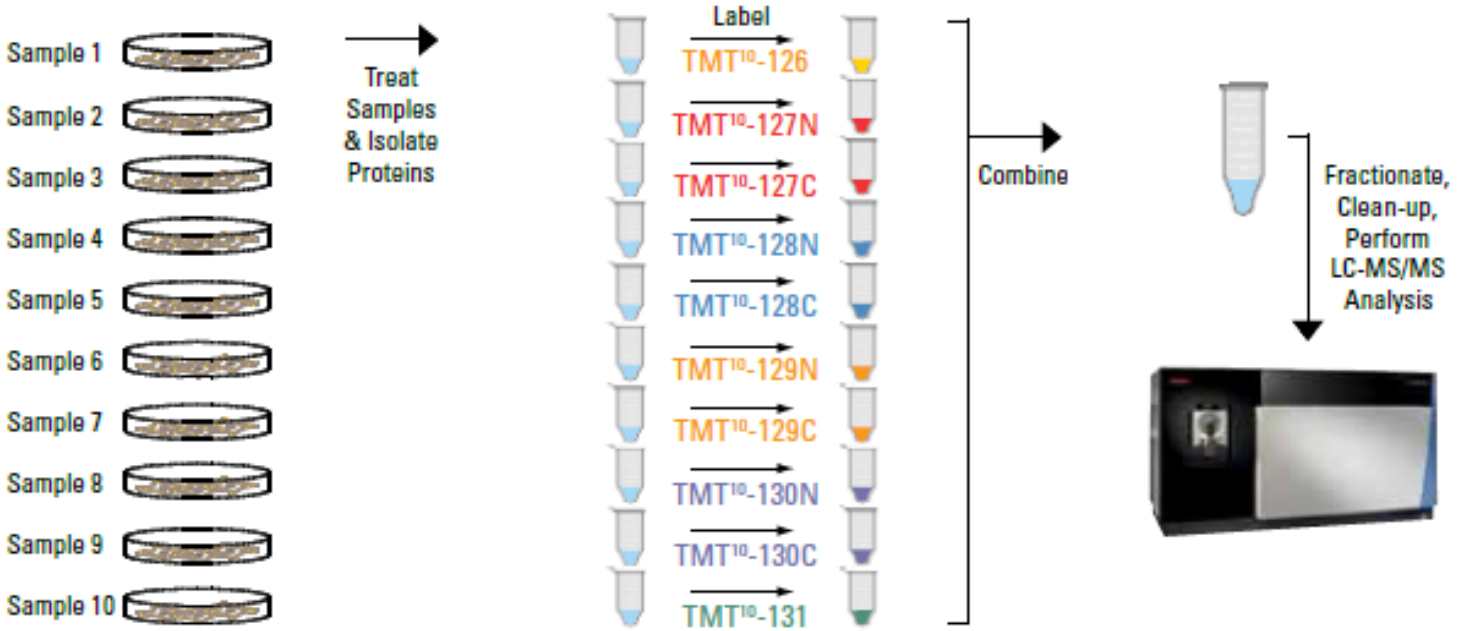


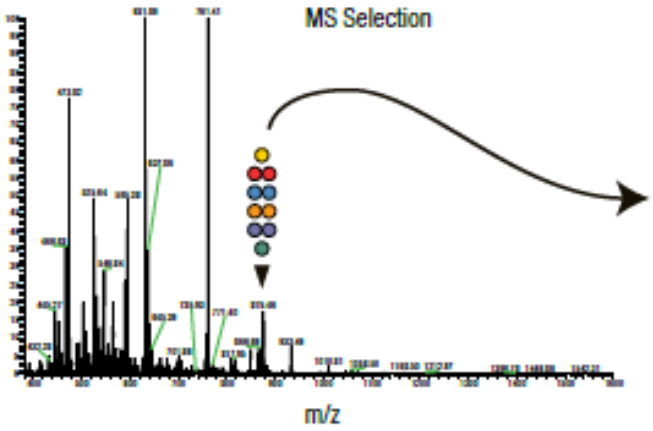
SILAC: a metabolic labelling technique

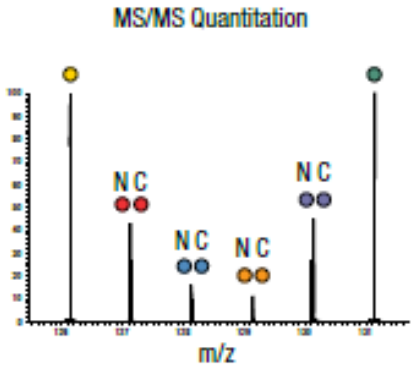


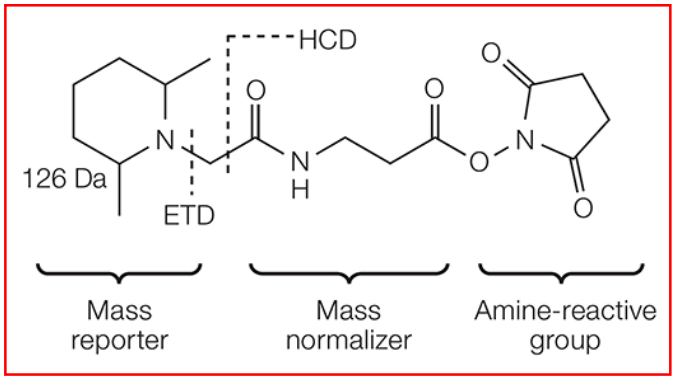


Isobaric labelling: Tandem Mass Tags









Question 2

What technique(s) can provide comparison of each individual peptide between experiments/samples?

- **Label-free techniques**
- **SILAC**
- **Isobaric labelling**
- MS gives directly the concentrations of analytes, which can be compared between the samples

Question 3

What does the spectral counting approach compare in different analyzed samples?

□ The number of all spectra associated with a specific protein
□ The sum of all precursor intensities of peptides associated with a specific protein
□ The precursor intensities of the 3 most intense detected peptides
□ The count of observed peptides versus all possible peptides

Spectra Counting – Ch. 4, Slide 7

Spectrum counting

spectral counting: quantification based on spectral counting and concomitant identification (MS/MS)
dynamic range with spectral counting: 3 orders of magnitude

Total spectra

This method uses the sum of all the spectra associated with a specific protein within a sample which includes also those spectra that are shared with other proteins and is referred to as the Total Spectrum Count

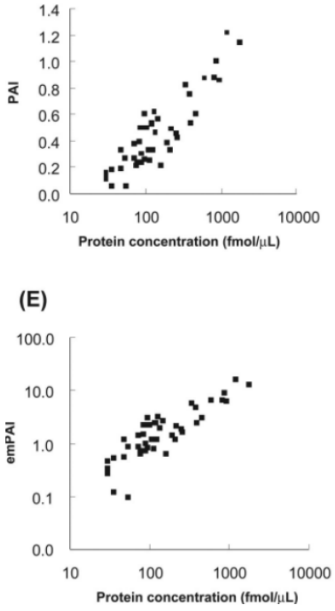
emPAI

Spectrum Counting methods can also be used in the determination of absolute abundance of proteins. The Protein Abundance Index (PAI) is defined as the number of observed peptides divided by the number of all possible tryptic peptides from a particular protein, that are within the mass range of the employed mass spectrometer:

$$PAI = \frac{N_{observed}}{N_{observable}}$$

where $N_{observed}$ is the number of experimentally observed peptides and $N_{observable}$ is the calculated number of observable peptides for each protein.
In a subsequent refinement PAI was transformed into an exponential form called emPAI and defined as follows:

$$emPAI = 10^{PAI} - 1$$



DOI: 10.1074/mcp.M500061-MCP200

7

Question 4

What are some advantages of the label-based techniques?

□ As the labeling occurs during sample preparation, quantitative artifacts are minimized
□ The techniques are usually very cheap
□ Multiplexing of samples is possible
□ They can be performed at protein or peptide level

Question 4

What are some advantages of the label-based techniques?

□ **As the labeling occurs during sample preparation, quantitative artifacts are minimized**

□ The techniques are usually very cheap (No, very expensive)

□ **Multiplexing of samples is possible** (example: combination of different isotopes + mixing into one sample)

□ **They can be performed at protein or peptide level** (example: SILAC: protein level, TMT: peptide level)

Question 5

What does SILAC stand for?

- □ Selective In-vivo Labeling After Chemical reaction
- □ Nothing in particular
- □ Static Isoforms Labels for Affinity Capture
- □ Stable Isotope Labeling with Amino Acids in Cell Culture

Answer:

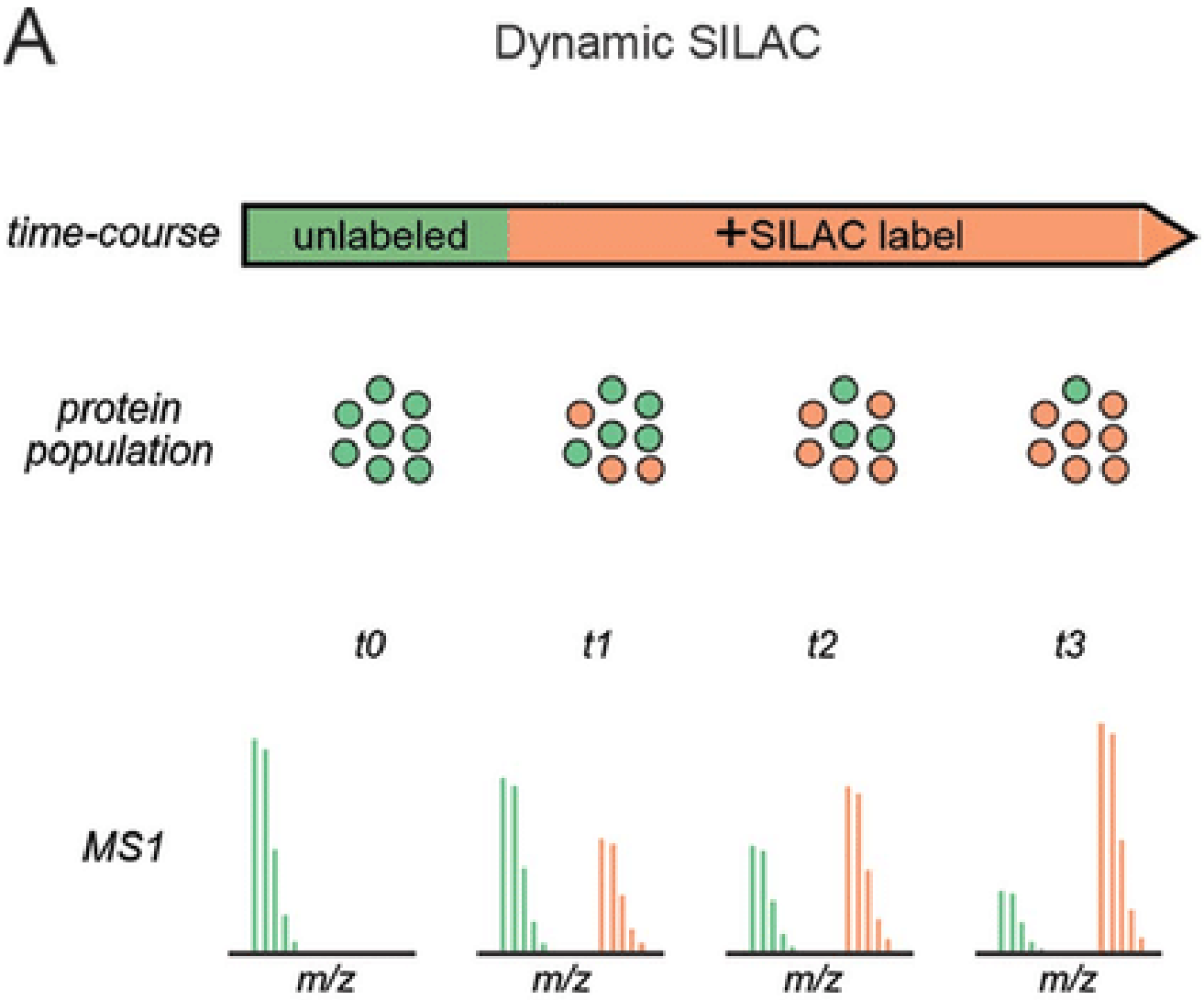
Stable Isotope Labeling with Amino Acids in Cell Culture

Question 6

At what MS level does SILAC quantification occur?

