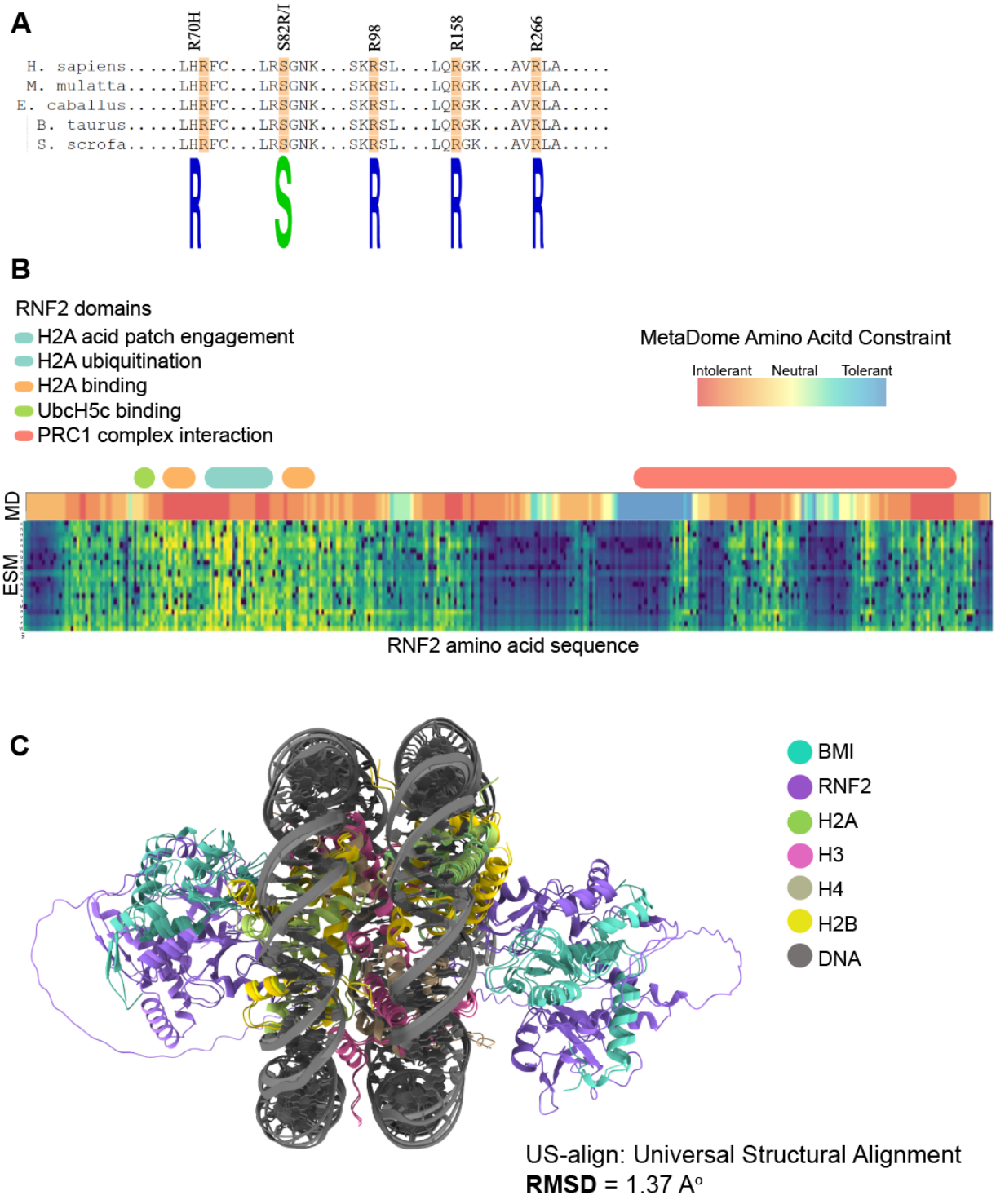
