

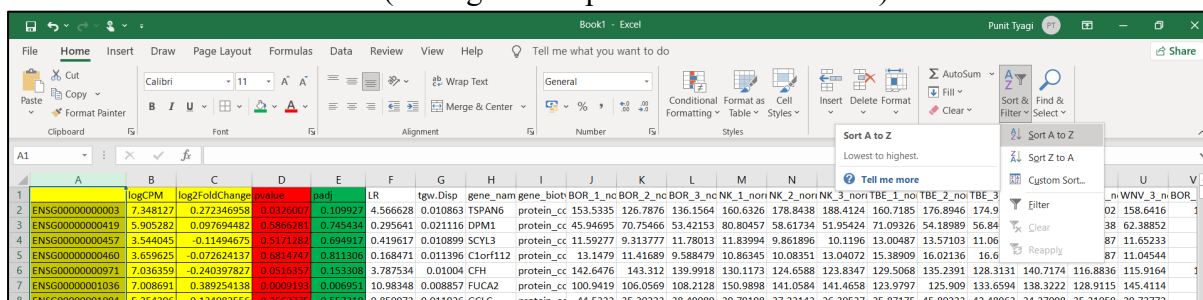
Detailed information on preparation of data to upload to OManalysis

Steps for transcriptomics: -

1. The differential data from the core facility or the commercial vendors will be in edgeR.tsv format.
2. Before going further make copies of edgeR.tsv files in one folder and rename them as per treatments.
3. Open edgeR.tsv file for each treatment in excel. Select and highlight columns belonging to ensemble ID (column A below), logCPM, log2Fold change and pvalue or padj as shown below

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		logCPM	log2FoldChange	pvalue	padj	LR	tgw.Disp	gene_name	gene_biotype	BOR_1_nc	BOR_2_nc	BOR_3_nc	NK_1_nori	NK_2_n
2	ENSG00000000003	7.348127	0.272346958	0.0326007	0.109927	4.566628	0.010863	TSPAN6	protein_cc	153.5335	126.7876	136.1564	160.6326	178.84
3	ENSG000000000419	5.905282	0.097694482	0.5866281	0.745434	0.295641	0.021116	DPM1	protein_cc	45.94695	70.75466	53.42153	80.80457	58.617
4	ENSG000000000457	3.544045	-0.11494675	0.5171282	0.694917	0.419617	0.010899	SCYL3	protein_cc	11.59277	9.313777	11.78013	11.83994	9.8618
5	ENSG000000000460	3.659625	-0.072624137	0.6814747	0.811306	0.168471	0.011396	C1orf112	protein_cc	13.1479	11.41689	9.588479	10.86345	10.083
6	ENSG000000000971	7.036359	-0.240397827	0.0516357	0.153308	3.787534	0.01004	CFH	protein_cc	142.6476	143.312	139.9918	130.1173	124.65
7	ENSG00000001036	7.008691	0.389254138	0.0009193	0.006951	10.98348	0.008857	FUCA2	protein_cc	100.9419	106.0569	108.2128	150.9898	141.05
8	ENSG00000001084	5.254396	-0.134083556	0.3562775	0.557318	0.850972	0.011936	GLCL	protein_cc	44.5332	35.30222	38.49089	39.79198	37.231
9	ENSG00000001167	4.909893	0.11829335	0.553343	0.721251	0.351362	0.024137	NFYA	protein_cc	17.24778	33.95022	27.94357	30.39326	30.472
10	ENSG00000001460	1.37171	0.053013071	0.8851381	0.937965	0.020868	0.040612	STPG1	protein_cc	1.979253	1.051556	3.287478	2.80741	1.4405

4. Delete all other columns.
5. sort column A in A-to-Z (sorting with expanded selection mode).



6. Repeat points 3 to 5 for each treatment edgeR.tsv file.
7. After sorting the information according to column A (yellow), check the first and last ensemble gene ID in all the treatments (it must be the same).