- □ MS^1
- □ MS^2
- □ MS^3
- □ MS^n

SILAC quantification happens at the MS^1 Level



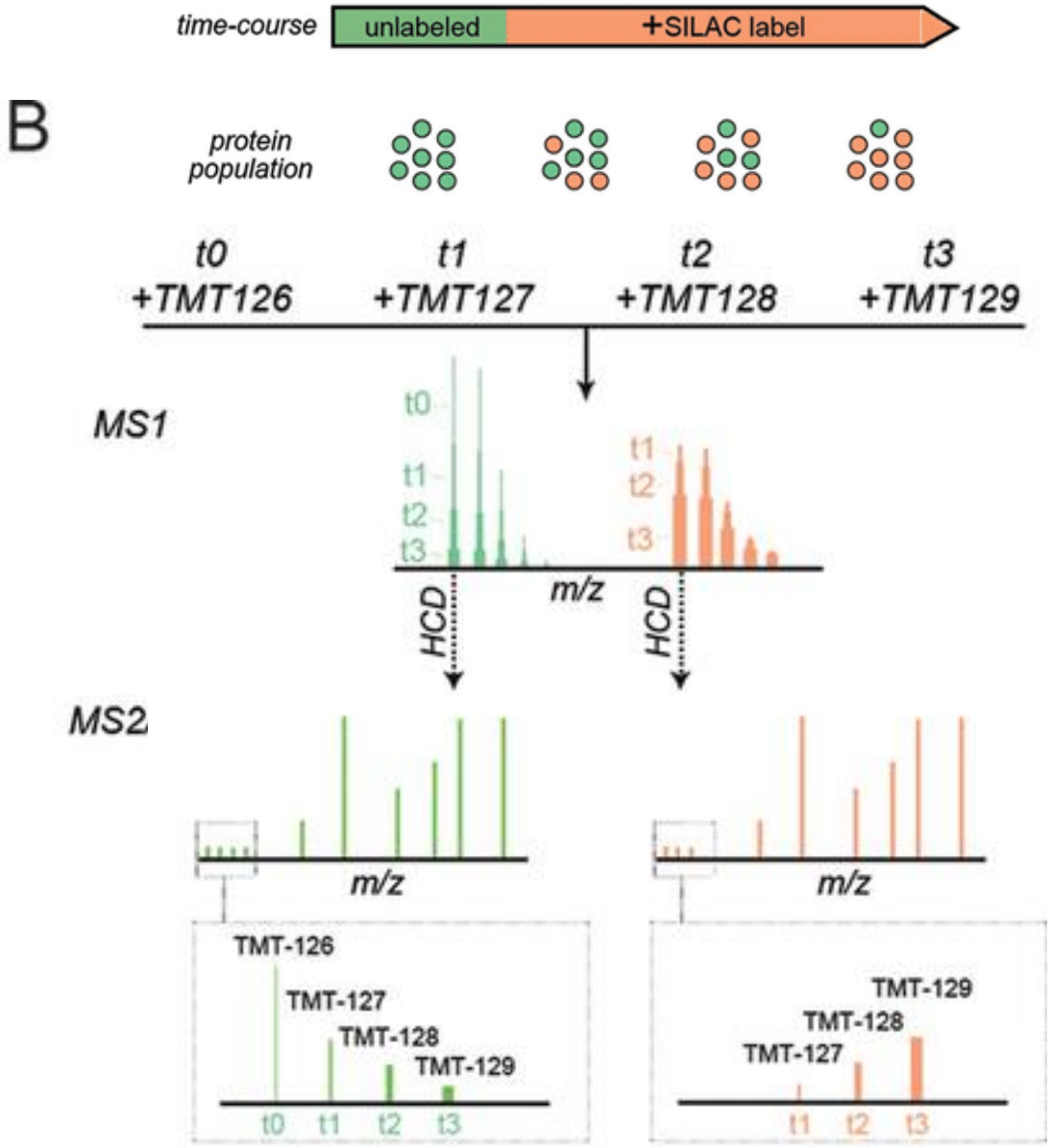
Welle, K. et al. Molecular & Cellular Proteomics 15(12), 2016

Question 7

At what MS level does TMT or ITRAQ quantification occur?

- □ MS^1
- □ MS^2
- □ MS^3
- □ MS^n

TMT and iTRAQ quantification happens at the MS^2 Level



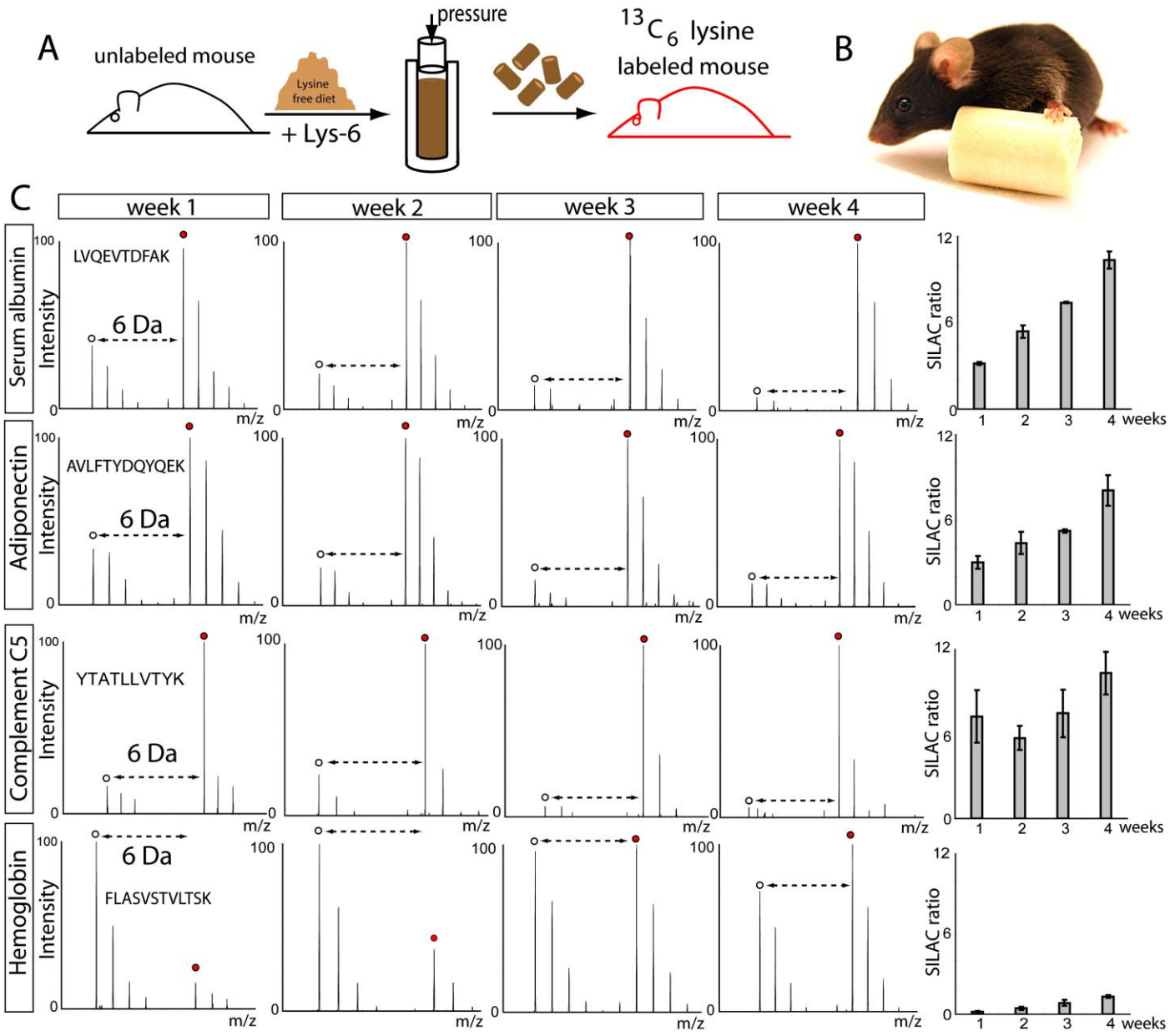
- A cool of example to show how TMT and SILAC can complement each other
- We add in SILAC labels to the cells producing protein, then we collect protein samples at given time points
- Longer exposure to SILAC labels = more population of proteins labelled by SILAC (By MS1)
- Then you digest these into peptides, label with TMT, fragment, count the relative abundance of tags
- We see that the tag becomes less abundant over time for the light species, but increases for the heavier SILAC labelled species, suggesting an increase in SILAC labelled protein as time of exposure to SILAC probe increases

Question 8

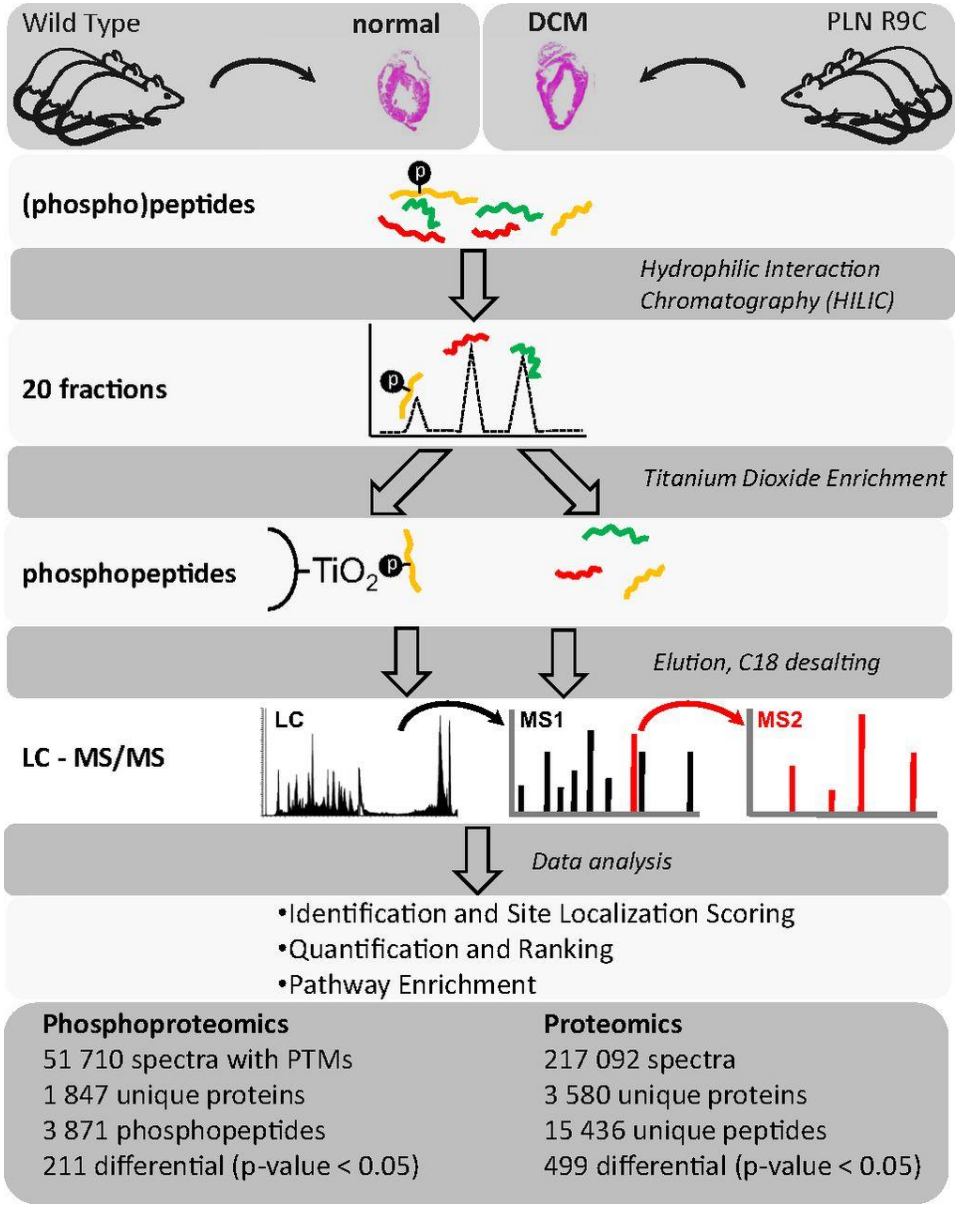
How is it possible to compare the liver proteome of two mice?

