

Summary statistics of post-processed MaxQuant output for internal QC.

Summary

PRIDE dataset identifier:	PXD001608
PRIDE dataset URL:	https://www.ebi.ac.uk/pride/archive/projects/PXD001608
PubMed ID:	25993694
Quantification method:	Label-free (baseline)
Search database:	Human Reference Proteome (UniProt, May 2019. 95,915 sequences)
Contaminant database:	MaxQuant contaminants database (conf/contaminants.fasta)
Analysis software:	MaxQuant v1.6.3.4
Operating system:	Red Hat Enterprise Linux Server

Table 1. Summary

	Pre-processed	Post-processed
Number of samples	10	10
Number of protein groups	6081	5043
Number of reverse decoys [^]	93	0
Number of contaminants [^]	45	0
Total number of proteins [§]	30088	25038
...SwissProt	13717	11351
...Trembl	16259	13651
Total number of non-isoforms ^a	22937	19008
...SwissProt.	6566	5321
...Trembl.	16259	13651
Total number of peptides [¶]	74046	65527
Total number of unique peptides [†]	57199	52349
Protein groups with 2 or more mapped peptides [*]	5647	5043
Protein groups mapped to unique gene [•]	NA	5378

[^] Some protein groups containing reverse decoys and/or contaminant entries present within are not marked as “+” in “Potential.contaminant” or “Reverse” fields in the output and therefore the total number of contaminants and/or reverse decoys counted here may be less than those present within the protein groups.

[§] The total number of proteins present across all protein groups. Number of SwissProt and Trembl entries within are counted. Total number of proteins in preprocessed data includes contaminants, reverse decoys, SwissProt and Trembl proteins. In post-processed data, total number of proteins includes only SwissProt and Trembl entries.

^a Total number of proteins excluding isoforms present across all protein groups. Total number of non-isoform SwissProt and Trembl entries within.

[¶] Sum of peptides that are mapped across all protein groups.

[†] The total number of unique peptides associated with the protein group (i.e. these peptides are not shared

with another protein group).

* Proteins within protein groups to which 2 or more peptides are mapped to.

- Proteins within protein groups which are mapped to an unique Ensembl Gene ID.

Pre-processed data

Output from MaxQuant without any downstream processing

Post-processed data

Processed MaxQuant output

Filters applied:

- Remove reverse decoys
- Remove contaminants
- Include protein groups that have 2 or more peptides mapped to protein
- Include protein groups wherein all protein IDs in the protein group are mapped to an unique Ensembl Gene ID

Normalisation method: Fraction Of Total.

In this normalisation method each protein iBAQ intensity value is scaled to the total amount of signal in a given MS run (column) and transformed to parts per billion (ppb)

