

Annotated Sequence	Modifications	# PSMs	1 Master Protein Accessions	2 Positions in Master Proteins	3 Modifications in Master Proteins	Master Protein Descriptions	# Missed Cleavages
[K].GLYGIKDDVFLSVPCILGQNGISDLVK.[V]	1xIAA [C15]; 1xDeamidated [N20]; 1xTrimethyl [K6]	2	P00338	P00338 [279-305]	P00338 1xTrimethyl	L-lactate dehydrogenase	1
[R].LGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLK.[T]	1xIAA [C8]; 1xDeamidated [N28]; 1xOxidation [M27]; 1xDimethyl [K35]	4	P00338	P00338 [178-212]	P00338 1xDimethyl [K212]	L-lactate dehydrogenase A chain	0
[R].LGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLK.[T]	1xIAA [C8]; 1xDeamidated [N28]; 1xOxidation [M27]; 1xTrimethyl [K35]	6	P00338	P00338 [178-212]	P00338 1xTrimethyl [K212]	L-lactate dehydrogenase A chain	0
[R].LGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLK.[T]	1xIAA [C8]; 1xDeamidated [N28]; 1xTrimethyl [K35]	3	P00338	P00338 [178-212]	P00338 1xTrimethyl	L-lactate dehydrogenase	0
[R].LGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLK.[T]	1xIAA [C8]; 1xOxidation [M27]; 1xDimethyl [K35]	4	P00338	P00338 [178-212]	P00338 1xDimethyl	L-lactate dehydrogenase	0
[R].LGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLK.[T]	1xIAA [C8]; 1xOxidation [M27]; 1xTrimethyl [K35]	16	P00338	P00338 [178-212]	P00338 1xTrimethyl	L-lactate dehydrogenase	0
[R].VGAFTMVCKDADEAK.[R]	1xIAA [C8]; 1xOxidation [M6]; 2xDimethyl [K9; K15]	2	P00505	P00505 [288-302]	P00505 2xDimethyl	Aspartate aminotransferase	1
[K].TGQATVASGIPAGWMGLDCGPESKK.[Y]	1xIAA [C19]; 1xOxidation [M15]; 1xTrimethyl [K25]	7	P00558	P00558 [298-323]	P00558 1xTrimethyl	Phosphoglycerate kinase 1	1
[K].DCVGPEVEKACANPAAGSVILLENLR.[F]	2xIAA [C2; C11]; 1xDeamidated [N13]; 1xTrimethyl [K9]	1	P00558	P00558 [98-123]	P00558 1xTrimethyl	Phosphoglycerate kinase 1	1
[K].TGQATVASGIPAGWMGLDCGPESKK.[K]	1xIAA [C19]; 1xOxidation [M15]; 1xDimethyl [K25]	9	P00558	P00558 [298-322]	P00558 1xDimethyl	Phosphoglycerate kinase 1	0
[K].SVWLMSHLGRPDGVPM.PDK.[Y]	1xOxidation [M]; 1xDimethyl [R/K]	5	P00558	P00558 [57-75]	P00558 1xDimethyl [R/K]	Phosphoglycerate kinase 1	0

Step 1: From the **Master Protein Accessions** column retrieve the protein sequence e.g. P00338

```
>sp|P00338|LDHA_HUMAN L-lactate dehydrogenase A chain OS=Homo sapiens OX=9606 GN=LDHA PE=1 SV=2
MATLKDQLIYNLLKEEQTPQNKITVVGVGAVGMACAISILMKDLADELALVDVIEDKLGEMMDLQHGSFLRTPKIVSGK
DYNVTANSKLVIITAGARQQEGESRLNLVQRNVNIFKFIIPNVVKYSPNCKLLIVSNPVDILTYVAWKISGFPKNRVIGSGCNL
DSARFRYLMGERLGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLKTLHPDLGTDKDKEQWKEVHKQVVESAYEVI
KLKGYTSWAIGLSVADLAESIMKNLRRVHPVSTMIKGLYGIKDDVFLSVPCILGQNGISDLVKVTLTSEEEARLKKSADTLWGI
QKELQF
```

Step 2: Map the identified peptides from the **Positions in Master Proteins** on the sequence:

```
MATLKDQLIYNLLKEEQTPQNKITVVGVGAVGMACAISILMKDLADELALVDVIEDKLGEMMDLQHGSFLRTPKIVSGK
DYNVTANSKLVIITAGARQQEGESRLNLVQRNVNIFKFIIPNVVKYSPNCKLLIVSNPVDILTYVAWKISGFPKNRVIGSGCNL
DSARFRYLMGERLGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLKTLHPDLGTDKDKEQWKEVHKQVVESAYEVI
KLKGYTSWAIGLSVADLAESIMKNLRRVHPVSTMIKGLYGIKDDVFLSVPCILGQNGISDLVKVTLTSEEEARLKKSADTLWGI
QKELQF
```

Step 3: Map the identified PTMs from the **Modifications in Master Proteins (only methylations) or Modifications (all PTMs - might be computationally difficult)** on the sequence:

```
MATLKDQLIYNLLKEEQTPQNKITVVGVGAVGMACAISILMKDLADELALVDVIEDKLGEMMDLQHGSFLRTPKIVSGK
DYNVTANSKLVIITAGARQQEGESRLNLVQRNVNIFKFIIPNVVKYSPNCKLLIVSNPVDILTYVAWKISGFPKNRVIGSGCNL
DSARFRYLMGERLGVHPLSCHGWVLGEHGDSSVPVWSGMNVAGVSLKme2/
me3TLHPDLGTDKDKEQWKEVHKQVVESAYEVIKLKGYTSWAIGLSVADLAESIMKNLRRVHPVSTMIKGLYGIKme3DD
VFLSVPCILGQNGISDLVKVTLTSEEEARLKKSADTLWGIQKELQF
```