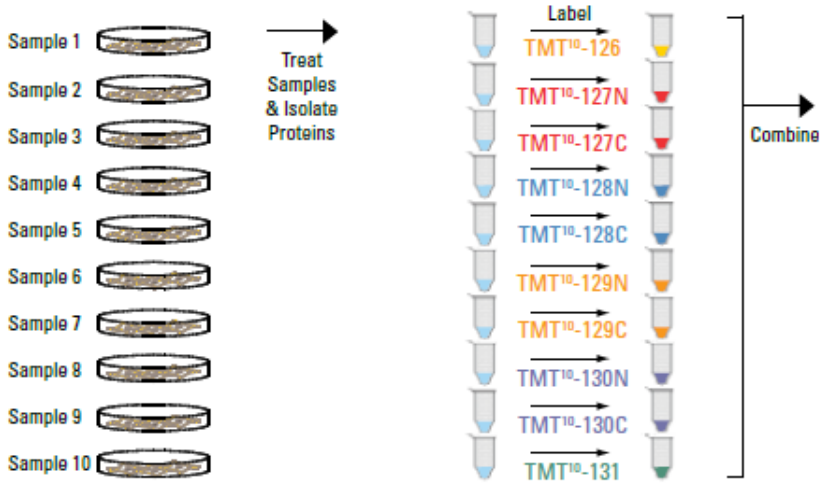
- □ Using *in-vivo* SILAC
- □ Using SILAC
- □ Using isobaric labeling
- □ Using a label-free approach

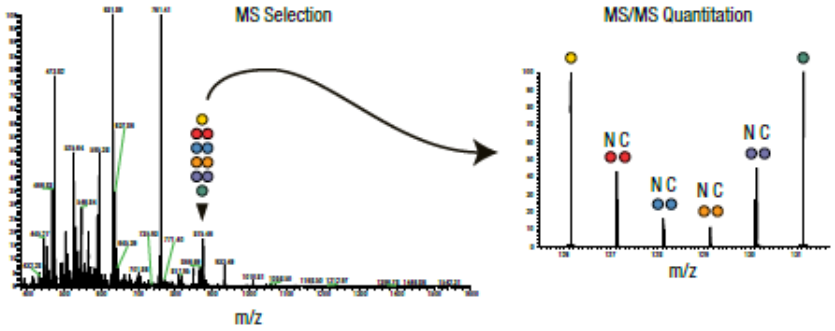
In-vivo SILAC: comparison of mice proteome



You can also first extract proteins from tissue sample, then perform LFQ or TMT





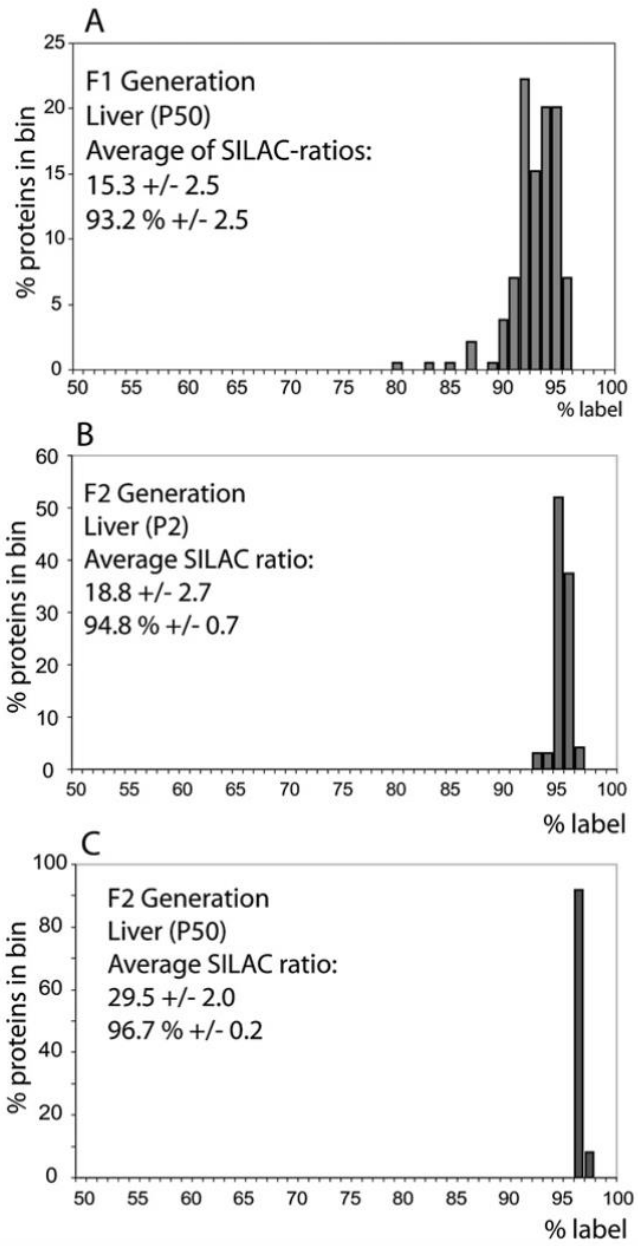


Question 9

In general, how many generations of animals are needed at least to completely label all organs in a SILAC mouse?

- □ 1
- □ 2
- □ 10
- □ 5

SILAC mice: complete organ labelling



- 2 generations of mice required for the complete labelling of all organs

Resource

Cell

SILAC Mouse for Quantitative Proteomics Uncovers Kindlin-3 as an Essential Factor for Red Blood Cell Function

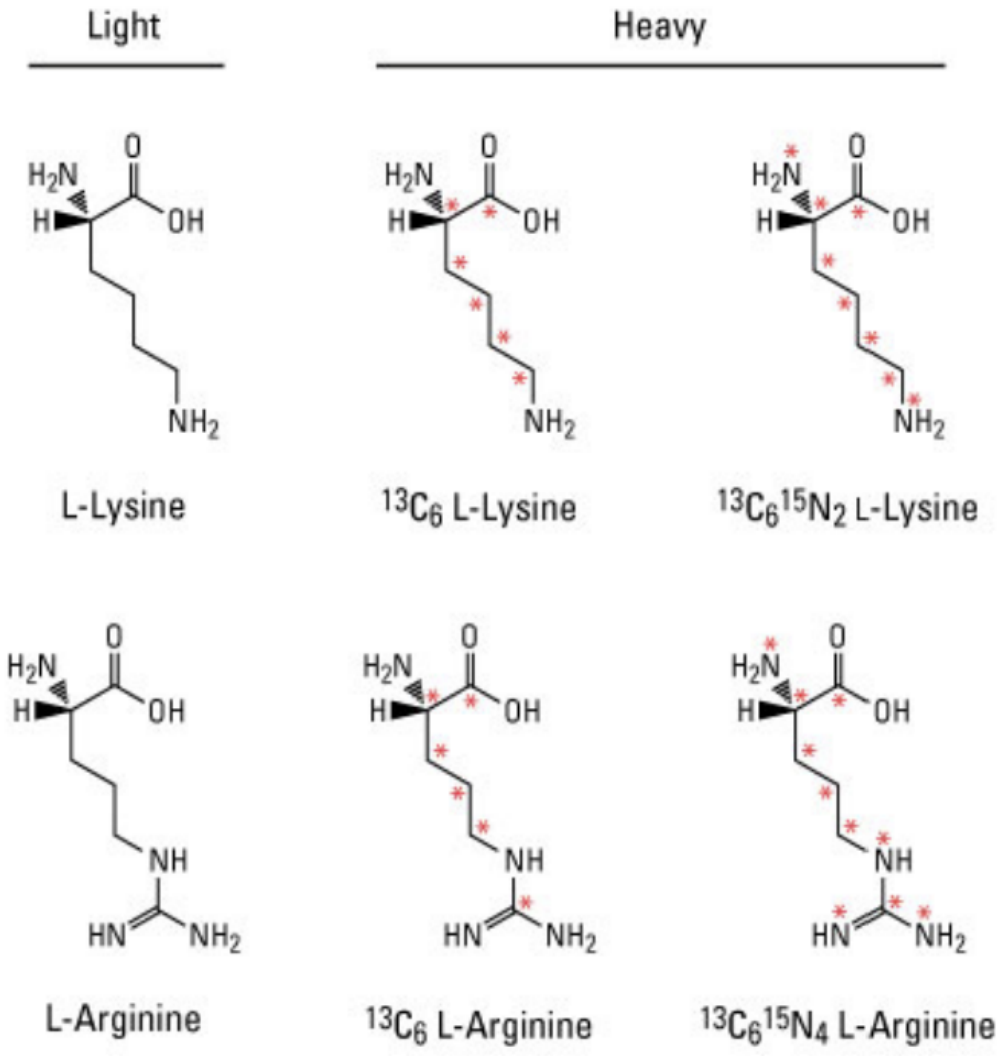
Marcus Krüger,[1,3] Markus Moser,[2,3] Siegfried Ussar,[2] Ingo Thievessen,[2] Christian A. Luber,[1] Francesca Forner,[1] Sarah Schmidt,[2] Sara Zanivan,[1] Reinhard Fässler,[2,*] and Matthias Mann[1,*]
[1]Department of Proteomics and Signal Transduction
[2]Department of Molecular Medicine
Max-Planck-Institute for Biochemistry, 82152 Martinsried, Germany
[3]These authors contributed equally to this work
*Correspondence: faessler@biochem.mpg.de (R.F.), mmann@biochem.mpg.de (M.M.)
DOI 10.1016/j.cell.2008.05.033

Question 10

What amino acids are generally available to perform a SILAC experiment?

- □ Lysine and arginine
- □ Glycine and Leucine
- □ Lysine and proline
- □ Leucine and isoleucine

Most commonly used amino acids for SILAC labelling are Arg and Lys



Question 11

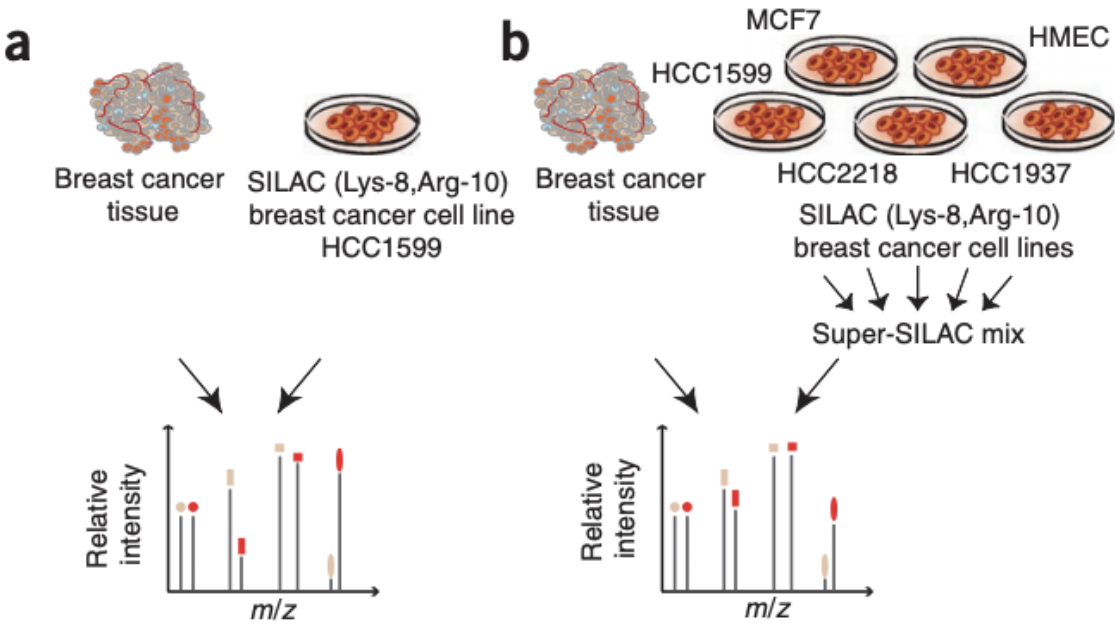
What is called a super-SILAC mix?

