

# Biology and Bias in Cell Type-Specific RNAseq of Nucleus Accumbens Medium Spiny Neurons

Kronman, Hope<sup>1\*</sup> & Richter, Felix<sup>2,3\*</sup>, Labonté, Benoit<sup>1</sup>, Chandra, Ramesh<sup>4</sup>, Zhao, Shan<sup>5</sup>, Hoffman, Gabriel<sup>3</sup>, Lobo, Mary Kay<sup>4</sup>, Schadt, Eric E.<sup>3</sup>, Nestler, Eric J.<sup>1#</sup>

\* These authors contributed equally to this work

# Corresponding author, e-mail: eric.nestler@mssm.edu

<sup>1</sup> Dept of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai

<sup>2</sup> Dept of Pediatrics, Icahn School of Medicine at Mount Sinai

<sup>3</sup> Dept of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai

<sup>4</sup> Dept of Anatomy and Neurobiology, University of Maryland School of Medicine

<sup>5</sup> Dept of Anesthesiology, Icahn School of Medicine at Mount Sinai

## Supplementary figure legends

### Supplementary Figure 1. Differences in genomic distribution and differential expression across methods.

- (A) Percent of intronic reads across samples for each gene, median indicated as a line in method color
- (B) Differential expression overlap using lenient filters on all (1 FPM in at least two samples)
- (C) Differential expression overlap using strict filters on all (FPKM  $\geq 1$  in at least one sample and all samples with FPKM  $< 5.0 \times 10^4$ )
- (D) Differential expression overlap using strict filter for whole cell and nuclear and a less strict filter for RiboTag (FPKM  $\geq 0.1$  in at least two samples and all samples with FPKM  $< 5.0 \times 10^4$ )
- (E) Density plots of gene lengths for all genes, protein-coding genes, and non-coding genes for each method (whole cell = black, nuclear = blue, RiboTag = red)

### Supplementary Figure 2. Hub network overlaps and nuclear DET-predicted TFBSs

- (A) Overlap of hub network genes for D1- and D2-correlated modules across method
- (B) Overlap of genes showing regulation by transcription, cytosolic mechanisms, or translation in D1- and D2-MSNs. 0 overlaps are not labeled.
- (C) List of 27 miRNAs predicted upstream of the (Transcription D1) / (Cytosol D2) overlap (overlap p-value  $< 0.05$ ) and FPKM values of the 4 miRNAs from this list that are differentially expressed in D2-nuclei (but not in D2-whole cells)

### Supplementary Figure 3. Variance in just genes detected in all three methods.

- (A) PCA of combined data showing separation of samples by cell type on PC5/PC6. Key is shown in figure
- (B) PCA of nuclear, whole cell, and RiboTag datasets from this study along with RiboTag RNA-sequencing from Song et al, 2018. Key is shown in figure
- (C) Variance explained by method, cell type and residuals across the transcriptome. Central bars represent median

### Supplementary Figure 4. Hub network overlaps and nuclear DET-predicted TFBSs

- (D) Overlap of hub network genes for D1- and D2-correlated modules across method
- (E) Overlap of predicted TFBSs for D2 whole cell and RiboTag hub networks with predicted TFBSs for nuclear D2-DEGs
- (F) Overlap of predicted TFBSs for D1 whole cell and RiboTag hub networks with predicted TFBSs for nuclear D1-DEGs

### Supplementary Figure 5. Variance distributions for D1- and D2-DETs across methods.

- (A) Variance distribution for nuclear D1- and D2-DEGs (top of the box = 3<sup>rd</sup> quartile, midline = median, bottom of the box = 1<sup>st</sup> quartile)

