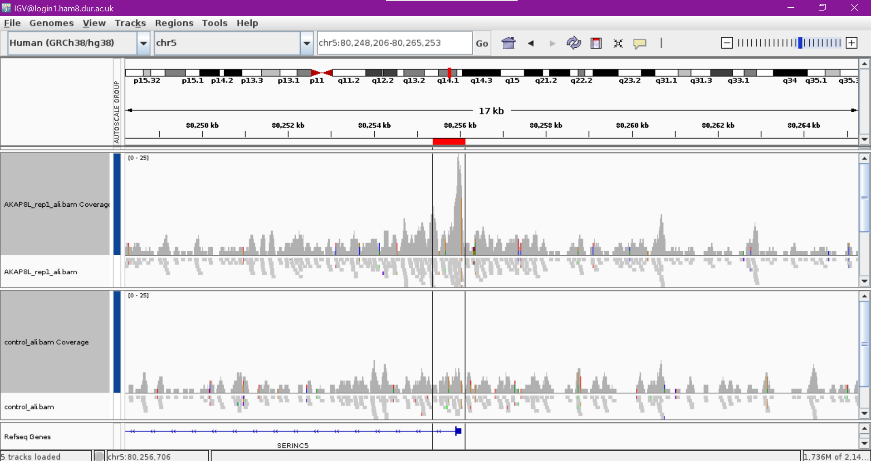
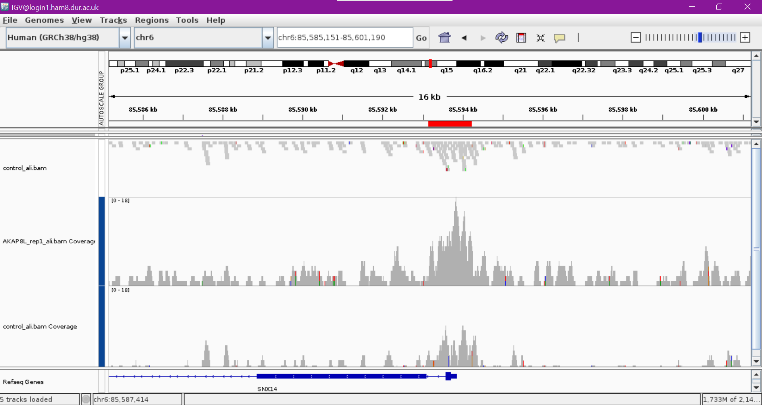
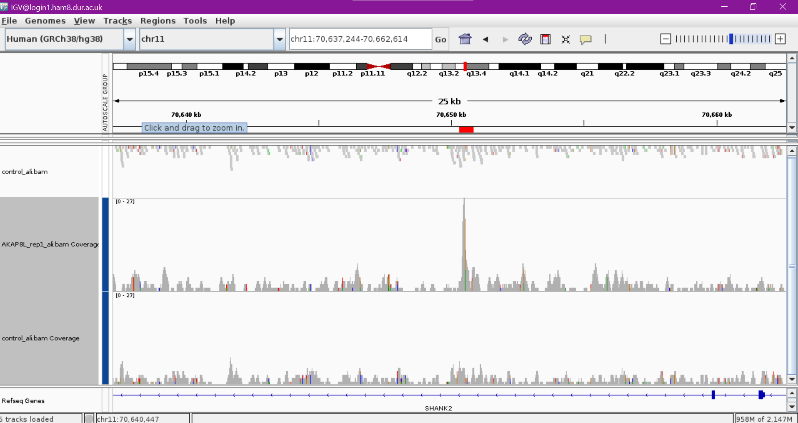
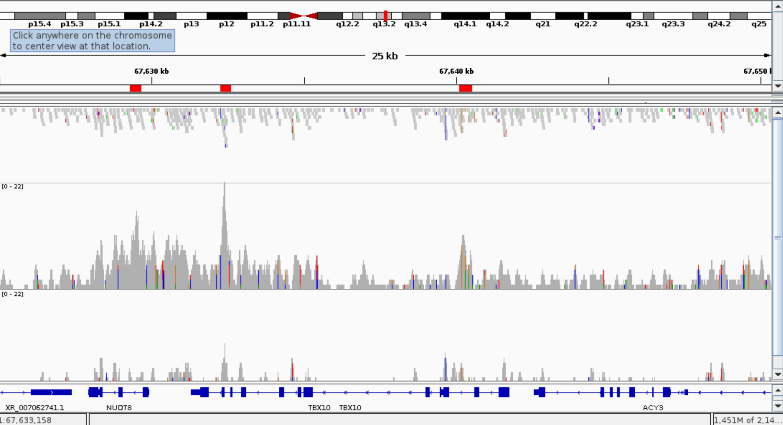