- □ A mixture of samples labelled with TMT
- □ A mixture of cell lines labelled by SILAC
- □ A mixture of nonlabeled healthy tissues

Super SILAC mix: superior internal standard

- The use of a single cell line as an internal standard is often not good enough for relative quantification since different cell lines of similar functions can express different proteins to a different level

- Super SILAC mix: a mixture of SILAC labelled cell lines (multiple) as an internal standard. That way, you can compare the expression level of proteins of your sample of interest versus the average between multiple cell lines.



Geiger, T. et al. Nature Methods volume 7, 2010

Question 12

What multiplexing capabilities are available with TMT?

- □ 2-plex
- □ 4-plex
- □ 6-plex
- □ 10-plex

Multiplex systems with tandem mass tags

ThermoFisher SCIENTIFIC Popular Applications & Techniques Shop All Products Services Support

Search All | Search by catalog number, product name, keyword, application

Home › Life Sciences › Protein Biology › Protein Mass Spectrometry Analysis › Protein Quantitation Using Mass Spectrometry › Tandem Mass Tag Systems

Tandem Mass Tag (TMT) Systems

‹ Protein Quantitation Using Mass Spectrometry

Tandem Mass Tag Systems

- TMT Publications

SILAC Metabolic Labeling Systems

SureQuant Targeted Mass Spec Assay Kits

The Thermo Scientific Tandem Mass Tag Reagents are designed to enable identification and quantitation of proteins in different samples using tandem mass spectrometry (MS). All mass tagging reagents within a set have the same nominal mass (i.e., are isobaric) and chemical structure composed of an amine-reactive NHS ester group, a spacer arm (mass normalizer), and a mass reporter (**Figure 1**).

The standard tandem mass tag (TMT) labeling kits and reagents enable multiplex (6-plex to 11-plex) relative quantitation using high-resolution MS for samples prepared from cells or tissues. The TMT 10plex and 11plex label reagents share an identical structure with TMTzero, TMTduplex, and TMT 6plex reagents but contain different numbers and combinations of 13C and 15N isotopes in the mass reporter.

View a list of TMT publications in peer-reviewed journals based on topic or disease ›

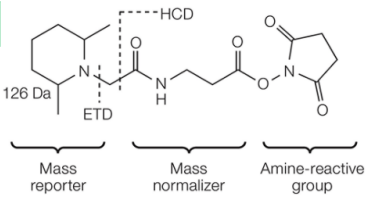


Figure 1. Functional regions of the TMT reagent's chemical structure, including MS/MS sites of fragmentation by HCD and ETD.

- 2 plex is also possible as mentioned in the lecture
- It is also possible to do a 16-plex with current technology known as TMT Pro

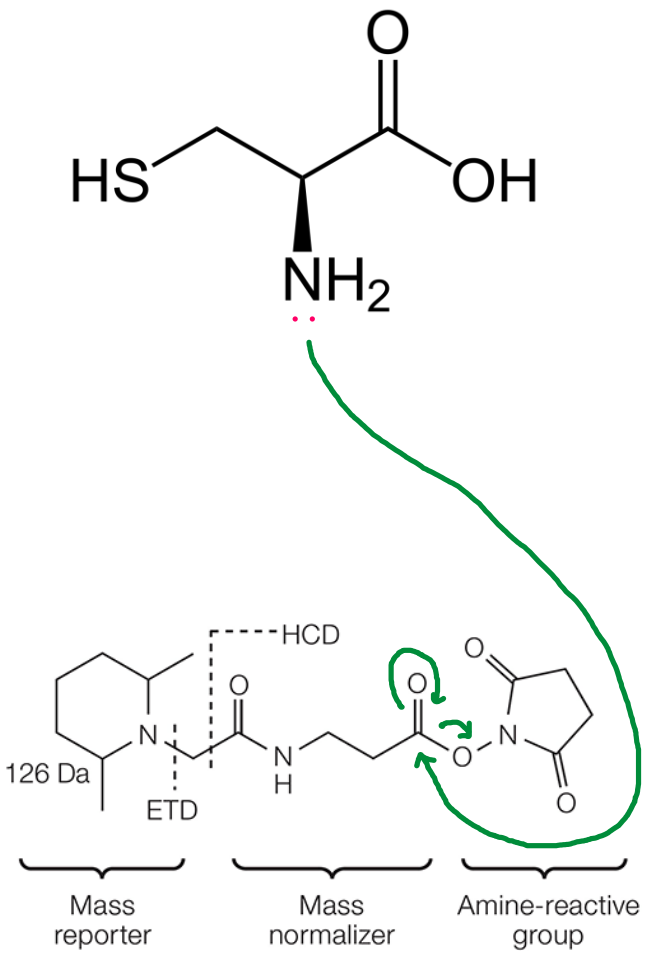
Question 13

What amino acid(s) is/are labeled with TMT or iTRAQ?

- □ Lysine
- □ Arginine
- □ N-terminus
- □ Cysteine

Labelling mechanism: tandem mass tags

Amino acid/ peptide/ protein



Mass Tag

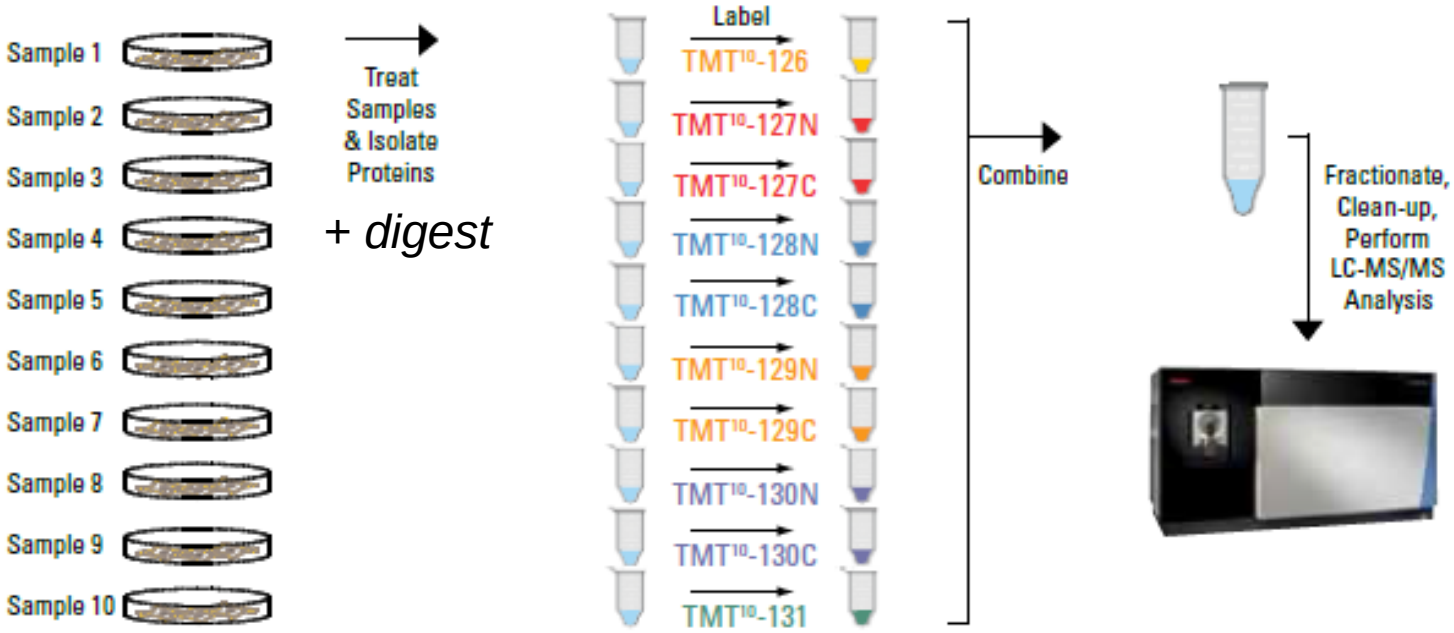
Amino acid/ peptide/ protein labelled at the N terminal

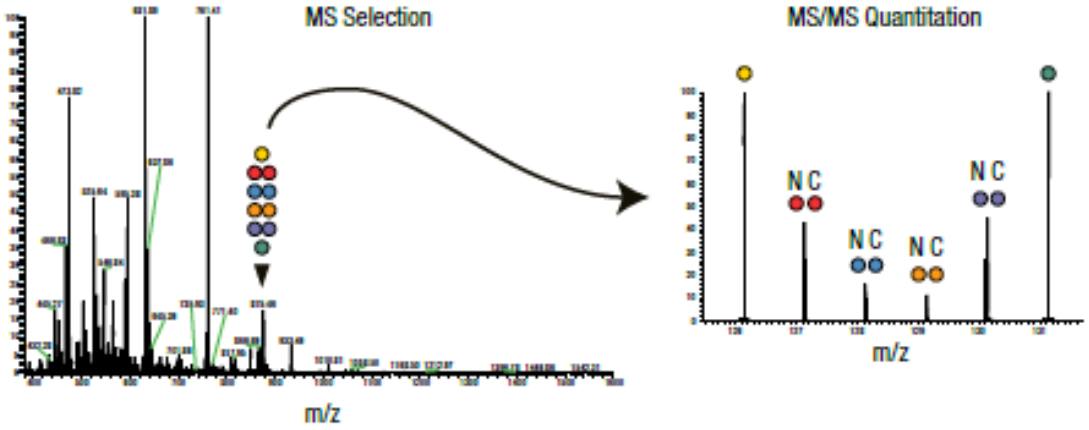
Question 14

At what level is TMT or iTRAQ labeling generally performed?

- □ Amino acid level
- □ Peptide level
- □ Protein level
- □ Cell culture level

TMT and iTRAQ are often done at the peptide level





Question 15

In order to decipher subtle changes in phosphorylation regulation in two cell lines, what quantitative approach(es) would you recommend?