Plots

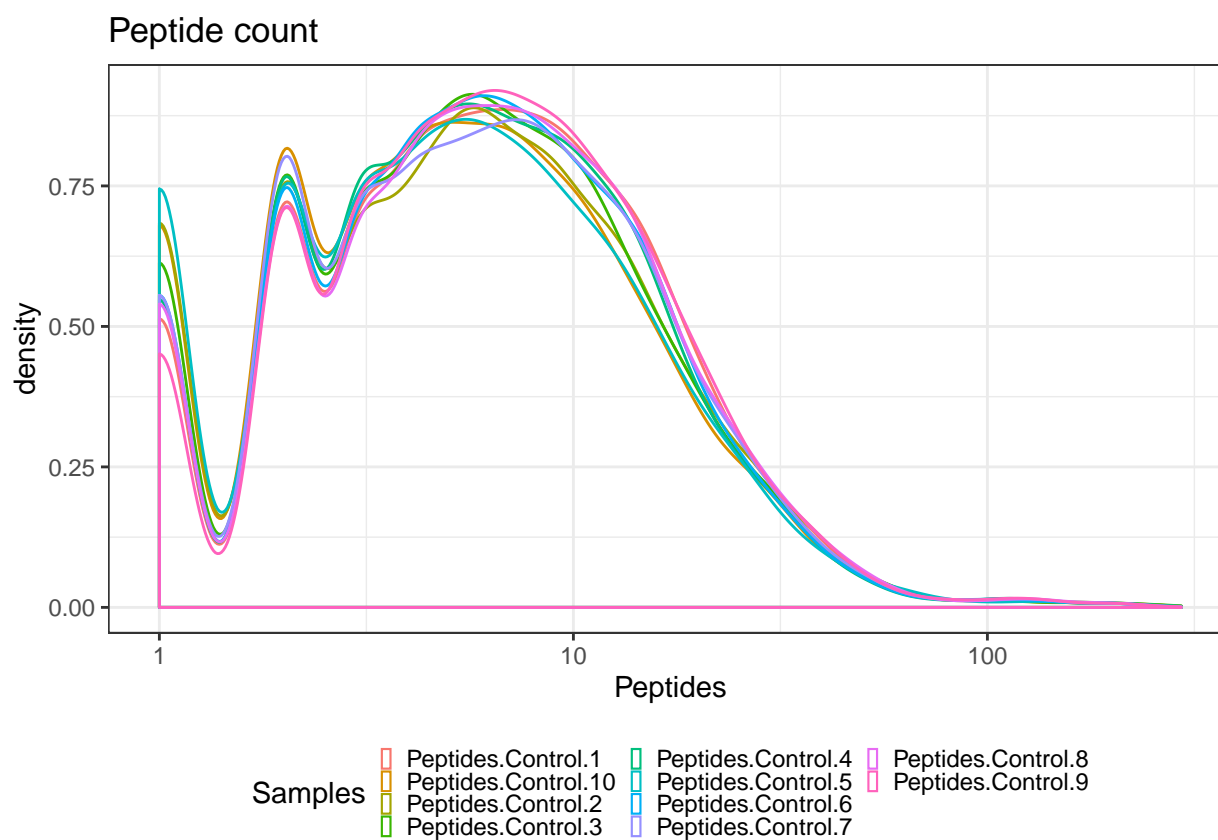


Figure 1. Peptide density.