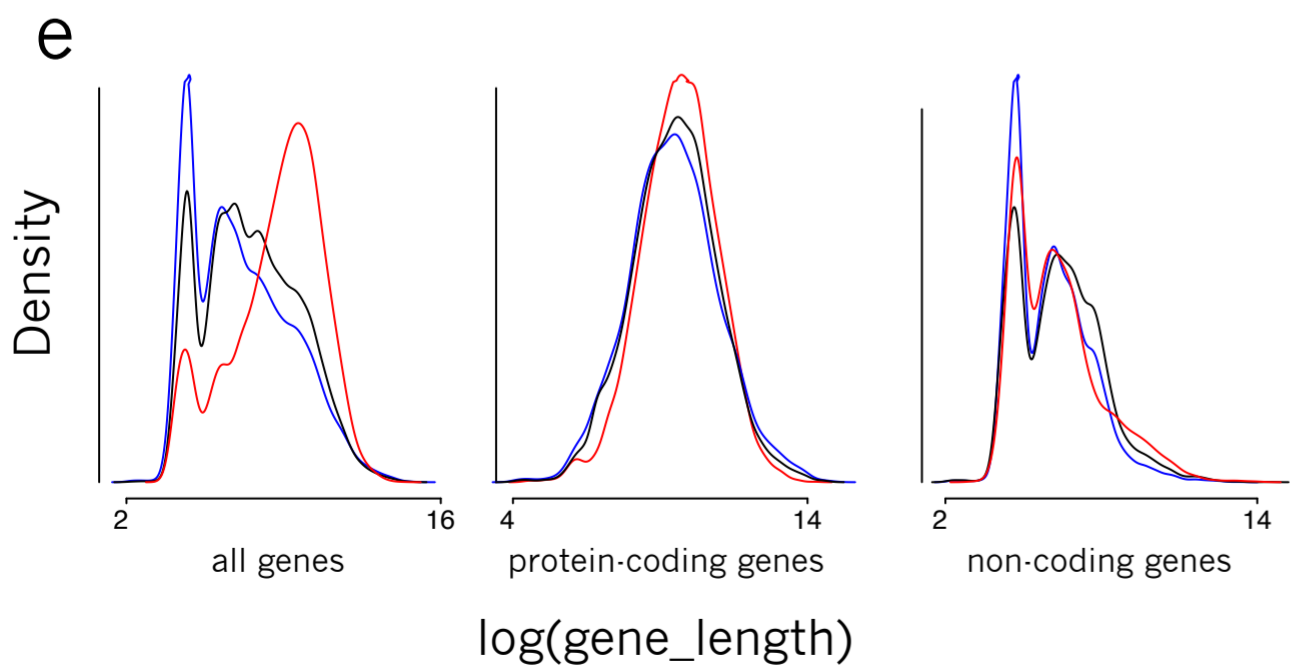
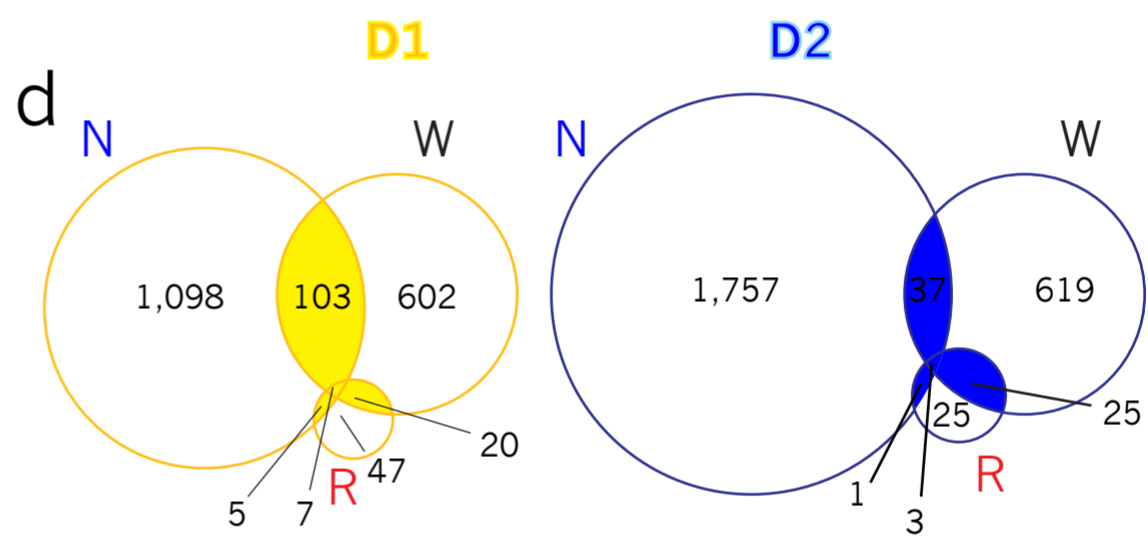
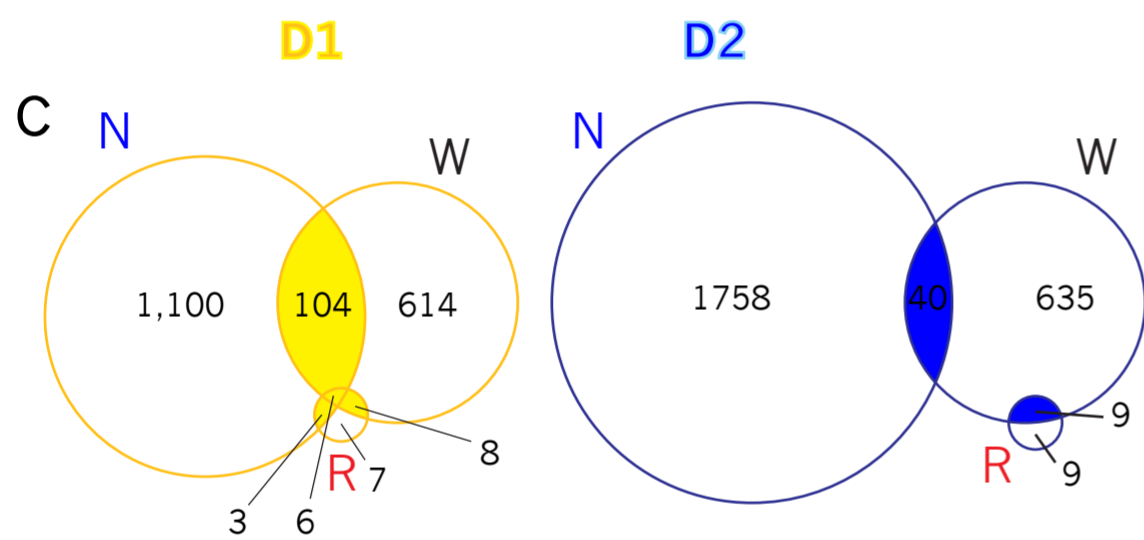
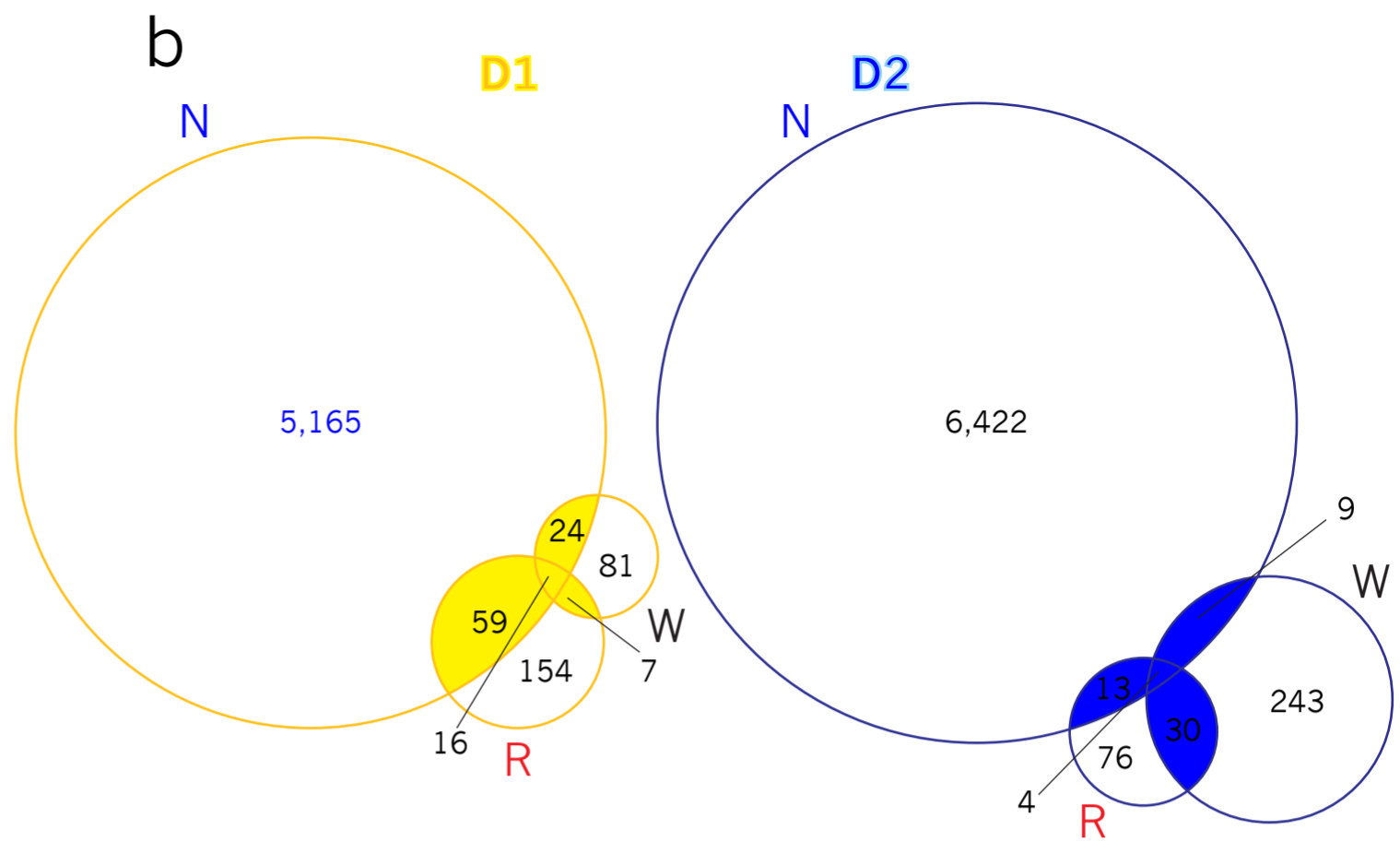