- □ SILAC
- □ In-vivo SILAC
- □ Isobaric labeling
- □ Spectral counting

Quantification of subtle protein changes

- Spectra counting: not accurate
- In vivo SILAC: we do not need full labelling of mice…
- **Isobaric labelling, SILAC: good accuracy**

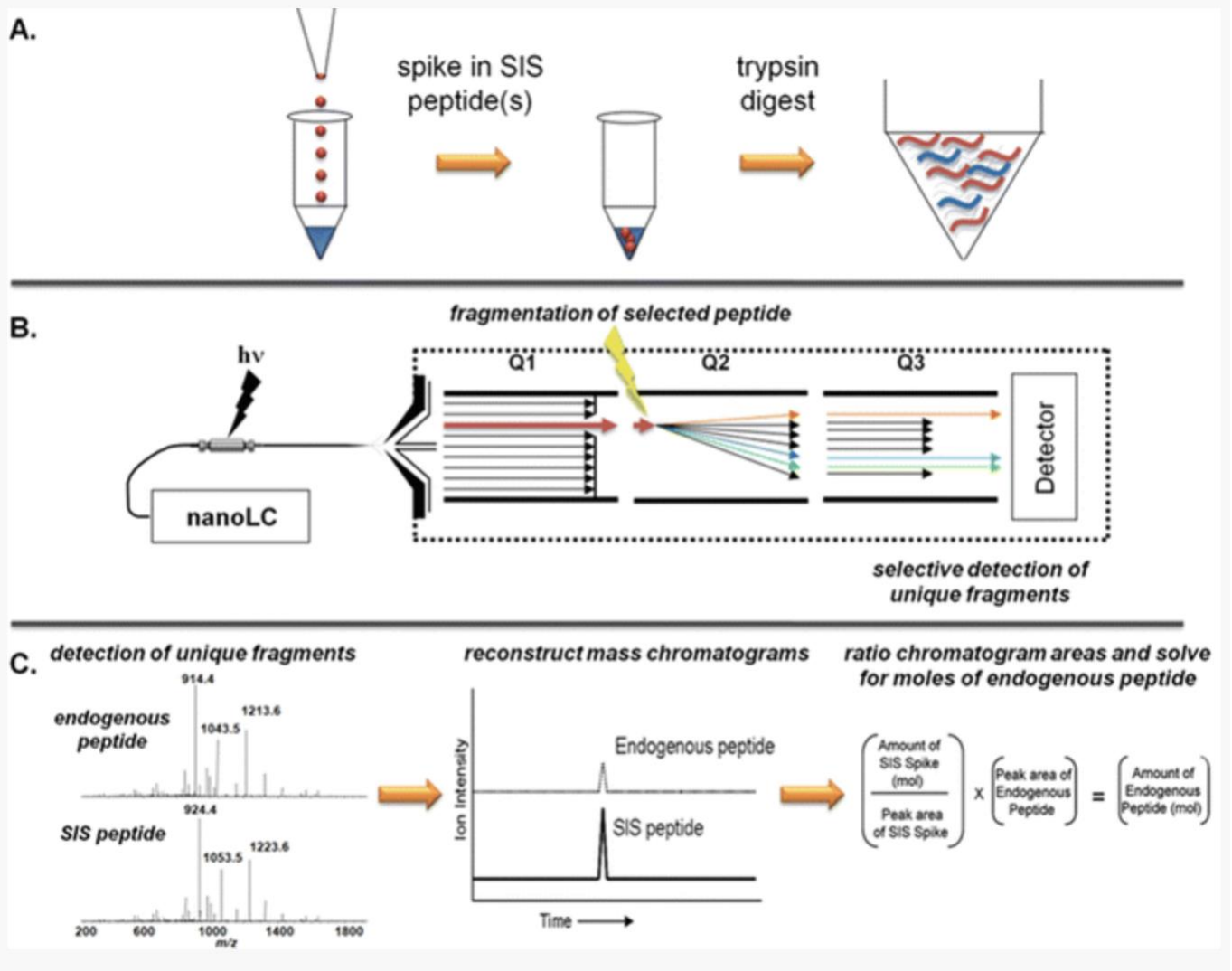
	Application	Accuracy (process)	Quantitative proteome coverage	Linear dynamic range[a]
Metabolic protein labeling	Complex biochemical workflows	+++	++	1–2 logs
	Comparison of 2–3 states			
	Cell culture systems only			
Chemical protein labeling (MS)	Medium to complex biochemical workflows	+++	++	1–2 logs
	Comparison of 2–3 states			
Chemical peptide labeling (MS)	Medium complexity biochemical workflows	++	++	2 logs
	Comparison of 2–3 states			
Chemical peptide labeling (MS/MS)	Medium complexity biochemical workflows	++	++	2 logs
	Comparison of 2–8 states			
Enzymatic labeling (MS)	Medium complexity biochemical workflows	++	++	1–2 logs
	Comparison of 2 states			
Spiked peptides	Medium complexity biochemical workflows	++	+	2 logs
	Targeted analysis of few proteins			
Label free (ion intensity)	Simple biochemical workflows	+	+++	2–3 logs
	Whole proteome analysis			
	Comparison of multiple states			
Label free (spectrum counting)	Simple biochemical workflows	+	+++	2–3 logs
	Whole proteome analysis			
	Comparison of multiple states			

Question 16

What mass spectrometer is commonly used for targeted protein quantification with stable isotope dilution?

- □ FT-ICR
- □ QqQ
- □ Ion trap
- □ MADLI-TOF

Stable isotope dilution with QqQ – SRM or MRM modes



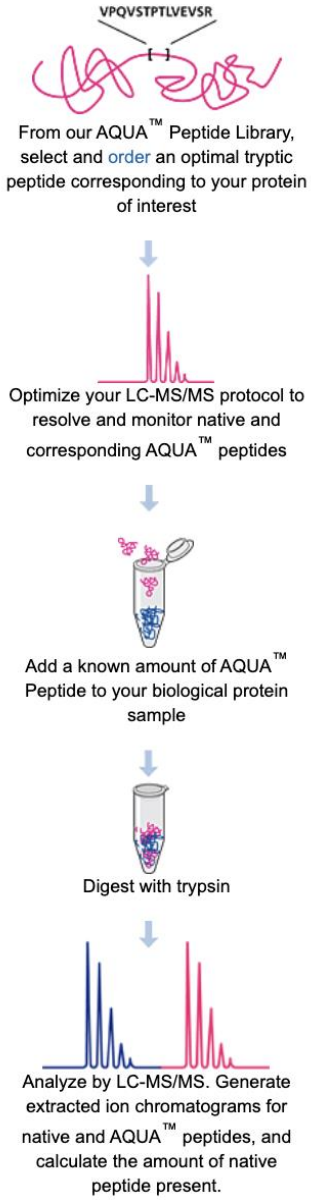
Russo, P. et al. Molecular Profiling. 2017

Question 17

How are peptides selected for their use as heavy AQUA standards?

- □ They need to be proteotypic
- □ They need to contain more than 25 amino acids
- □ They need to be heavily modified post-translationally
- □ They need to fragment efficiently

Absolute quantification peptides: AQUA



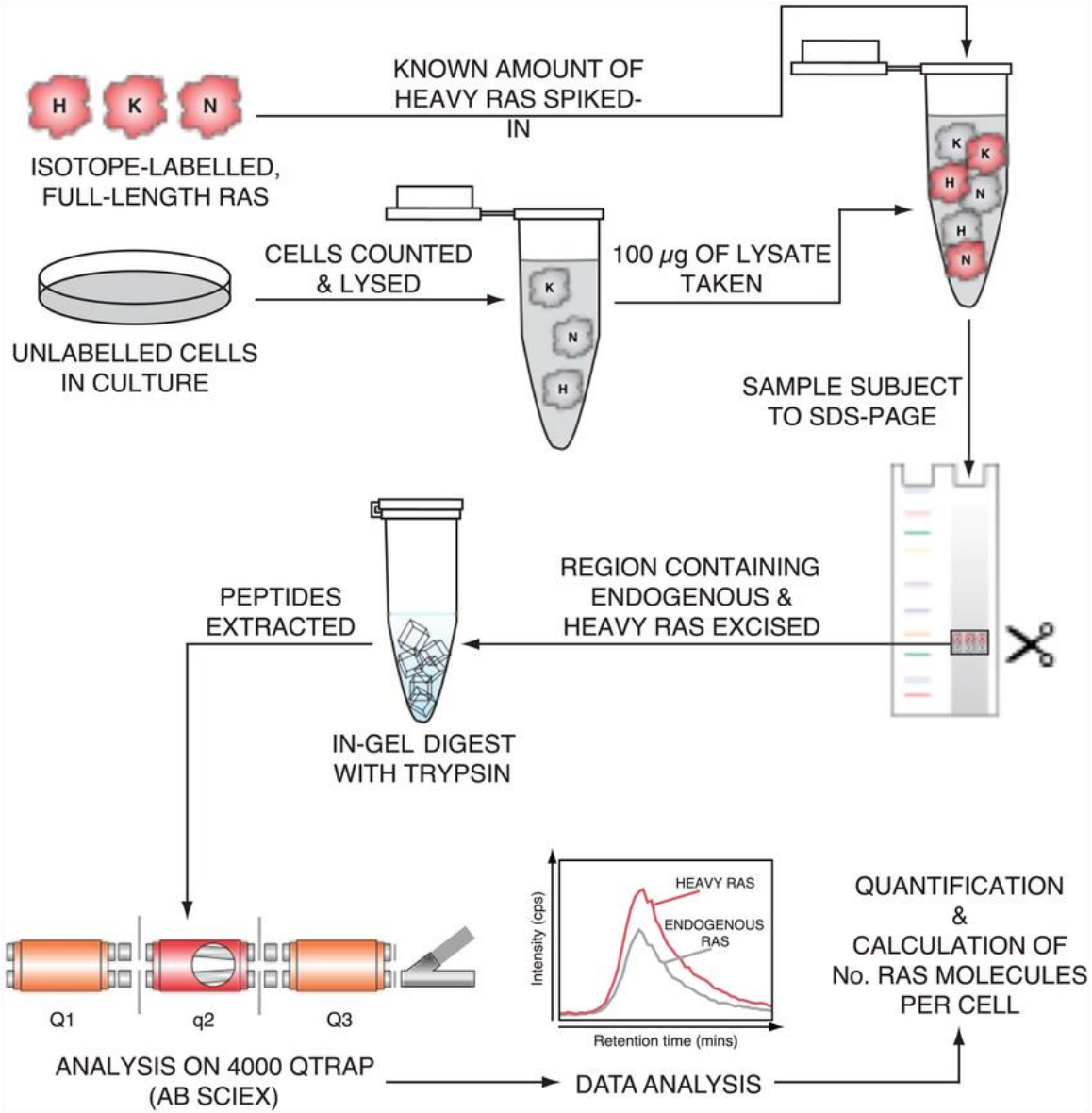
- An AQUA™ Peptide is simply **a synthetic tryptic peptide** corresponding to a peptide of interest. Each AQUA™ peptide incorporates one stable isotope labeled amino acid, creating a slight increase (6-10 daltons) in molecular weight
- Must be proteotypic.
- Definition of proteotypic: a peptide sequence that is found in only a single known protein and therefore serves to identify that protein
- So AQUA probes must be a peptide which is a heavier version (isotopically) of a peptide which is part of your protein of interest.

Question 18

What does PSAQ stand for?

- □ Protein Standard Absolute Quantification
- □ Protein for Stable Accurate Quantification
- □ Peptide Standard for Absolute Quantification
- □ None of these

PSAQ: Protein standard absolute quantification



Question 19

How is abbreviated multiple selected reaction monitoring?