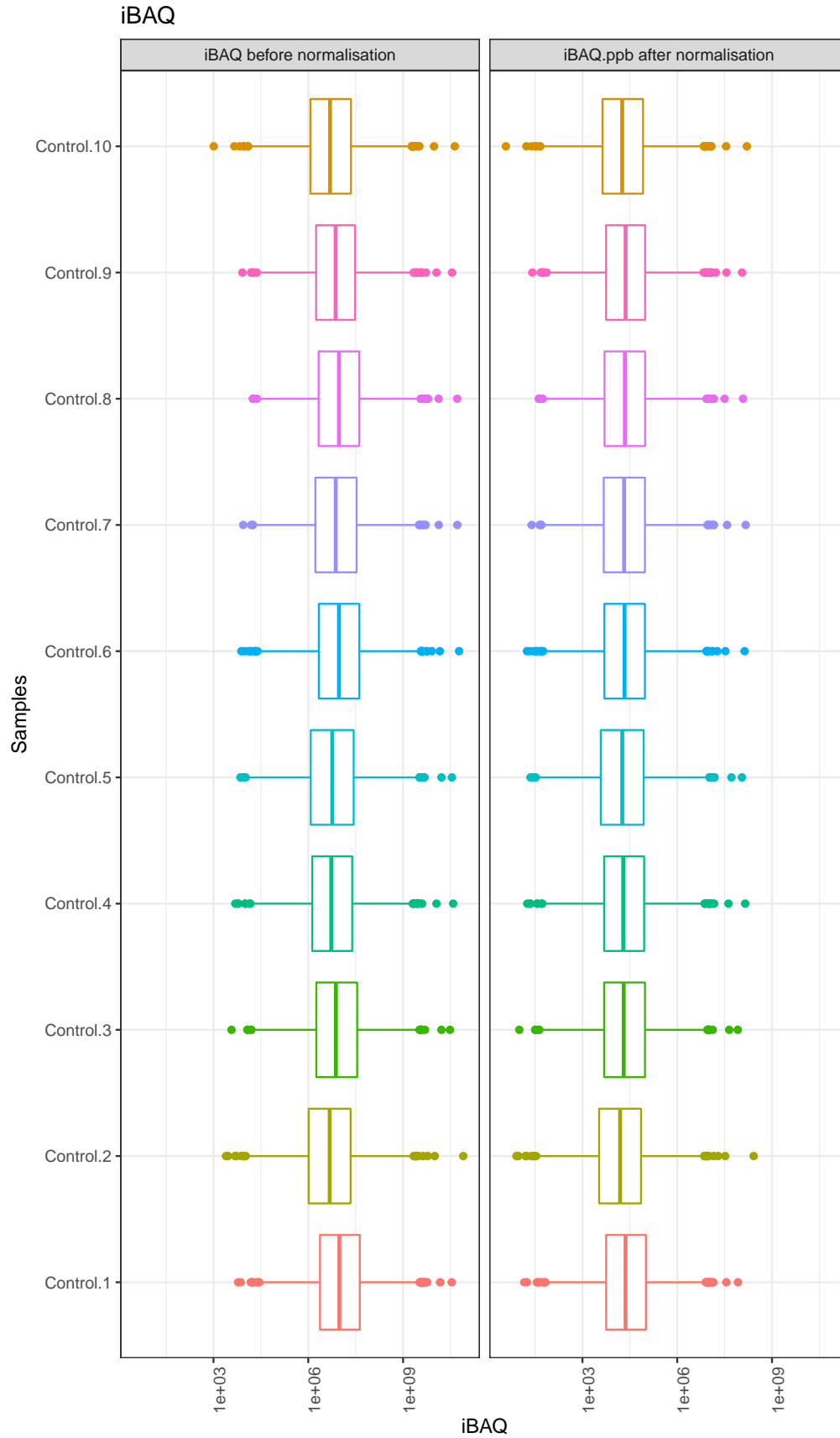


Figure 2. Boxplots

with distribution of iBAQ values for each sample before and after FOT normalisation.