Template.xlsx

8. Open a template excel file () and paste column A (yellow) of first treatment and name the column header as ENSEMBLGENE.
9. Copy in order log2FoldChnage column C (yellow), logCPM column B and pvalue column D (red) or padj column E (green) from each treatment excel file.
10. Paste the copied columns next to the ENSEMBLGENE in template excel file.
11. **The master file must look like the below screenshot.** Yellow color columns belong to treatment 1, green color columns belong to treatment 2, red color columns belong to treatment 3 and blue color columns belong to treatment 4.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ENSEMBLGENE	logFC	logCPM	PValue	logFC	logCPM	PValue	logFC	logCPM	PValue	logFC	logCPM	PValue
2	ENSG00000000003	0.561283	4.557846	0.438597	-0.90616	4.557846	0.141092	-0.11978	4.557846	0.880781	-0.35996	4.557846	0.640167
3	ENSG000000000419	0.602523	4.758842	0.402304	0.096469	4.758842	0.902259	0.001348	4.758842	0.997336	-0.78139	4.758842	0.254466
4	ENSG000000000460	0.826338	1.782839	0.256336	-0.13903	1.782839	0.8607	0.076887	1.782839	0.923429	0.792963	1.782839	0.277646
5	ENSG000000000971	0.780918	6.933399	0.223946	-0.32462	6.933399	0.676356	-0.21808	6.933399	0.782282	-0.62059	6.933399	0.380078
6	ENSG000000001036	0.869587	5.462408	0.151536	-0.35845	5.462408	0.642259	-0.29408	5.462408	0.7065	0.171774	5.462408	0.827248
7	ENSG000000001084	0.429974	4.100634	0.56991	-0.62329	4.100634	0.381325	0.636583	4.100634	0.36521	-0.01379	4.100634	0.986338
8	ENSG000000001167	0.413473	3.653166	0.587001	0.34253	3.653166	0.656407	0.252641	3.653166	0.746001	0.178518	3.653166	0.820939
9	ENSG000000001461	-1.51866	5.259942	0.000472	-1.37751	5.259942	0.002377	-0.57455	5.259942	0.426293	-1.1324	5.259942	0.0298
10	ENSG000000001497	0.34803	4.674989	0.652222	0.013774	4.674989	0.985861	-0.6163	4.674989	0.385452	-0.54674	4.674989	0.454072

- Column headers and the order of the columns (logFC, logCPM, Pvalue) must be the same as shown in the above screenshot.
- The template file of transcriptomics **must be saved as CSV file** to upload to OManalysis.

Steps for Proteomics: -

- The output from the differential proteomics analysis pipeline for each treatment is in a separate xlsx file.
- Open each treatment file in excel.
- Each treatment file must contain the column headers, UniProt ID, FDR-adjusted P-value and Fold Change in the same order presented in the below screenshot.

	A	B	C
1	UniProt ID	FDR-adjusted P-value	Fold Change
2	Q8SPP7	4.5E-10	3304.8
3	P54229	1.9E-08	1443.5
4	P56425	1.6E-06	1217.0
5	P22226	2.8E-08	1026.2
6	Q2TBU0	3.8E-08	996.8
7	F1N465	1.5E-03	527.2
8	E1BCU6	1.5E-06	401.1
9	Q9TU03	1.6E-04	312.8
10	P52176	1.1E-04	219.2

- Open a new .xlsx file (master file) and name the first excel sheet as Treatment1 (Screenshot below)

	A	B	C	D	E	F	G
1	UniProt ID	FDR-adjusted P-value	Fold Change				
2	Q8SPP7	4.5E-10	3304.8				
3	P54229	1.9E-08	1443.5				
4	P56425	1.6E-06	1217.0				
5	P22226	2.8E-08	1026.2				
6	Q2TBU0	3.8E-08	996.8				
7	F1N465	1.5E-03	527.2				
8	E1BCU6	1.5E-06	401.1				
9	Q9TU03	1.6E-04	312.8				
10	P52176	1.1E-04	219.2				

- Paste the columns from each treatment into the master file with a new excel sheet and naming them as Treatment1 for the first treatment, Treatment2 for the second treatment, Treatment3 for the third treatment, Treatment4 for the fourth treatment.
- Save the master file in .xlsx format.**