- □ SRM
- □ MRM
- □ mSRM
- □ MSM

Answer: MRM or mSRM

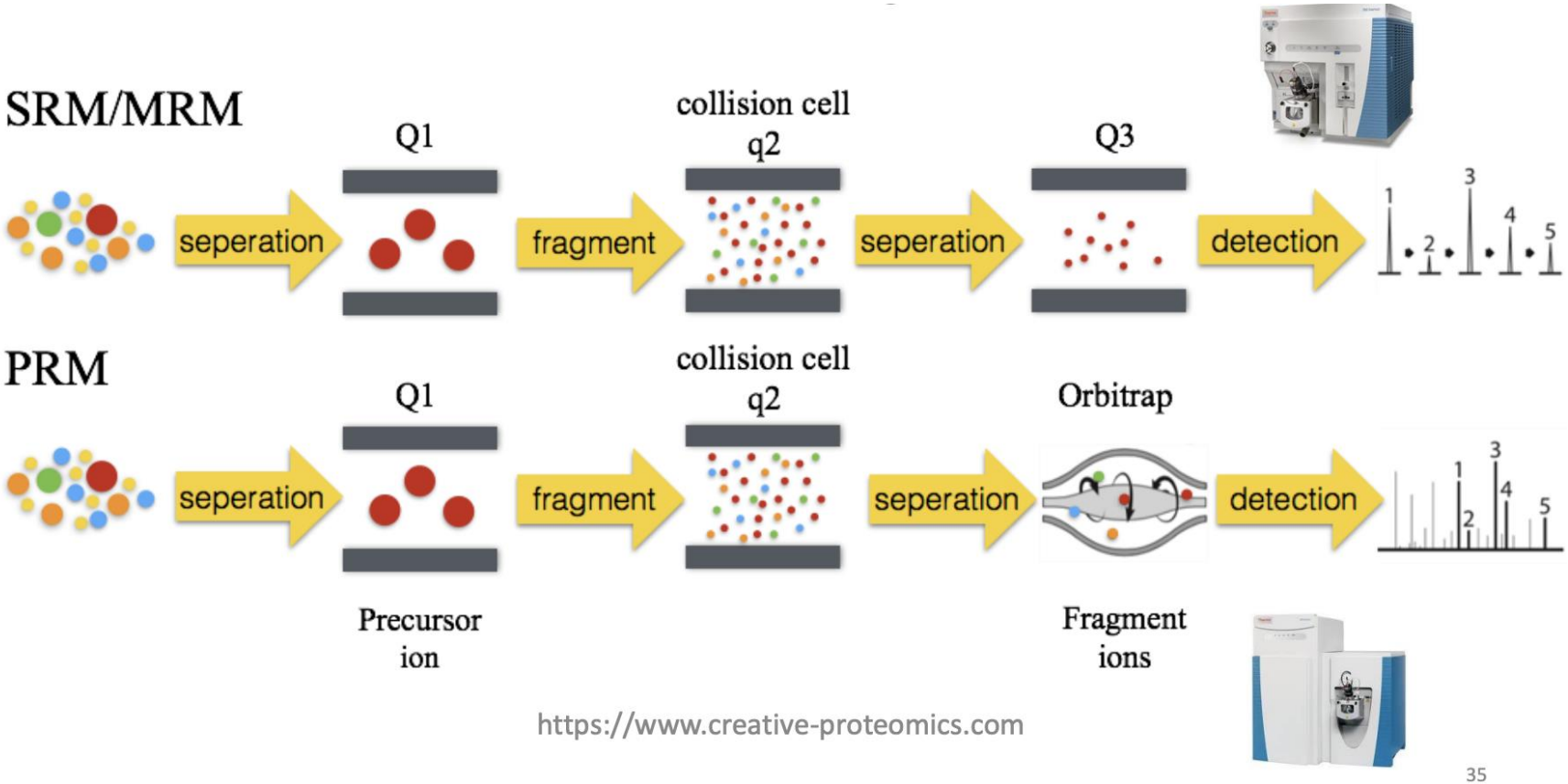
Question 20

What type of mass spectrometer is used for parallel reaction monitoring?

- □ QqQ
- □ FT-ICR
- □ Orbitrap
- □ Ion trap

PRM: Parallel Reaction Monitoring

- We use a q-OrbitTrap instrument



35

Question 21

What does PRM required for its development?

- □ Selection of transitions
- □ Optimization of collision energies
- □ Selection of peptides to be used as heavy standards
- □ An LC system

Parallel Reaction Monitoring (PRM) development

Parallel Reaction Monitoring (PRM)

Peptide selection – Fragmentation – Fragment analysis

Proteins and peptides selection

Transitions selection

Collision Energy optimisation

Sample preparation optimisation

LC-MS method development

Selectivity

Sensitivity

Higher Specificity

Accuracy

No missing values

Method development

Method characteristic

Method development time divided by 3
Simplified workflow
More confident in quantification results

Courtesy of Charlotte Macron

36

Question 22 (Chapter 5 Q1)

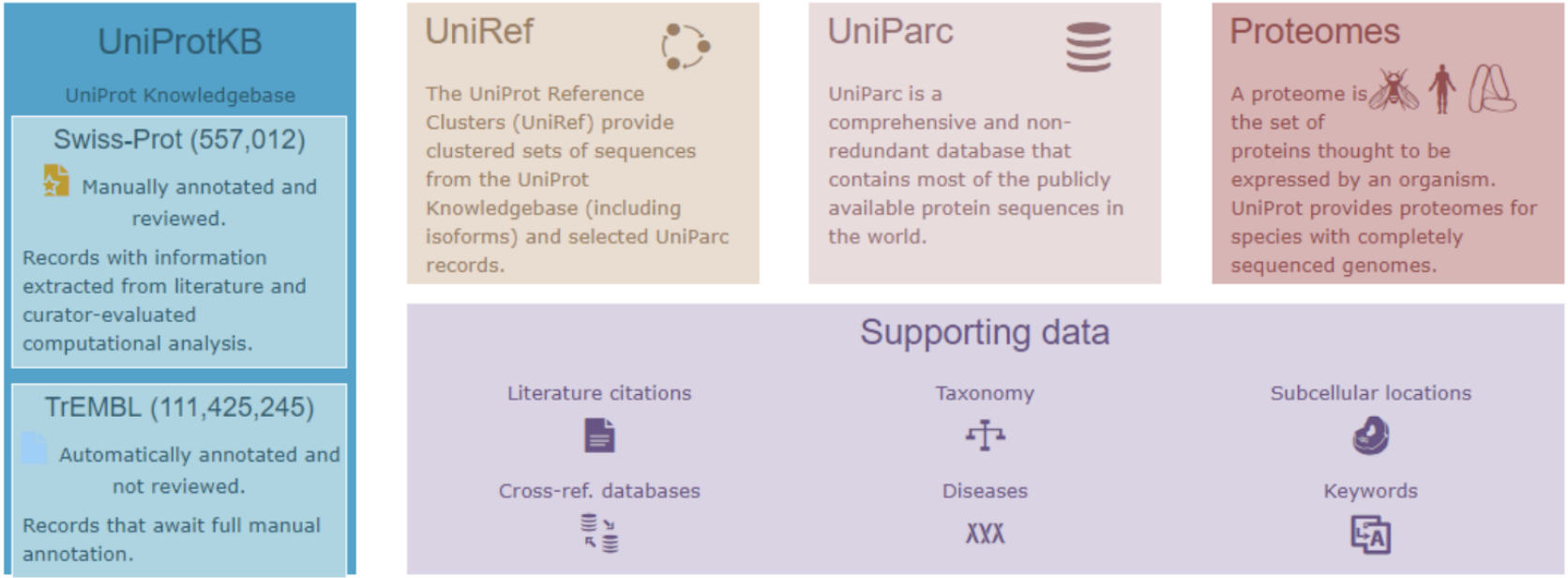
What is UniprotKB?

□ A DNA database
□ A search engine
□ A database of functional information on proteins with accurate, consistent and rich annotation
□ A TV program about proteins on the Discovery Channel

UniprotKB is a database of functional information on proteins

UniProtKB

The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation. In addition to capturing the core data mandatory for each UniProtKB entry (mainly, the amino acid sequence, protein name or description, taxonomic data and citation information), as much annotation information as possible is added.



http://www.uniprot.org/

5

Question 23

What is a protein FASTA file?

□ A file containing the DNA sequences coding for specific proteins
□ A file containing protein sequences
□ A result of protein identifications
□ All functional information available in UniprotKB

Protein FASTA files contain protein sequences

Example of FASTA files

In bioinformatics, FASTA format is a text-based format for representing either nucleotide sequences or peptide sequences, in which nucleotides or amino acids are represented using single-letter codes.

```
>db|UniqueIdentifier|EntryName ProteinName OS=OrganismName OX=OrganismIdentifier [GN=GeneName
]PE=ProteinExistence SV=SequenceVersion
```

```
>sp|Q99497|PARK7_HUMAN Protein DJ-1 OS=Homo sapiens GN=PARK7 PE=1
SV=2
MASKRALVILAKGAEEMETVIPVDVMRRAGIKVTVAGLAGKDPVQCSRDVVICPDASLED
AKKEGPYDVVVLPGGNLGAQNLSESAAVKEILKEQENRKGLIAAICAGPTALLAHEIGFG
SKVTTHPLAKDKMMNGGHYTYSENRVEKDGLILTSRGPGTSFEFALAIVEALNGKEVAAQ
VKAPLVLKD
```


>sp|Q99497|PARK7_HUMAN Protein DJ-1 OS=Homo sapiens GN=PARK7 PE=1
SV=2
MASKRALVILAKGAEEMETVIPVDVMRRAGIKVTVAGLAGKDPVQCSRDVVICPDASLED
AKKEGPYDVVVLPGGNLGAQNLSESAAVKEILKEQENRKGLIAAICAGPTALLAHEIGFG
SKVTTHPLAKDKMMNGGHYTYSENRVEKDGLILTSRGPGTSFEFALAIVEALNGKEVAAQ
VKAPLVLKD

>sp|Q9497|PARK7_HUMAN Protein DJ-1 OS=Homo sapiens GN=PARK7 PE=1
SV=2
MASKRALVILAKGAEEMETVIPVDVMRRAGIKVTVAGLAGKDPVQCSRDVVICPDASLED
AKKEGPYDVVVLPGGNLGAQNLSESAAVKEILKEQENRKGLIAAICAGPTALLAHEIGFG
SKVTTHPLAKDKMMNGGHYTYSENRVEKDGLILTSRGPGTSFEFALAIVEALNGKEVAAQ
VKAPLVLKD
>sp|Q6UWI2|PARM1_HUMAN Prostate androgen-regulated mucin-like
protein 1 OS=Homo sapiens GN=PARM1 PE=1 SV=1
MVYKTLFALCILTAGWRVQSLPTSAPLSVSLPTNIVPPTTIWTSSPQNTDADTASPSNGT
HNNSVLPVTASAPTSLLPKNISIESREEEITSPGSNWEGTNTDPSPSGFSSTSGGVHLTT
TLEEHSSGTPEAGVAATLSQSAAEPPTLISPQAPASSPSSLSTSPPEVFSASVTTNHSST
VTSTQPTGAPTAPESPTEESSSDHTPTSHATAEPVPQEKTPPTTVSGKVMCELIDMETTT
TFPRVIMQEVEHALSSGSIAAITVTVIAVVLLVFGVAAYLKIRHSSYGRLLDDHDYGSWG
NYNNPLYDDS
>sp|Q9H300|PARL_HUMAN Presenilins-associated rhomboid-like
protein, mitochondrial OS=Homo sapiens GN=PARL PE=1 SV=2
MAWRGWAQRGWGCGQAWGASVGGRSCEELTAVLTPPQLLGRRFNFFIQQKCGFRKAPRKV
EPRRSDPGTSGEAYKRSALIPPVEETVFYPSPYPIRSLIKPLFFTVGFTGCAFGSAAIWQ
YESLKSRVQSYFDGIKADWLDSIRPQKEGDFRKEINKWWNNLSDGQRTVTGIIAANVLVF
CLWRVPSLQRTMIRYFTSNPASKVLCSPMLLSTFSHFSLFHMAANMYVLWSFSSSIVNIL
GQEQFMAVYLSAGVISNFVSYVGKVATGRYGPSLGASGAIMTVLAAVCTKIPEGRLAIIF
LPMFTFTAGNALKAIIAMDTAGMILGWKFFDHAAHLGGALFGIWYVTYGHELIWKNREPL
VKIWHEIRTNGPKKGGGSK
>sp|O95453|PARN_HUMAN Poly(A)-specific ribonuclease PARN OS=Homo
sapiens GN=PARN PE=1 SV=1
MEIIRSNFKSNLHKVYQAIEEADFFAIDGEFSGISDGPSVSALTNGFDTPEERYQKLKKH
SMDFLLFQFGLCTFKYDYTDSKYITKSFNFYVFPKPFNRSSPDVKFVCQSSSIDFLASQG
FDFNKVFRNGIPYLNQEEERQLREQYDEKRSQANGAGALSYVSPNTSKCPVTIPEDQKKF
IDQVVEKIEDLLQSEENKNLDLEPCTGFQRKLIYQTLSWKYPKGIHVETLETEKKERYIV
ISKVDEEEERKRREQQKHAKEQEELNDAVGFSRVIHAIANSGKLVIGHNMLLDVMHTVHQF
YCPLPADLSEFKEMTTCVFPRLLDTKLMASTQPFKDIINNTSLAELEKRLKETPFNPPKV
ESAEGFPSYDTASEQLHEAGYDAYITGLCFISMANYLGSFLSPPKIHVSARSKLIEPFFN
KLFLMRVMDIPYLNLEGPDLQPKRDHVLHVTFPKEWKTSDLYQLFSAFGNIQISWIDDTS
AFVSLSQPEQVKIAVNTSKYAESYRIQTYAEYMGRKQEEKQIKRKWTEDSWKEADSKRLN
PQCIPYTLQNHYYRNNSFTAPSTVGKRNLSPSQEEAGLEDGVSGEISDTELEQTDSCAEP
LSEGRKKAKKLKRMKKELSPAGSISKNSPATLFEVPDTW
>sp|P09874|PARP1_HUMAN Poly [ADP-ribose] polymerase 1 OS=Homo
sapiens GN=PARP1 PE=1 SV=4
MAESSDKLYRVEYAKSGRASCKKCSESIPKDSLRMAIMVQSPMFDGKVPHWYHFSCFWKV
GHSIRHPDVEVDGFSELRWDDQQKVKKTAEAGGVTGKGQDGIGSKAEKTLGDFAAEYAKS
NRSTCKGCMEKIEKGQVRLSKKMVDPEKPQLGMIDRWYHPGCFVKNREELGFRPEYSASQ
LKGFSLLATEDKEALKKQLPGVKSEGKRKGDEVDGVDEVAKKKSKKEKDKDSKLEKALKA
QNDLIWNIKDELKKVCSTNDLKELLIFNKQQVPSGESAILDRVADGMVFGALLPCEECSG
QLVFKSDAYYCTGDVTAWTKCMVKTQTPNRKEWVTPKEFREISYLKKLKVKKQDRIFPPE
TSASVAATPPPSTASAPAAVNSSASADKPLSNMKILTLGKLSRNKDEVKAMIEKLGGKLT
GTANKASLCISTKKEVEKMNKKMEEVKEANIRVVSEDFLQDVSASTKSLQELFLAHILSP
WGAEVKAEPVEVVAPRGKSGAALSKKSKGQVKEEGINKSEKRMKLTLKGGAAVDPDSGLE
HSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSWGRVGTVIGSNK
LEQMPSKEDAIEHFMKLYEEKTGNAWHSKNFTKYPKKFYPLEIDYGQDEEAVKKLTVNPG
TKSKLPKPVQDLIKMIFDVESMKKAMVEYEIDLQKMPLGKLSKRQIQAAYSILSEVQQAV
SQGSSDSQILDLSNRFYTLIPHDFGMKKPPLLNNADSVQAKVEMLDNLLDIEVAYSLLRG
GSDDSSKDPIDVNYEKLKTDIKVVDRDSEEAEIIRKYVKNTHATTHNAYDLEVIDIFKIE