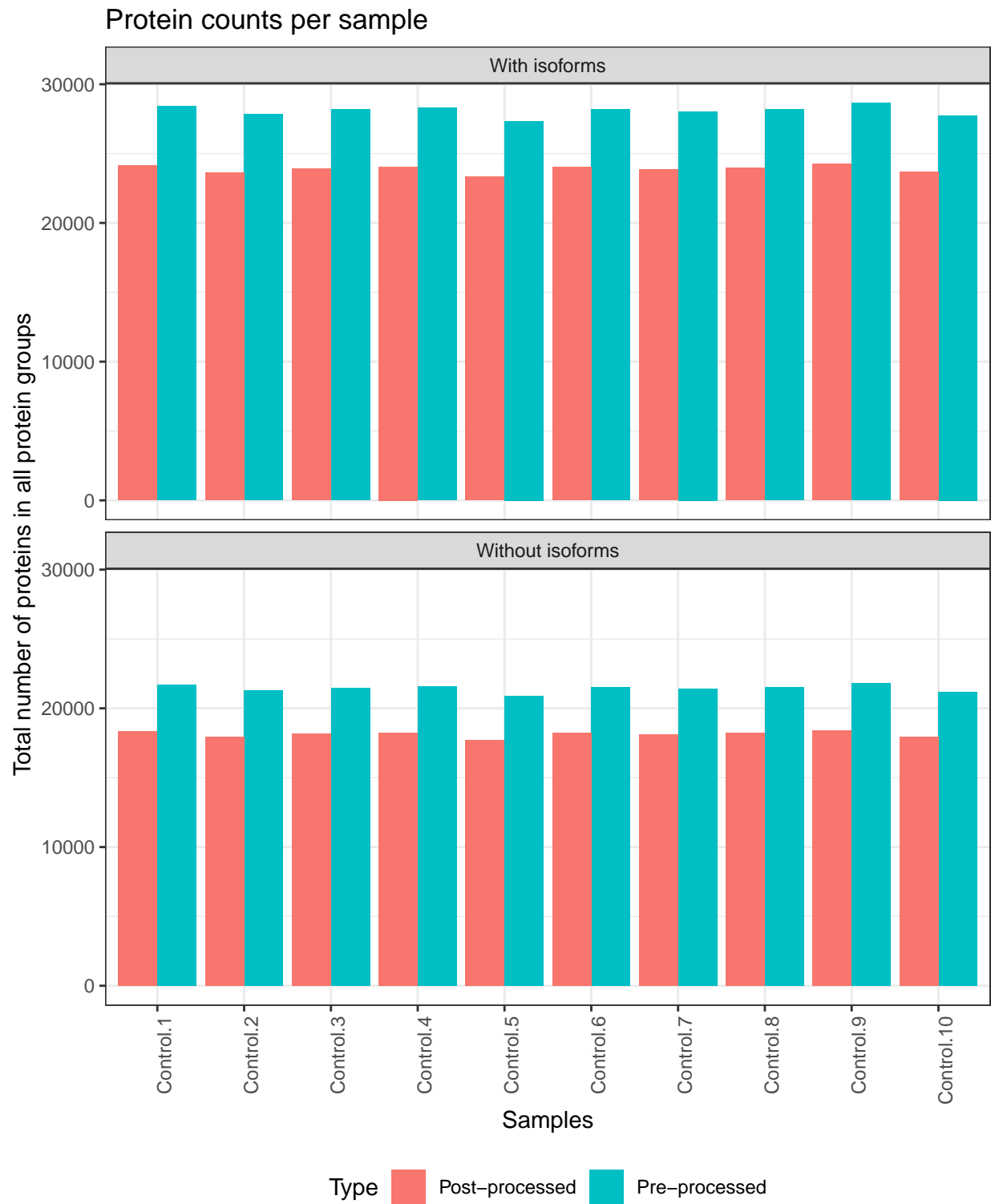


Figure 3. Protein count plot

Table 2. Protein counts.

Total number of proteins before and after downstream processing

	Samples	Pre. ALL§	Post. ALL¶	Pre. non-iso.†	Post. non-iso.€
1	Control.1	28433	24125	21699	18340
2	Control.10	27759	23662	21176	17966
3	Control.2	27844	23634	21272	17966
4	Control.3	28173	23920	21486	18148
5	Control.4	28287	24044	21602	18253
6	Control.5	27352	23337	20895	17741
7	Control.6	28187	24015	21516	18251
8	Control.7	28030	23845	21395	18108
9	Control.8	28187	23968	21521	18223
10	Control.9	28633	24258	21815	18410