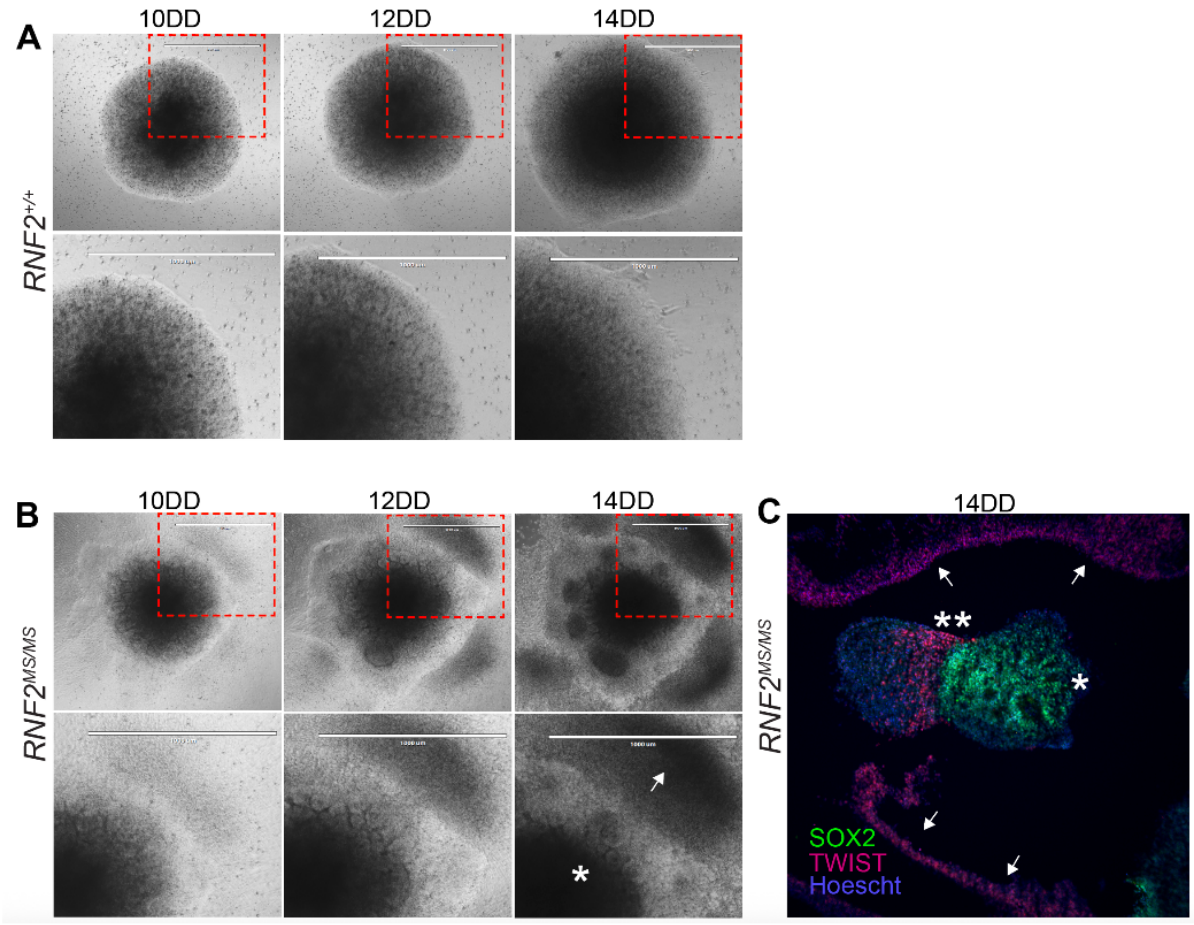


