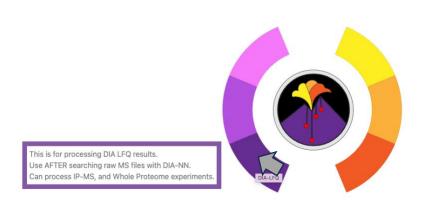
DIA-LFQ (Protein level)

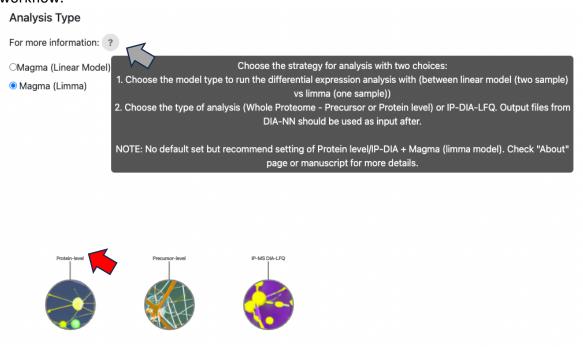
Step 1: Click on the TMT tab on the home page to access all the TMT workflows (as shown in figure below with the grey arrow).



Step 2: Choose the type of analysis to run. You can hover over the question mark to get more details on the different choices and recommended settings (as shown by the grey arrow below). The choices to be made here include:

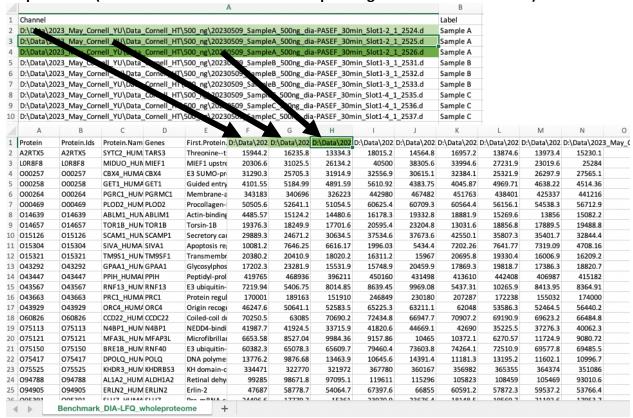
1. The choice of analysis type (MAGMa LM for a two-sample t-test and MAGMa Limma for a one sample t-test)

Click on the Protein-level arm (highlighted with red arrow in the figure below) to access to workflow.



Step 3: Upload the annotation file. This is made using the experimental setup.

An example can be seen below taken from a benchmarking whole proteome (human with yeast spike in) DIA-LFQ experiment. Here each Channel is associated with a separate machine run. So, in the example below, condition "Sample A" (Human: Yeast:: 1:1) was run thrice and thus has three separate files generated which correspond to two columns in the protein level quantification file (highlighted in green). These are headers of the files that would be input in step 5 below (see second screen shot for corresponding color in column header).

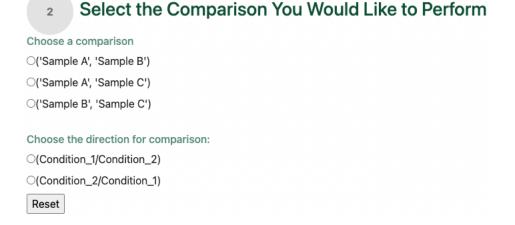


(the pg matrix file is inputted in step 5)

You can also hover over question mark for more details. Click on the "Sample File" in blue to download this example annotation file.



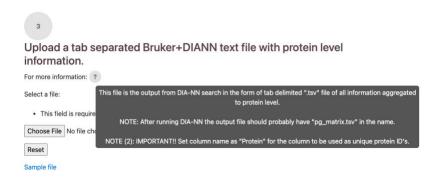
Step 4: Choose the conditions you want to compare in a pairwise fashion. And choose the direction of comparison. So, choosing comparison ('Sample A, 'Sample B') and direction (Condition_1/Condition_2) means the following comparison will run – (Sample A/Sample B).



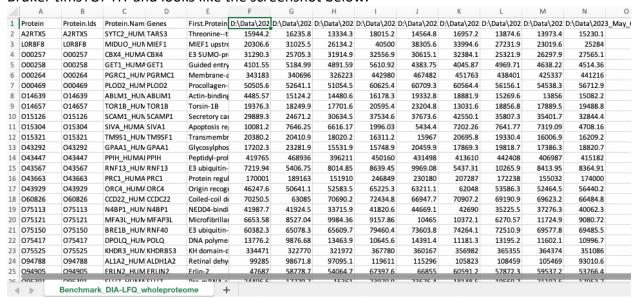
Step 5: Upload the input files (protein level aggregated) for quantification.

For DIA-NN search, upload a tab separated file of quantification file (aggregated to protein level). The file usually has "pg_matrix.tsv" in the name. **Set the column name as "Protein" that you want to use as the unique identifier** (every point on your volcano plot would be associated with a unique value in that column). Make sure the files have the following columns – 'Protein', 'Genes'.

As before hover over question mark for more information. Click on the "Sample File" in blue to download this example protein level aggregated file.

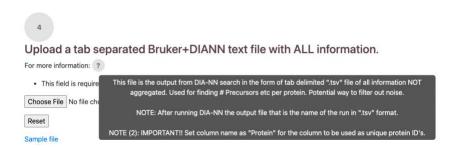


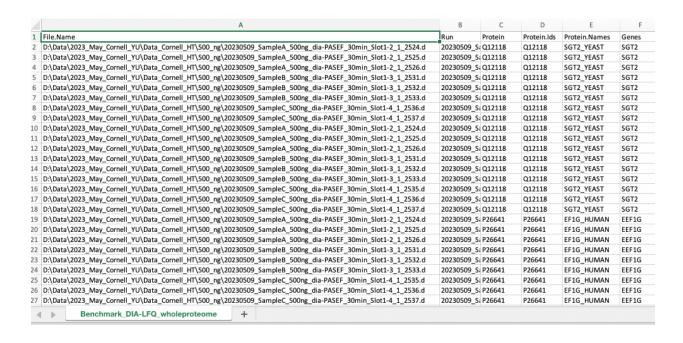
The example is taken from a down-sampling of real benchmarking DIA-LFQ experiment run on Bruker timsTOF HT and looks like the screenshot below.