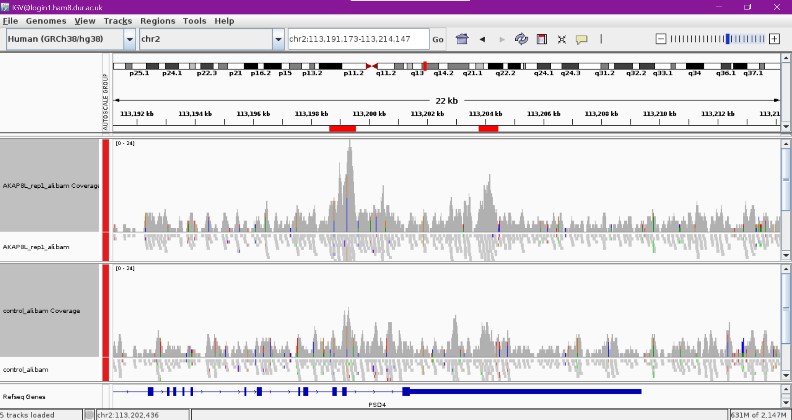
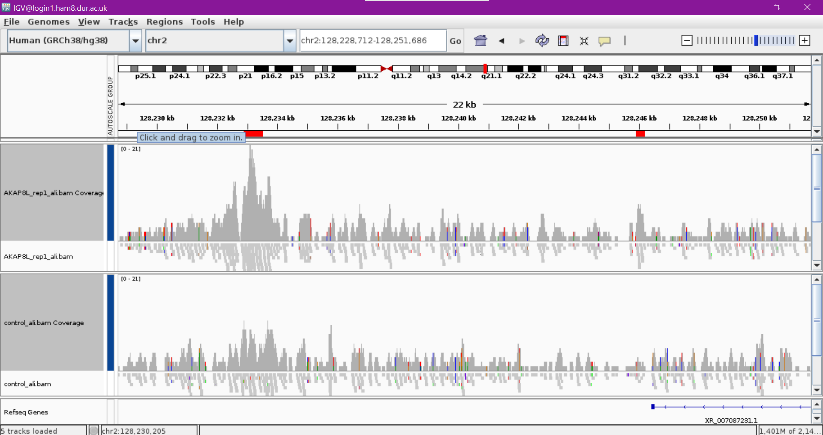
Pics of stages of pipeline not seen?  
> in order each command line output (each m&m stage)