Supplemental Data Figure 1: RNF2 conservation and constraint.

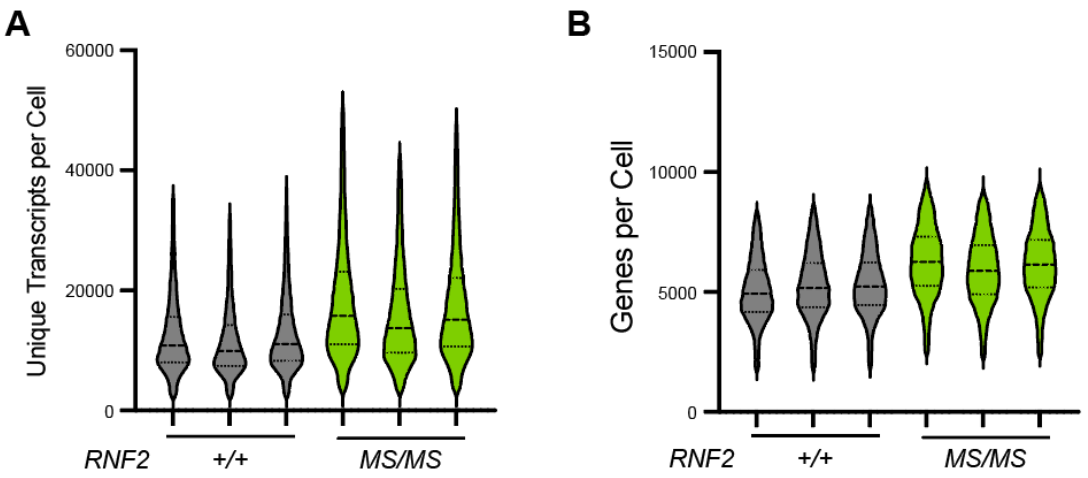
(A) Amino acid sequences of RNF2 across species, showing conservation of residues associated with the missense variants reported in Table 1. **(B)** MetaDome and ESM2

protein language model-based prediction of the effects of variants on the fold and function of RNF2. **(C)** Model of RNF2. Alphafold3 was used for modelling the structure of the PRC1 complex, and the Universal Structure alignment tool was used for a 3D structural alignment and cryo-EM structure of PRC1 complex.



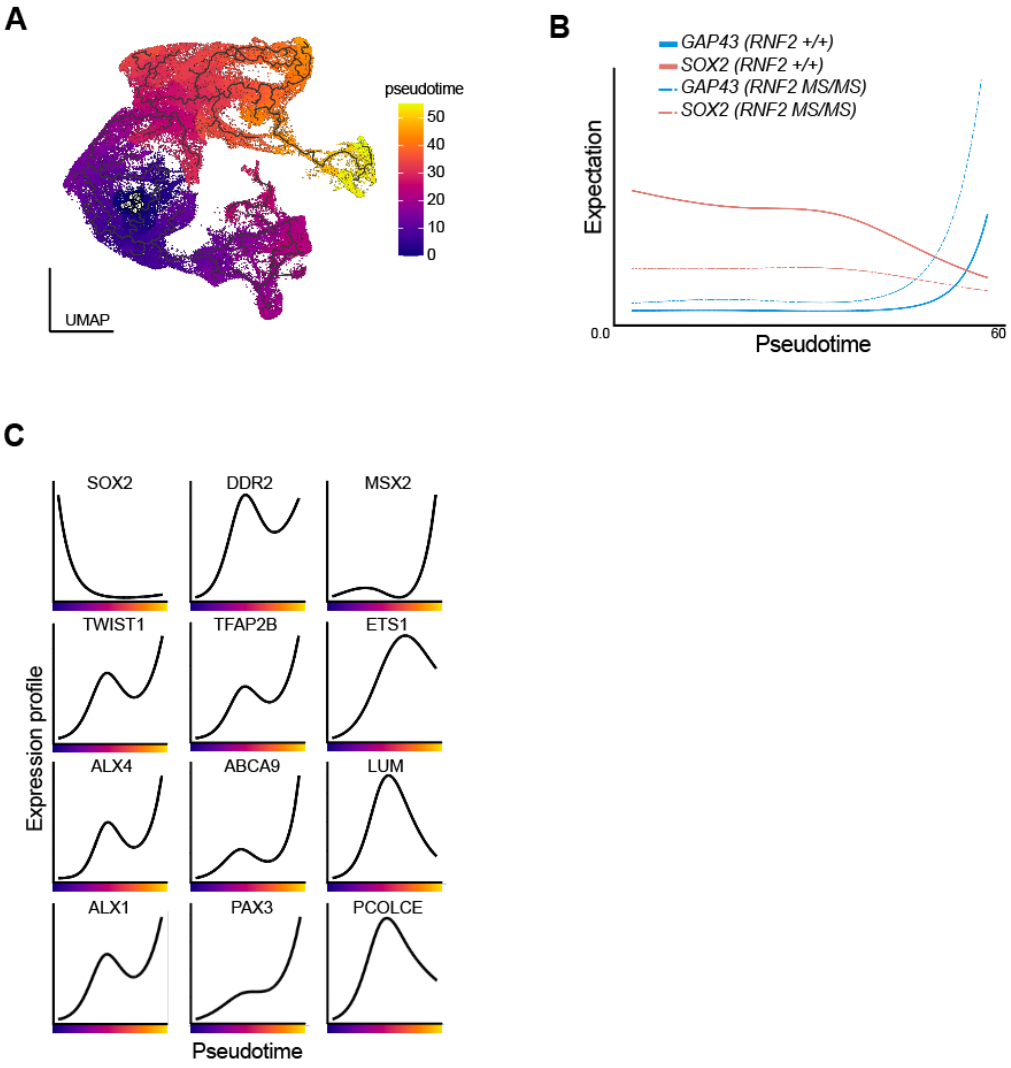
Supplemental Data Figure 2: Rosette mounds for *RNF2*$^{+/+}$ and *RNF2*$^{MS/MS}$ over dual SMAD inhibition neural differentiation.