Step 6: Enter the full file generated after DIA-NN search. This is used to compute the number of precursors identified per label type (specified in the annotation file in step 3). **Here also, set the column name as 'Protein' for the column to be used for unique identifiers (like step 5).**Usually, this file is very large so you can remove extra columns to reduce size. Just make sure the file have the following columns – 'File.Name', 'Run', 'Protein', 'Protein.lds', 'Protein.Names', 'Genes'.

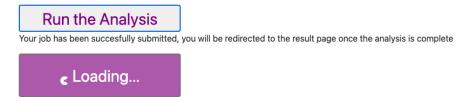
As before hover over question mark for more information. Click on the "Sample File" in blue to download this example file for this example (screenshot of what the file looks like where Protein. Groups set to Protein).





Step 7: Choose whether to do imputation on your data or not. Recommended to not penalize proteins that are not going to be identified/quantified in the Control. In certain situations, this is a given (for example, viral bait would never be endogenously expressed and hence might be completely missing in Control runs depending on type of control being compared to).

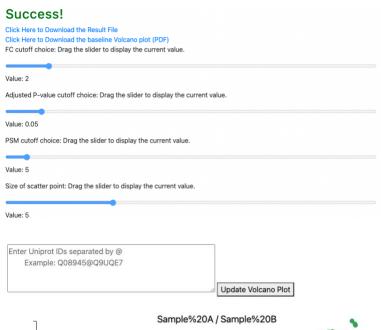
A successful start of run with look like this -

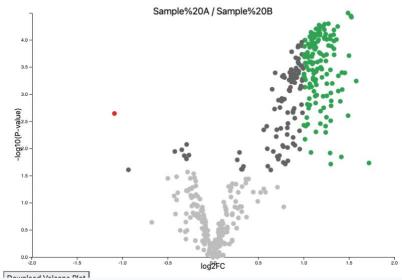


The output page (given a successful run) looks like the screenshot below. You can download the protein level differential expression analysis results as well as the pdf of resulting volcano plot (set at 5 PSM cutoff – where # PSMs is set as the # of Precursors of the numerator condition.

So, if comparison is Sample A vs Sample B then #PSMs is #Precursors for Sample A). You can highlight specific proteins on the volcano plot by entering the associated Uniprot IDs. You can dynamically set the thresholds and download the resulting file, as well as subset the result file with user-defined thresholds. —

_





The output file looks like this -

	A	В	С	D	E	F	G	Н	1	J
L	Protein	log2FC	pval	adjpval	#Precursors-	#Precursors-	#Precursors-	# PSMs	PSM Cutoff	Gene Symbo
2	P39966	1.10159864	0.00018238	0.00131265	37	29	38	37	5	PTC2
3	P00445	1.52273032	3.64E-05	0.00117586	47	45	44	47	5	SOD1
ļ	P25635	0.97145326	0.00070753	0.00233014	33	16	31	33	5	PWP2
,	P13186	0.82488338	0.00218761	0.00559936	23	7	18	23	5	KIN2
5	P53110	1.05599636	0.00181425	0.00481849	8	1	4	8	5	YGL159W
7	P13861	0.03717888	0.59514466	0.65479168	77	81	81	77	5	PRKAR2A
}	Q12502	1.13465163	0.00053816	0.00198398	14	6	13	14	5	LDB19
)	P36104	0.77748475	0.00929662	0.02005472	17	3	16	17	5	SWD2
0	Q9Y5X4	-1.1916566	0.00150702	0.00415904	3	3	3	3	5	NR2E3
1	Q9NYL9	0.14576137	0.14016324	0.20855615	47	59	52	47	5	TMOD3
2	Q9NQL2	-0.6744407	0.22962786	0.31163781	8	12	12	8	5	RRAGD
3	P28791	0.81897296	0.0055468	0.01239873	6	0	4	6	5	SEC20
1	P40357	0.84109681	0.00054833	0.00198547	33	17	34	33	5	SEC9
5	Q9Y597	-0.1153102	0.50408832	0.58592854	12	18	12	12	5	KCTD3
6	Q96MY7	0.26524255	0.0217614	0.04103104	3	3	3	3	5	FAM161B
7	P37838	1.20147305	0.00034735	0.00170026	36	23	36	36	5	NOP4
8	Q2V2M9	-0.2856924	0.32382714	0.41123549	5	6	4	5	5	FHOD3
9	Q96KQ4	-0.2925342	0.06000278	0.10221164	19	21	20	19	5	PPP1R13B
0	P03872	1.08172197	0.01186634	0.02452709	7	2	7	7	5	REP2
1	Q9BXW7	-0.0346588	0.56008565	0.63025584	77	87	91	77	5	HDHD5
2	Q8TF01	0.13667344	0.57201054	0.64075557	9	12	12	9	5	PNISR
3	P40040	1.13585458	0.00013806	0.00117586	15	13	15	15	5	THO1
4	P19880	1.37645951	5.67E-05	0.00117586	25	14	22	25	5	YAP1
5	P35180	1.12828748	0.0005825	0.00203529	17	10	18	17	5	TOM20
c	001000	1 12002771	0.00100764	n nnaa1722	E1	20	AC	E1		DDA 40