§ Total number of proteins (including isoforms) in pre-processed samples

¶ Total number of proteins (including isoforms) in post-processed samples

† Total number of proteins (without isoforms) in pre-processed samples

€ Total number of proteins (without isoforms) in post-processed samples

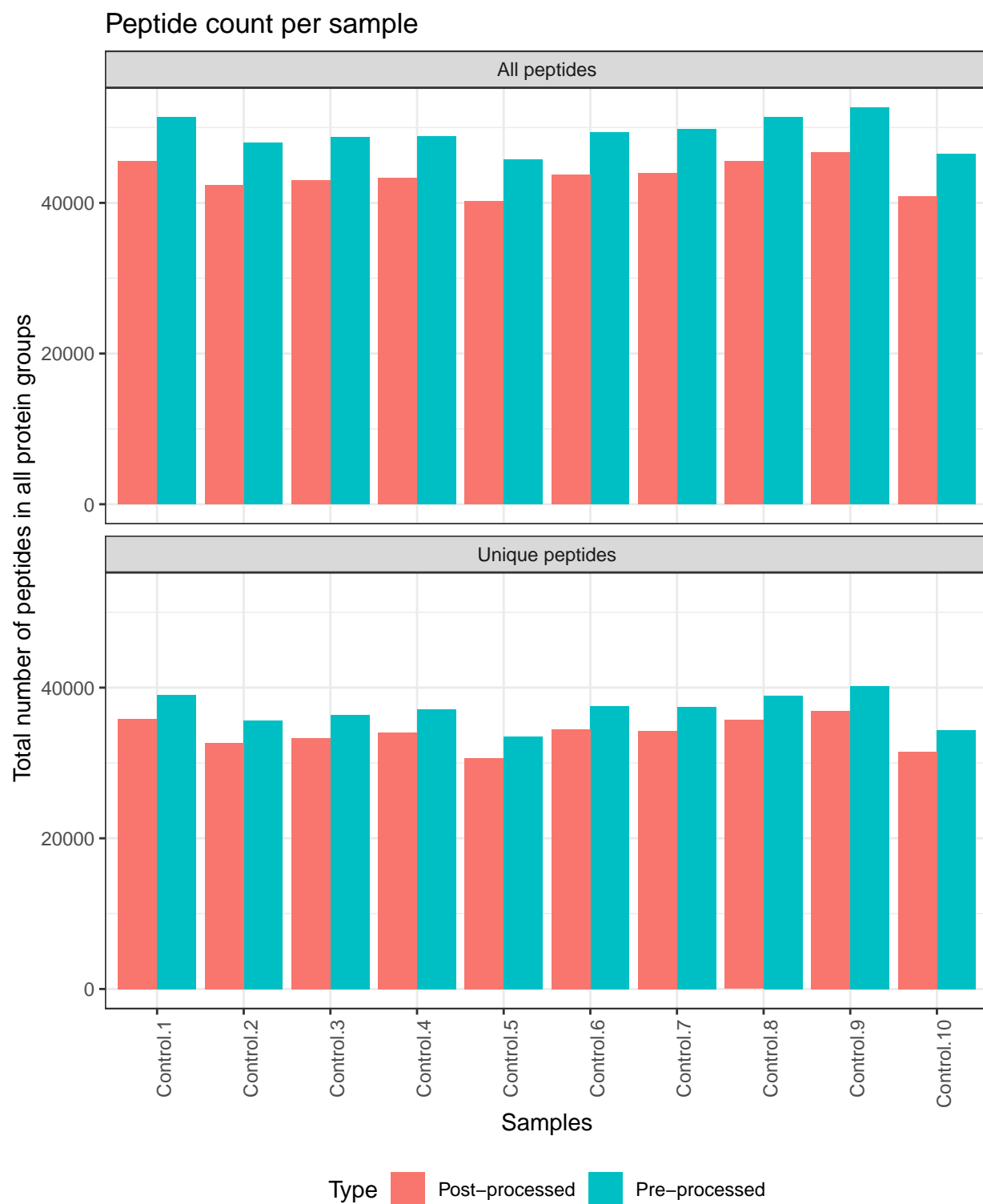


Figure 4. Peptide count plot

Table 3. Peptide counts.