* Training graphs/tensorboard/optimisation (per generation)
* Additional background math
* Figures of how model works/additional diagrams
* Datasets used in table format
* Check supplementary data from lanceotron/cnn-peaks/anka

Labelling AKAP8L

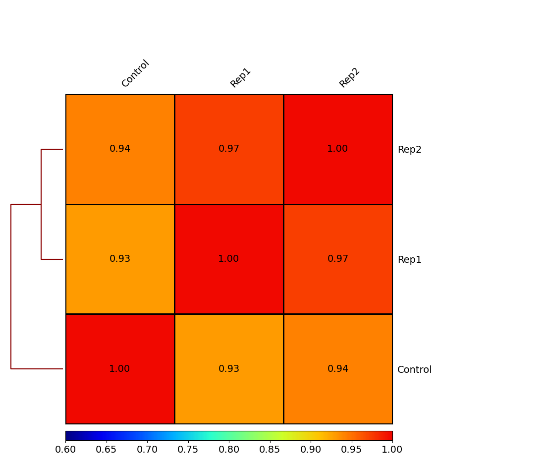




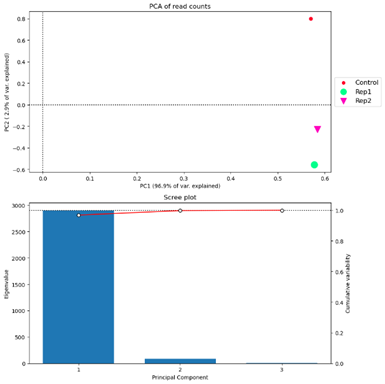
SUPPLEMENTARY

1. Read coverage

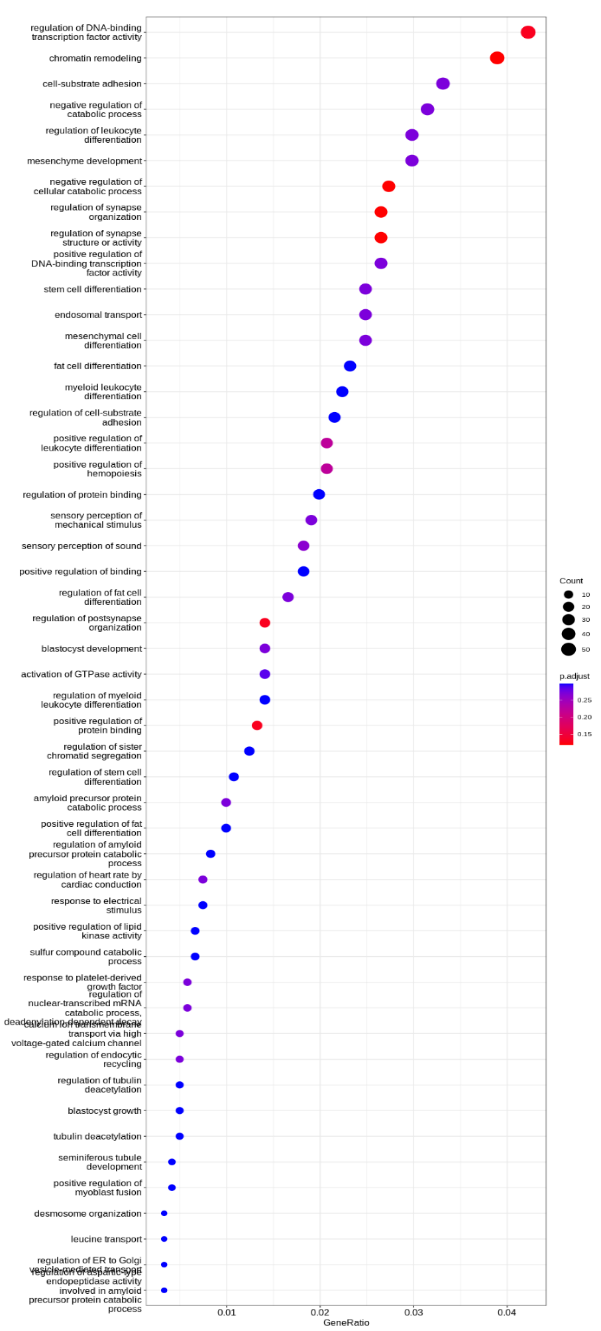
1a) Heatmap plotCorrelation used to examine read coverage similarity to perform heirarchial clustering (clustering of reps and control – similarity).