Question 24

Is all spectral information used for database search?

□ Yes, directly from the raw spectral files
□ No, raw spectral files are converted into filtered peak lists
□ No, the user defines which spectra he wants to keep for database search
□ None of these

For database search, raw spectra are first converted into filtered peak lists

Conversion of mass spectra into peak lists (2)

Conversion tool

1. General Settings	
Precursor Selection	Use MS1 Precursor
2. Spectrum Properties Filter	
Lower RT Limit	0
Upper RT Limit	0
First Scan	0
Last Scan	0
Lowest Charge State	0
Highest Charge State	0
Min. Precursor Mass	350 Da
Max. Precursor Mass	5000 Da
Total Intensity Threshold	0
Minimum Peak Count	1
3. Scan Event Filters	
Mass Analyzer	Any
MS Order	Is MS2
Activation Type	Any
Min. Collision Energy	0
Max. Collision Energy	1000
Scan Type	Is Full
Ionization Source	Any
Polarity Mode	Any
4. Peak Filters	
S/N Threshold (FT-only)	1.5
5. Replacements for Unrecognized Properties	
Unrecognized Charge Replacements	Automatic
Unrecognized Mass Analyzer Replacements	ITMS
Unrecognized MS Order Replacements	MS2
Unrecognized Activation Type Replacements	CID
Unrecognized Polarity Replacements	+

Peak list

```
BEGIN IONS
TITLE=Spectrum41093 scans:49959,
PEPMASS=314.71024002909854
CHARGE=2+
INSTRUMENT=Default
107.049 814.9
110.07123 25070.0
110.07507 891.1
111.05541 3336.0
112.07589 3249.0
112.08685 868.8
115.08643 758.8
116.07055 1225.0
117.07008 1346.0
120.08083 1692.0
123.11691 1327.0
124.07595 939.9
126.12762 19310.0
127.12467 19000.0
127.13026 1033.0
128.12878 2244.0
128.13458 8418.0
129.07013 812.0
129.13162 5651.0
130.14146 4147.0
131.08568 1146.0
131.13841 1642.0
136.0756 1564.0
138.09106 3231.0
141.10226 1699.0
143.08368 1916.0
145.05014 921.6
146.93361 915.2
149.05997 943.9
150.0914 1632.0
155.11803 3612.0
155.9746 1289.0
156.09131 7085.0
156.11272 942.8
158.09264 1781.0
158.76318 952.6
159.09698 1776.0
159.11717 1217.0
162.06616 1309.0
...
1093.6023 16590.0
1107.6099 64190.0
1108.6106 194100.0
1109.6106 119800.0
1110.6208 171500.0
1111.6179 126700.0
1112.6206 5077.0
1178.5533 1390.0
END IONS

BEGIN IONS
TITLE=Spectrum40214 scans:49000,
PEPMASS=633.4154700290985
CHARGE=2+
INSTRUMENT=Default
101.07073 2256.0
102.05476 2600.0
110.06744 2277.0
110.0713 54330.0
112.07542 7924.0
120.08075 19940.0
121.08384 1454.0
126.12772 244400.0
127.12476 280500.0
128.12831 13070.0
128.13437 205900.0
129.13147 147000.0
129.13731 11300.0
130.06488 1983.0
130.13487 7022.0
130.1411 121000.0
131.08524 1335.0
131.13818 82400.0
131.14351 2647.0
136.07568 19370.0
137.07887 2764.0
138.0912 9590.0
142.08595 8976.0
152.10661 8178.0
156.12227 2307.0
159.09155 14260.0
...
```

Question 25

What entries are necessary for an MS/MS ion search?

- □ A protein database
- □ A file containing tandem mass spectra as a peak list
- □ A chromatographic elution order file
- □ A mass spectrometry method file

To perform MS/MS ion search, you need a protein database and a tandem mass spectra peak list

MATRIX SCIENCE

Search this site

Home | Mascot database search | Products | Technical support | Training | News | Blog | Newsletter | Contact

Access Mascot Server | Database search help

Mascot database search > Access Mascot Server > MS/MS Ions Search

MASCOT MS/MS Ions Search

Your name | Email

Search title

Database(s): contaminants (AA)

Amino acid (AA)
cRAP
SARS-CoV-2
SwissProt
UP186698_X_laevis
UP1940_C_elegans
UP2195_D_discoideum
UP219602_F_oxysporum
UP2311_S_cerevisiae
UP241690_T_harzianum

Taxonomy: All entries

Enzyme: Trypsin | Allow up to 1 missed cleavages

Quantitation: None

Crosslinking: None

Fixed modifications: --- none selected ---

Acetyl (K)
Acetyl (N-term)
Acetyl (Protein N-term)
Amidated (C-term)
Amidated (Protein C-term)
Ammonia-loss (N-term C)
Carbamidomethyl (C)
Carbamidomethyl (N-term)
Carbamyl (K)
Carbamyl (N-term)
Carboxymethyl (C)

Display all modifications

Variable modifications: --- none selected ---

Peptide tol. ±: 1.2 Da | # 13C: 0 | MS/MS tol. ±: 0.6 Da

Peptide charge: 2+ | Monoisotopic | Average

Data file: Choose File | No file chosen

Data format: Mascot generic | Precursor: m/z

Instrument: Default | Error tolerant

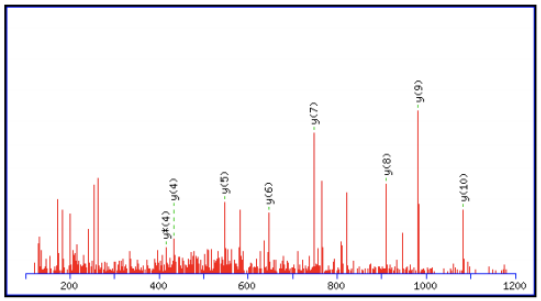
Decoy | Report top: AUTO hits

Start Search ... | Reset Form

MS/MS Fragmentation of VTAYTVDVTGR
Found in Q09666 in SP2017_07_Human, Neuroblast differentiation-associated protein AHNAK OS=Homo sapiens GN=AHNAK PE=1 SV=2

Match to Query 13429: 1180.609228 from(591.311890,2+) index(13392)
Title: Spectrum13394 scans:17257,
Data file File Name: Control_Hela_Helios_03.raw

Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 100 to 1200 Da Full range
Label all possible matches | Label matches used for scoring
Show Y-axis



Monoisotopic mass of neutral peptide Mr(calc): 1180.6088
Fixed modifications: Carbamidomethyl (C) (apply to specified residues or termini only)
Ions Score: 57 Expect: 3.6e-05
Matches : 8/94 fragment ions using 11 most intense peaks (help)

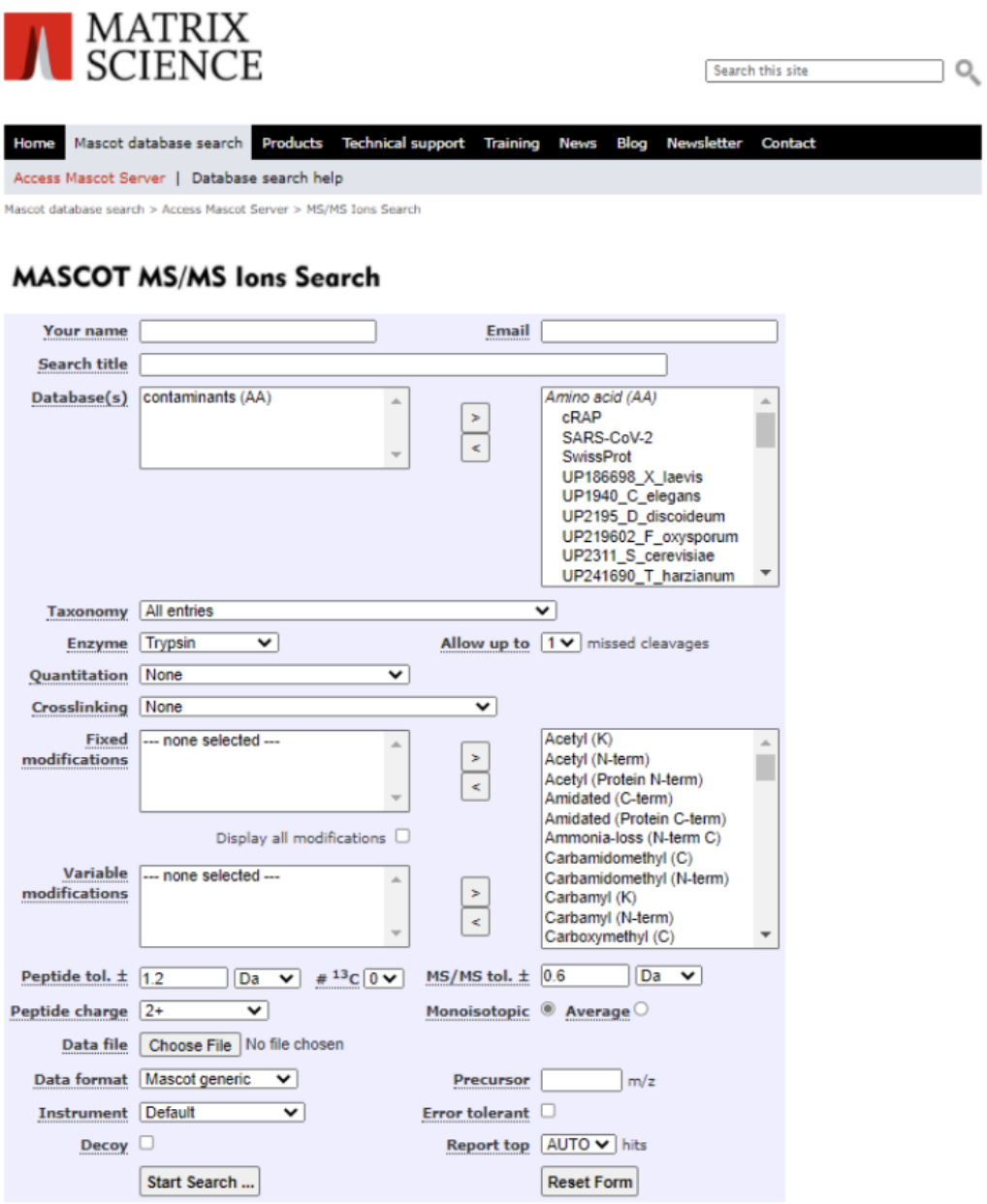
#	b	b^{++}	b^0	b^{0++}	Seq.	y	y^{++}	y^*	y^{*++}	y^0	y^{0++}	#
1	100.0757	50.5415			V							11
2	201.1234	101.0653	183.1128	92.0600	T	*1082.5477*	541.7775	1065.5211	533.2642	1064.5371	532.7722	10
3	272.1605	136.5839	254.1499	127.5786	A	*981.5000*	491.2536	964.4734	482.7404	963.4894	482.2483	9
4	435.2238	218.1155	417.2132	209.1103	Y	*910.4629*	455.7351	893.4363	447.2218	892.4523	446.7298	8
5	536.2715	268.6394	518.2609	259.6341	T	*747.3995*	374.2034	730.3730	365.6901	729.3890	365.1981	7
6	635.3399	318.1736	617.3293	309.1683	V	*646.3519*	323.6796	629.3253	315.1663	628.3413	314.6743	6
7	750.3668	375.6871	732.3563	366.6818	D	*547.2835*	274.1454	530.2569	265.6321	529.2729	265.1401	5
8	849.4353	425.2213	831.4247	416.2160	V	*432.2565*	216.6319	415.2300	208.1186	414.2459	207.6266	4
9	950.4829	475.7451	932.4724	466.7398	T	333.1881	167.0977	316.1615	158.5844	315.1775	158.0924	3
10	1007.5044	504.2558	989.4938	495.2506	G	232.1404	116.5738	215.1139	108.0606			2
11					R	175.1190	88.0631	158.0924	79.5498			1