(A) Representative images of *RNF2*$^{+/+}$ rosette mounds on differentiation days 10, 12, and 14. **(B)** Representative images of *RNF2*$^{MS/MS}$ rosette mounds on differentiation days 10, 12, and 14. Red insets highlight delamination and migration of cells to the periphery. Asterisks indicate neural rosette formation with the arrow indicating migration of cells. **(C)** Representative immunohistochemistry showing SOX2 (green), TWIST1 (red), and Hoescht (blue) immunostaining of *RNF2*$^{MS/MS}$ neural rosette cytoarchitecture. This experiment was repeated 3 times with similar results.



Supplemental Data Figure 3: scRNAseq Quality Control.

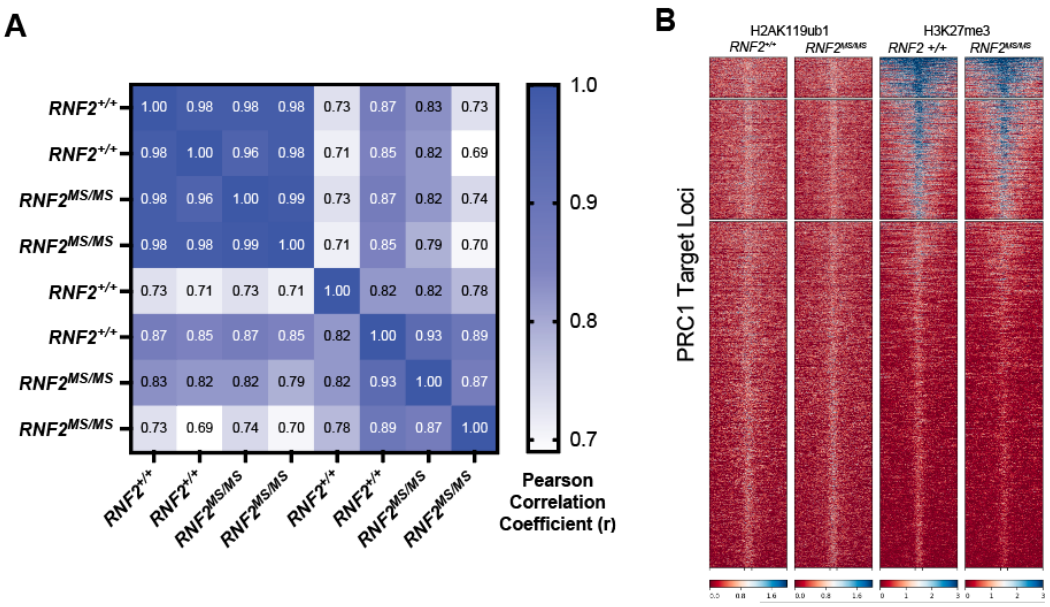
(A) Violin plot demonstrating the number of unique transcripts per cell per replicate. **(B)** Violin plot demonstrating the number of genes detected per cell for each replicate.



Supplemental Data Figure 4: H2AK119ub1 and H3K27me3 correlation plots.