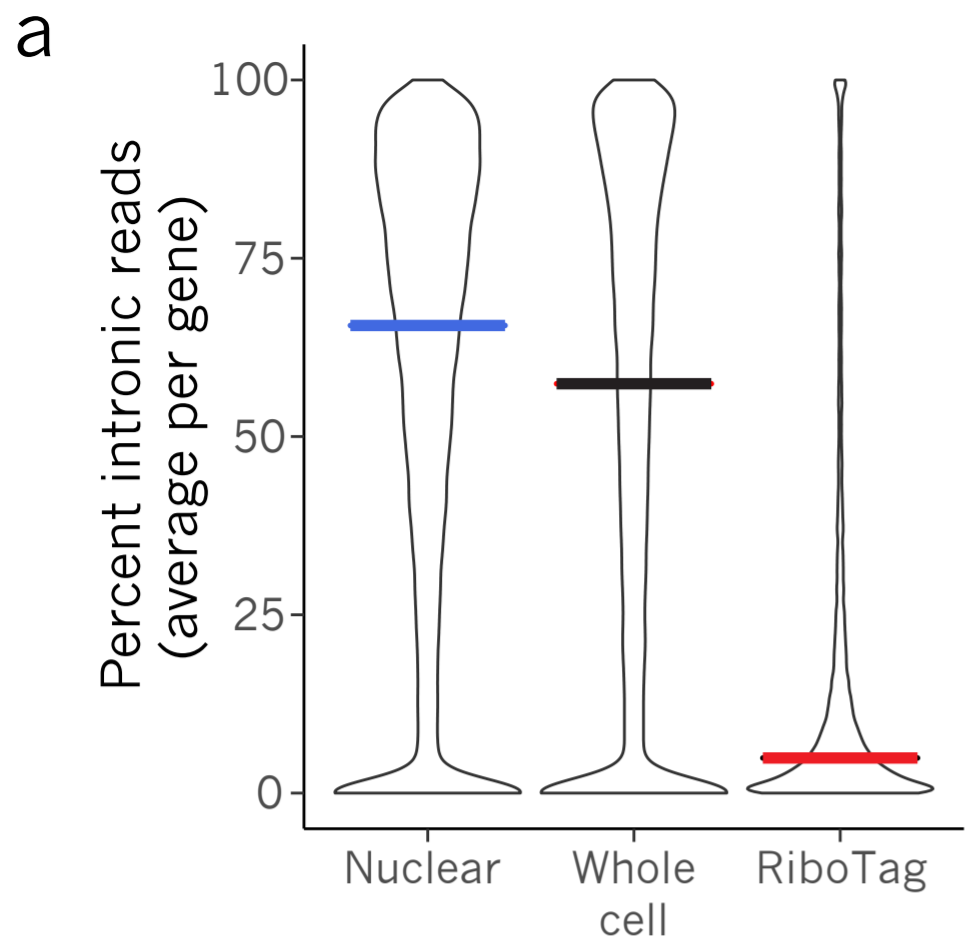
- (B) Variance distribution for RiboTag D1- and D2-DEGs (top of the box = 3<sup>rd</sup> quartile, midline = median, bottom of the box = 1<sup>st</sup> quartile)
- (C) Variance distribution for whole cell D1- and D2-DEGs (top of the box = 3<sup>rd</sup> quartile, midline = median, bottom of the box = 1<sup>st</sup> quartile)