Question 26

What parameters are usually indicated for the search?

- □ The used enzyme
- □ The mass tolerances for MS and MS/MS levels
- □ The fixed and variable amino acid modifications
- □ The type of chromatography used

What paramaters are indicated for the search?



- The used enzyme
- The mass tolerances for MS and MS/MS levels
- The fixed and variable amino acid modifications

Question 27

What would be the effect to specify too stringent mass tolerances?

Higher mass tolerance means less discrimination

- Remember that we are matching our peptides or MS/MS spectra to a large database of many proteins

- If we **increase the mass tolerance**, it means that our peptide can potentially match with more peptides in the database with the larger range of potential masses (Peptide mass fingerprinting)

- This leads to many more potential matches and will lead to a lowering of scoring of the match target on MASCOT due to the increasing probability that this peptide can also match to other potential proteins in our database

- If we **decrease the mass tolerance**, we are increasing the discrimination, meaning there are less peptides from the database which can fall within this mass error range, leading to more “accurate” matching

- However, one problem with this is if our mass spectrometer is not 100% accurate, and some peptides detected are off slightly in terms of m/z, then we miss out on them. Again, leading to a significant decrease in the overall protein match score

Question 28

Explain the principle of estimating FDR with the target-decoy method?

The target-decoy method and FDR

- Step 1: Concatenate a real protein database and a decoy protein database
- The decoy protein database can be generated by many different ways. One possibility being inverting the reading of all proteins (read from right to left), generating a “false-hit” database
- Step 2: Search your peptides against this concatenated database

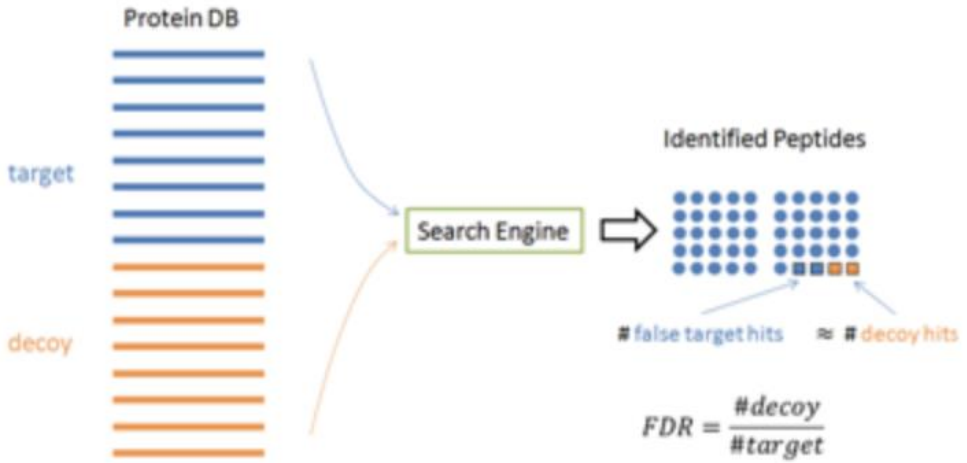


Figure 3: With a properly constructed decoy, the false identifications distribute evenly on the target and decoy. Thus, the amount of decoy hits can be used to estimate the FDR.

The target-decoy method and FDR

- Anything that matches to the decoy database are considered false hits
- With a properly generated decoy, false identifications distribute evenly on the target and decoy. Decoy hits = false hits
- Therefore, false discovery rate FDR = (#decoy hits/ #target hits)

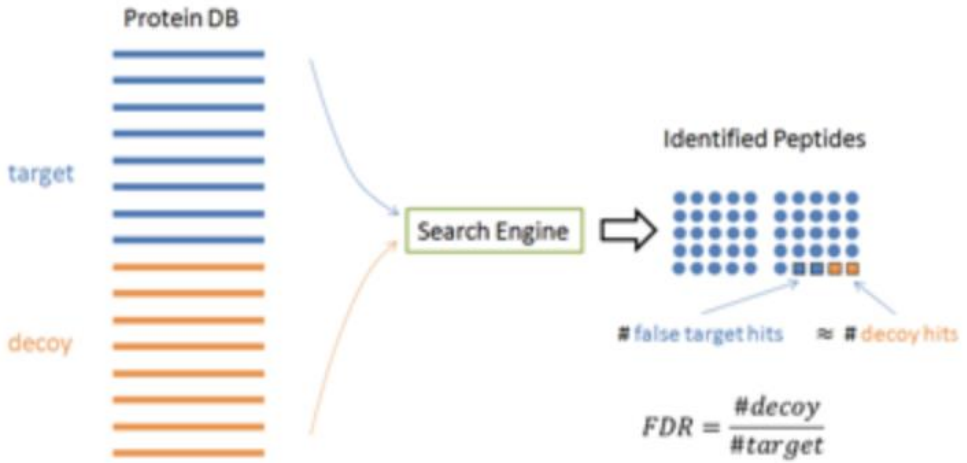


Figure 3: With a properly constructed decoy, the false identifications distribute evenly on the target and decoy. Thus, the amount of decoy hits can be used to estimate the FDR.

Question 29

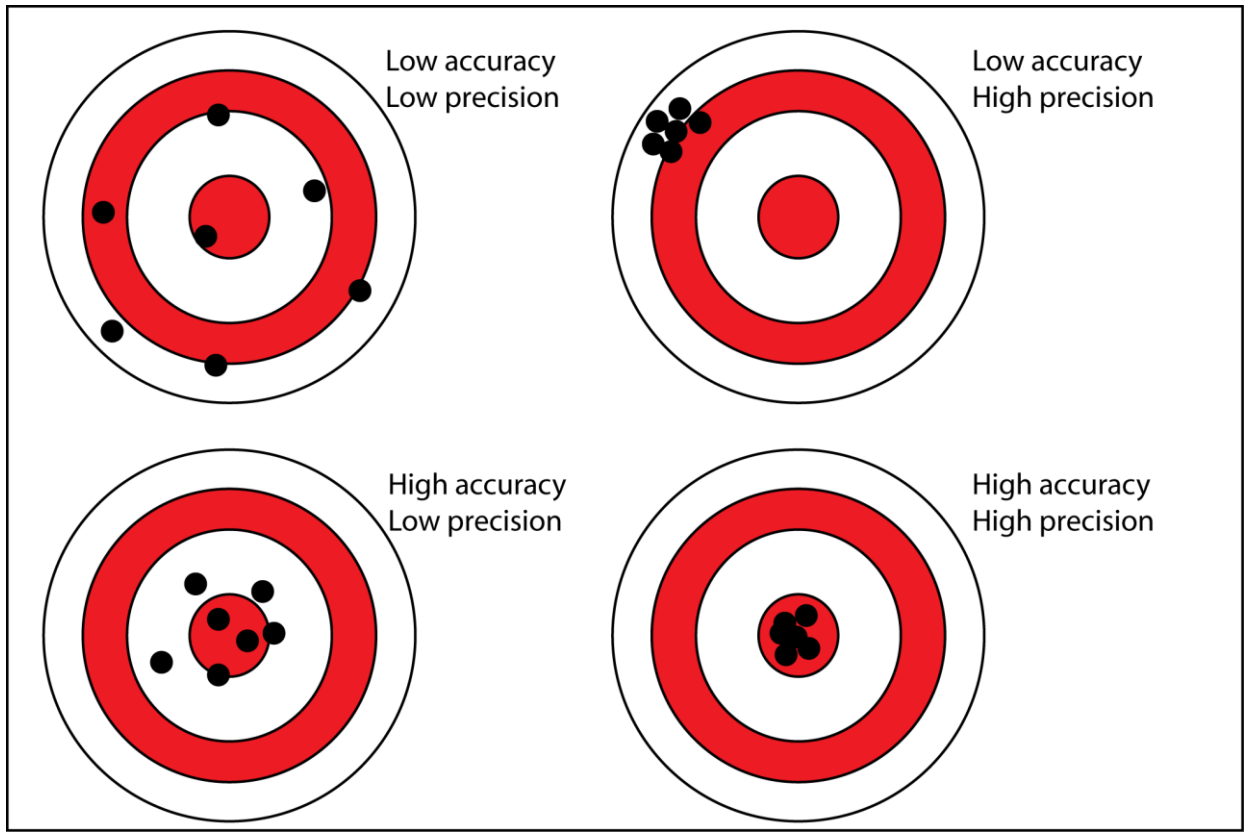
What is the characteristics of the following quantitative measurement?

□ It is imprecise □ It is accurate □ It is precise □ It is inaccurate

Precision versus accuracy

- Precision: a measurement is precise if you are able to get similar values across multiple replicates

- Accuracy: a measurement is accurate if you are able to get data very close to the true value of the analyte



Question 29

What is the characteristics of the following quantitative measurement?

□ It is imprecise □ It is accurate □ It is precise □ It is inaccurate

Question 30

What is the characteristics of the following quantitative measurement?

□ It is imprecise □ It is accurate □ It is precise □ It is inaccurate

Question 31

What is often abbreviated as FC?

- □ Quantitative fold change
- □ False change
- □ Free count
- □ Free cells

Answer: quantitative fold change

Question 32

What is commonly used to express FC?

- □ Log_2
- □ Log_{10}
- □ 10_n

Answer: log_2

Question 33

What does CV stand for?

- □ Coefficient of variation
- □ Coefficient of variability
- □ Change of variance
- □ None of those

Answer: coefficient of variation