Total number of all peptides before and after downstream processing

	Samples	Pre. ALL.§	Post. ALL.¶	Pre. unique.†	Post. unique.€
1	Control.1	51335	45530	39078	35841
2	Control.10	46466	40903	34391	31467
3	Control.2	48029	42336	35651	32623
4	Control.3	48706	43029	36403	33340
5	Control.4	48801	43327	37138	34085
6	Control.5	45728	40178	33499	30634
7	Control.6	49379	43774	37557	34439
8	Control.7	49760	43951	37458	34264
9	Control.8	51414	45526	38940	35703
10	Control.9	52649	46666	40234	36911

§ Total number of all peptides in pre-processed samples

¶ Total number of peptides in post-processed samples

† Total number of unique peptides in pre-processed samples

€ Total number of unique peptides in post-processed samples

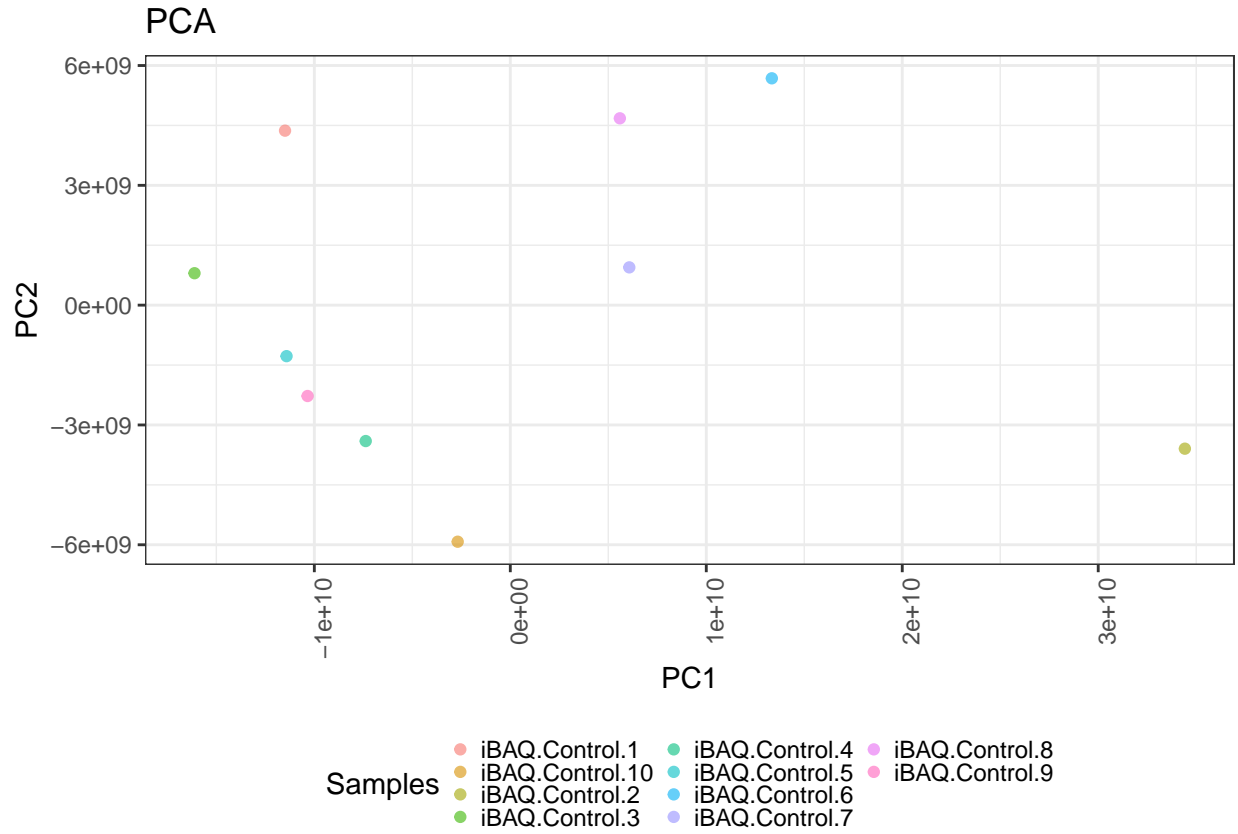


Figure 5. PCA plot. FOT normalised iBAQ values were used.