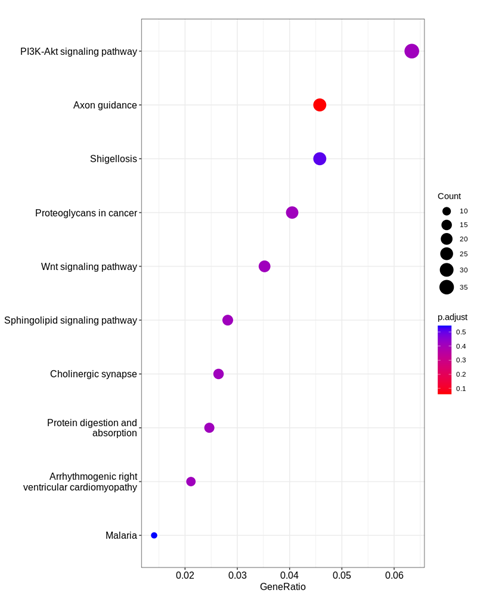
1b) PCA (principal component analysis) for read coverage calculations. Used to determine whether samples display greater variability between experimental conditions then between replicates of the same treatment based on read coverage values information.



2

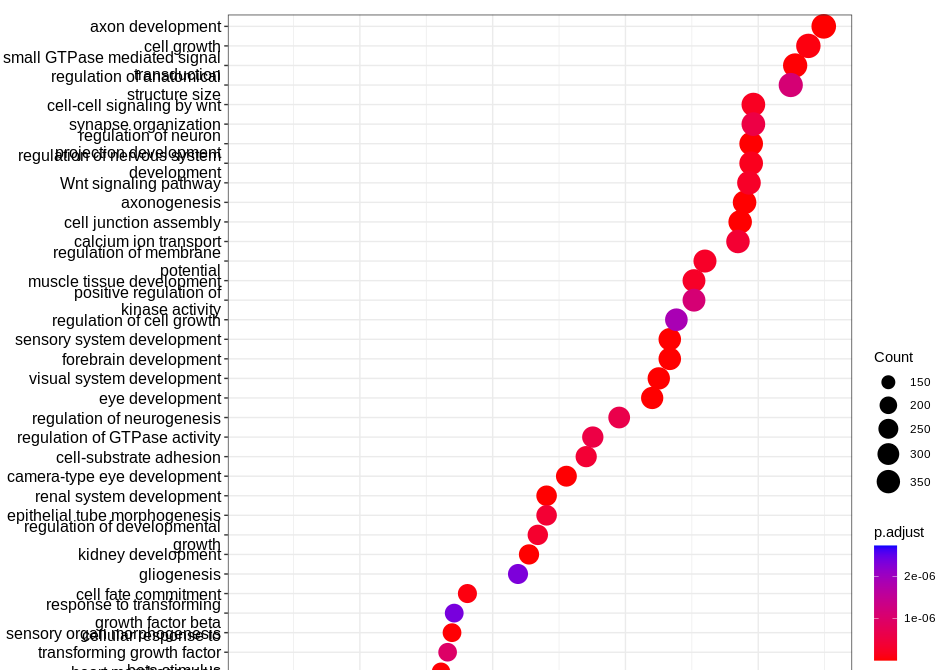


3b)