*\*can we add how many times a methylation event was identified?*

Step 4: Produce an annotated output list for all the sequences

# Output example:

	Modification (and which amino acid is methylated + Probability)	Positions in Master Proteins	Miscleavage
Control Sample Peptide GFLHSGTAKS VTCTYSPALN KMFCQLAKTC		NM_005324 [84-128]	3
Cancer Sample Peptide GFLHSGTAKS VTCTYSPALN KMFCQLAKTC	1xDimethyl [K9(99)] 1xMethyl [K27(100)]	NM_005324 [84-128]	3
	Which Sample Replicate it was identified in	Abundance	PSMs
	Rep 1	2 E10	2
	Rep 3	2.1 E10	
	Rep 1	4 E10	4
	Rep 2	4 E10	
	Rep 3	3.5 E10	

**Ideal Information per site**

**Degree of methylation:** M (Mono), D (Di), Tri (T).

**Probability for PTM site:** usually 99-100% but might not always be provided  
(not essential)

**Positions in Master Proteins:** the residue numbers of the peptide the PTM is on

**Miscleavage:** number

**Which Replicate/Sample it was identified in:** replicate number

**Abundance:** e-value

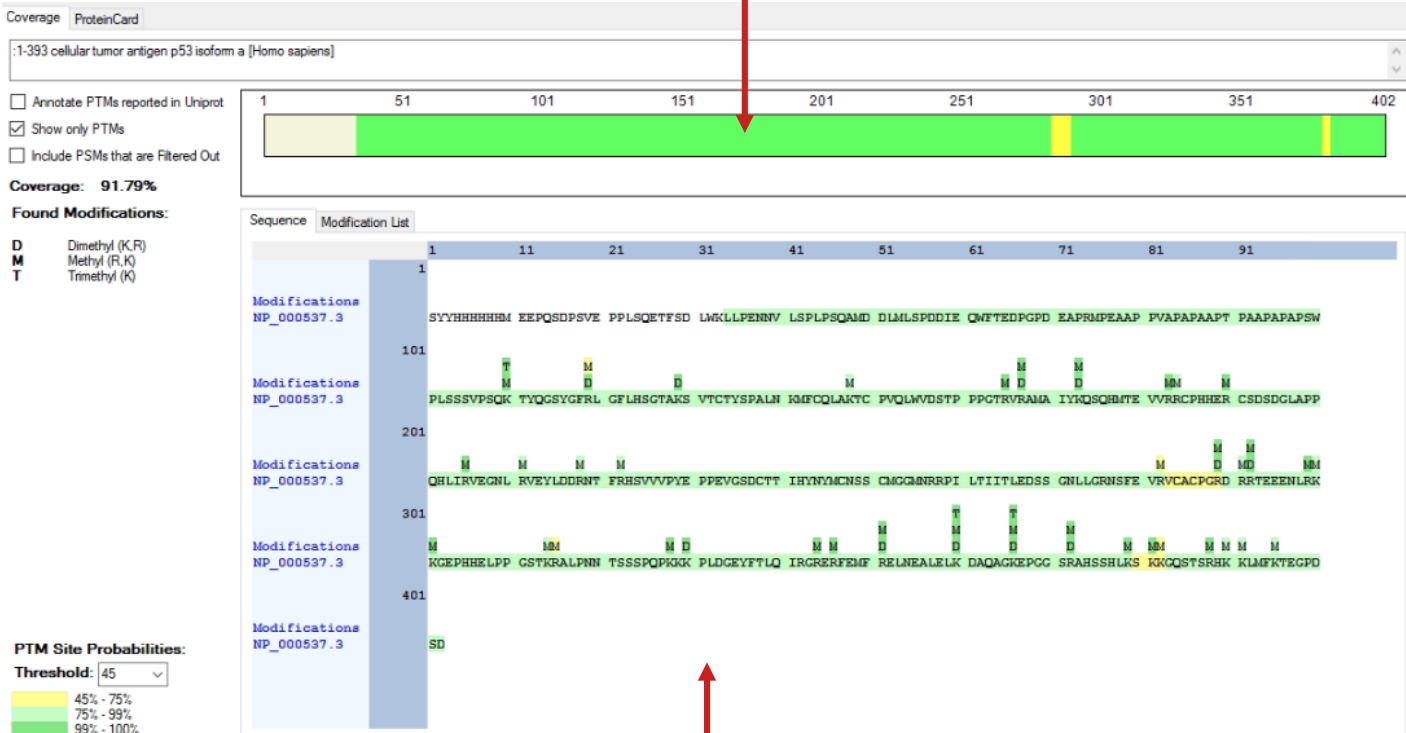
**PSMs:** number

# If you want to attempt visualisation:

## **#1) Protein Coverage:**

→ Highlighted = sequence in the excel list;

→ Grey = sequence not in the excel list



## **#2) PTM Inclusion:**

→ M = mono, D = di, T = tri

→ Highlighted = Methylation found