**Supplementary Figure 6. FACS plots demonstrating gating strategy.**

- (A) Gating strategy for D1-whole cells. From the left, selecting whole neurons (SSC vs. FSC), selecting live neurons (DAPI vs. FSC), selecting GFP-negative, tdTomato-positive whole neurons (GFP vs. Texas Red)
- (B) Gating strategy for D2-whole cells. From the left, selecting whole neurons (SSC vs. FSC), selecting live neurons (DAPI vs. FSC), selecting GFP-positive, tdTomato-negative whole neurons (GFP vs. Texas Red)
- (C) Gating strategy for D1-nuclei. From the left, selecting nuclei (SSC vs. FSC), selecting GFP-positive nuclei (GFP vs. FSC)
- (D) Gating strategy for D2-nuclei. From the left, selecting nuclei (SSC vs. FSC), selecting GFP-positive nuclei (GFP vs. FSC)

**Supplementary table list**

- Supplementary Table 1: Nuclear DEG logistic regression
- Supplementary Table 2: Differential expression by method
- Supplementary Table 3: FPKM by method
- Supplementary Table 4: FPKM and variance correlations across methods
- Supplementary Table 5: Variance explained by gene
- Supplementary Table 6: Biotypes of genes from variance partitioning analysis
- Supplementary Table 7: gProfiler enrichments
- Supplementary Table 8: Genes enriched in each method
- Supplementary Table 9: GWAS risk loci overlaps