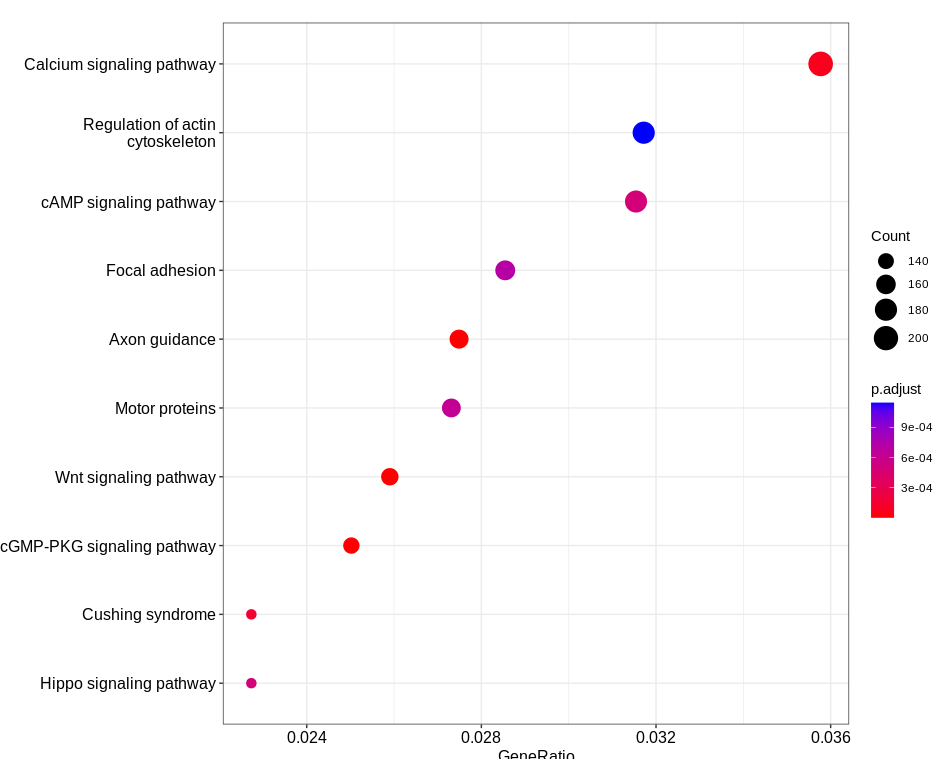


4. CNN Training/Testing

4a)



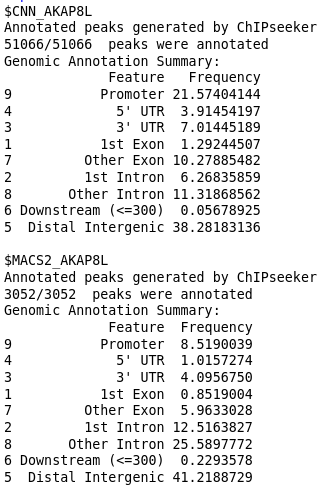
4b)



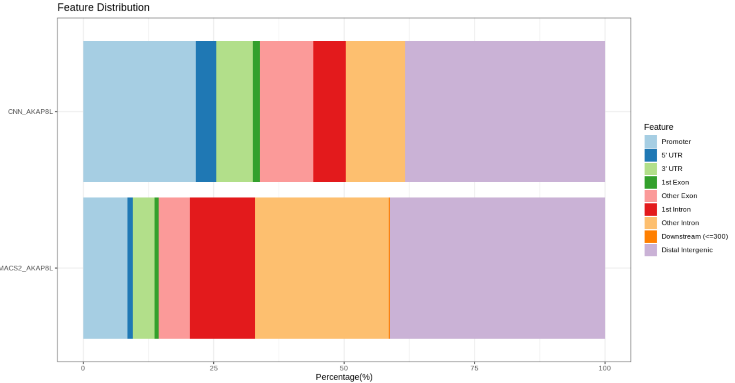
5. Comparative Analysis

5a) **INCLUDE IGV SHOWING MACS2, CNN AND CONTROL TRACKS**

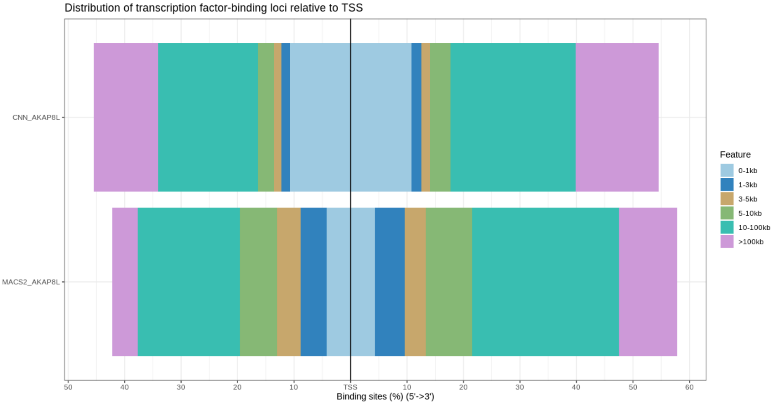
5a)



5b)



5c)



5d)



**Figure 5d.*****Gene over-representation comparison between MACS2 and CNN peak calling models.*** Combined enrichment analysis between AKAP8L peak calling models using the KEGG database.

5e)