(A) UMAP with pseudotime trajectory, depicting paths between similar single cell transcriptomes connected by a black line, with cells colored by pseudotime (their distance

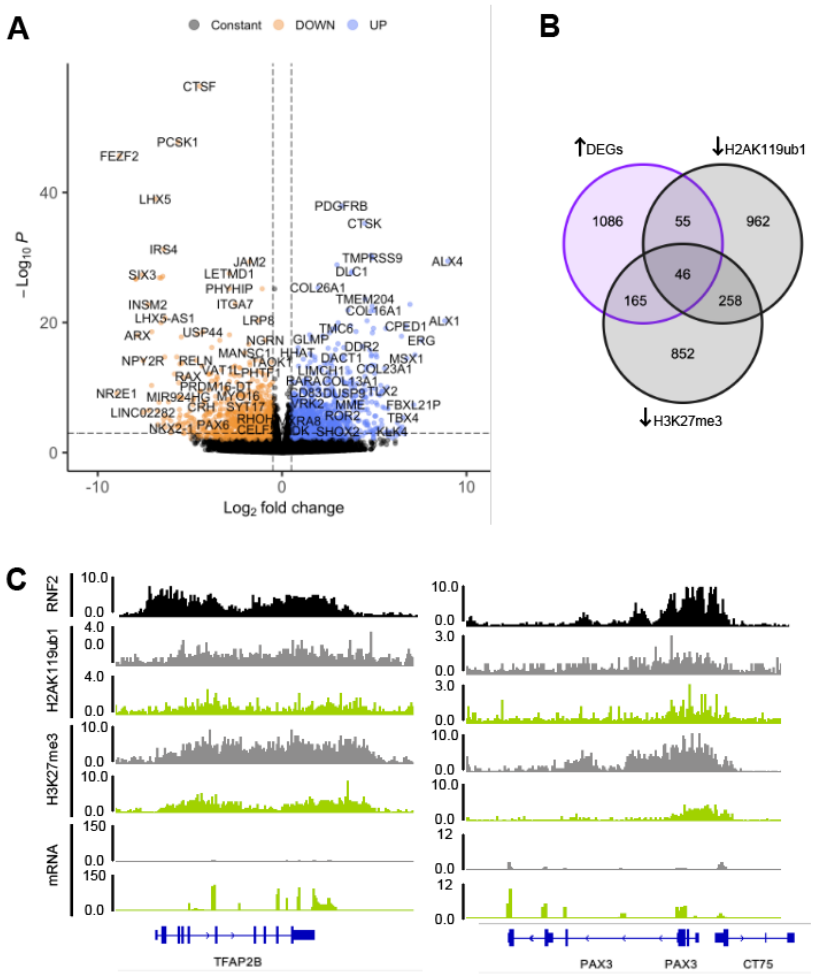
from the origin of the lines). **(B)** The expression of SOX2 (red) and GAP43 (blue) over pseudotime along the neural trajectory, plotted as Monocle3 expectation values, for *RNF2*$^{+/+}$ (solid line) and *RNF2*$^{MS/MS}$ (dashed line). **(C)** The expression of genes as a function of pseudotime along the mesenchymal trajectory, plotted as Monocle3 expectation values.



Supplemental Data Figure 5: H2AK119ub1 and H3K27me3 correlation plots.

(A) Pearson's correlation matrix with coefficients calculated between individual H3K27me3 CUT&RUN replicates using counts from 4255 polycomb target domains. **(B)**

Heatmap of H2AK119ub1 and H3K27me3 occupancy in *RNF2*$^{+/+}$ and *RNF2*$^{MS/MS}$ dissociated neural rosettes across PRC1 target domains. Loci are divided into three clusters, navy, turquoise, and yellow, via k-means clustering.



Supplemental Figure 6: Differentially expressed loci in *RNF2*$^{MS/MS}$ correspond to altered H2AK119ub1 and H3K27me3 occupancy.

(A) Volcano plot of bulk RNA sequencing results, highlighting genes with significantly increased expression (purple) and decreased expression (orange). **(B)** Overlap between the number of genes with significantly increased expression in bulk RNA sequencing, genes with the 5% greatest reduction of H2AK119ub1 in *RNF2*$^{MS/MS}$, and genes with the 5% greatest reduction of H3K27me3 in *RNF2*$^{MS/MS}$. Histogram demonstrating the distribution of overlap following 100,000 simulated 3-way overlaps of randomly selected genes, with the observed overlap between these three groups represented by the dashed vertical purple line. **(C)** RNF2, H2AK119ub1, and H3K27me3 CUT&RUN tracks with bulk RNA sequencing tracks at the *TFAP2B and PAX3* locus. CUT&RUN data are normalized to counts per 10 million mapped reads. Genic signal from two merged biological CUT&RUN replicates.