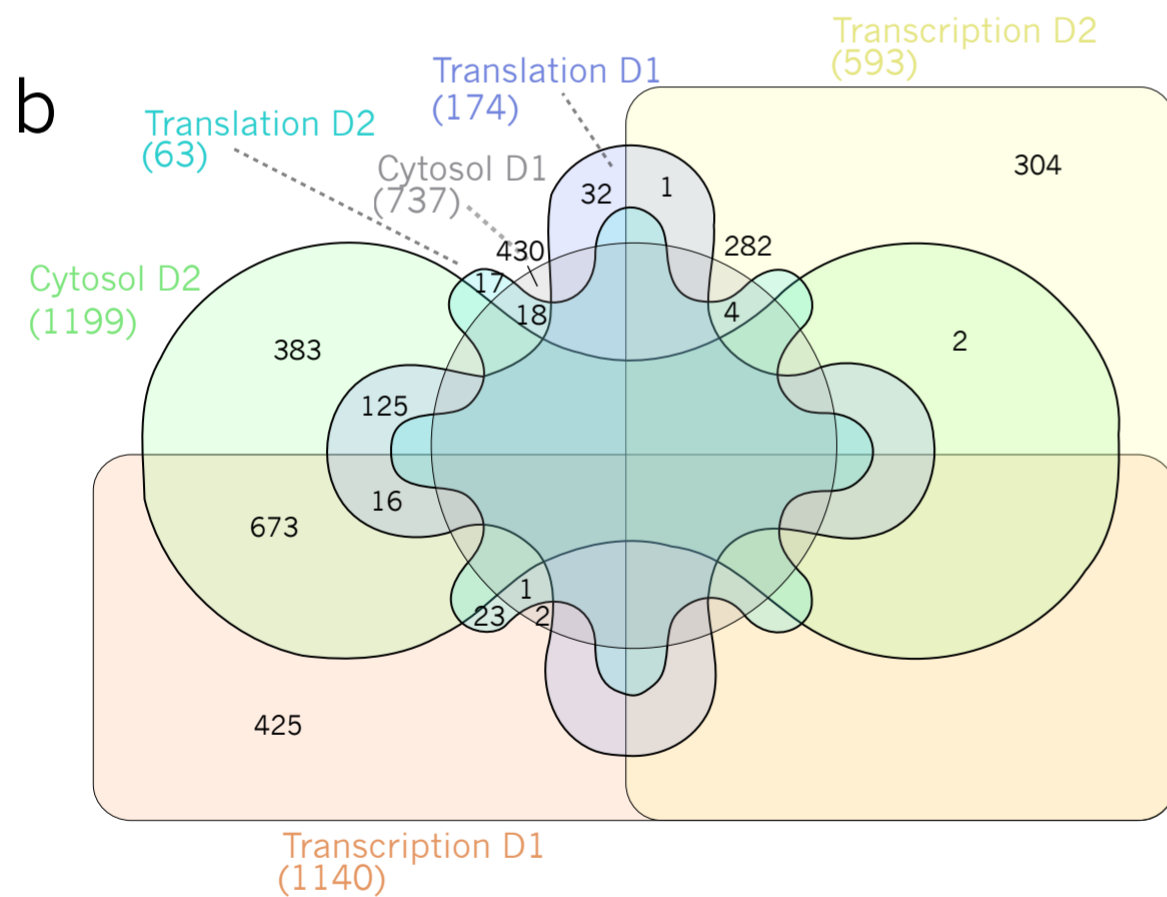
a

$$y_{expr} = \beta_0 + \beta_1 x_{d2} + \beta_2 x_{ribo} + \beta_3 x_{wc} + \beta_4 x_{wc} \cdot x_{d2} + \beta_5 x_{ribo} \cdot x_{d2}$$

Sequencing method...	Measures changes in...	Relevant independent variables are...			Coefficients being solved for are...
		$x_{d2}$	$x_{ribo}$	$x_{wc}$	
Nuclear	<b>Transcription</b>	0 and 1	0	0	$\beta_1$
WC	<b>Transcription + Post-transcriptional processes</b>	0 and 1	0	1	$\beta_1 + \beta_4$
RiboSeq	<b>Transcription + Post-transcriptional processes + Translation</b>	0 and 1	1	0	$\beta_1 + \beta_5$

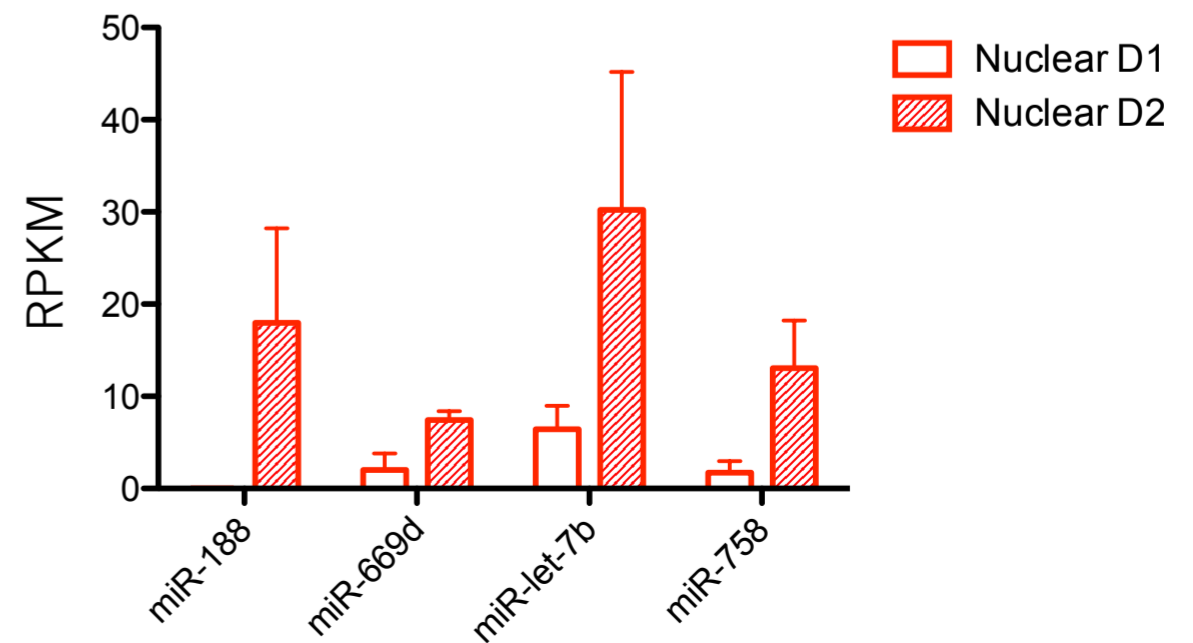
Therefore, genes with significant...  $\beta_1$   
 $(\beta_1 + \beta_4) - \beta_1 = \beta_4$   
 $(\beta_1 + \beta_5) - (\beta_1 + \beta_4) = \beta_5 - \beta_4$

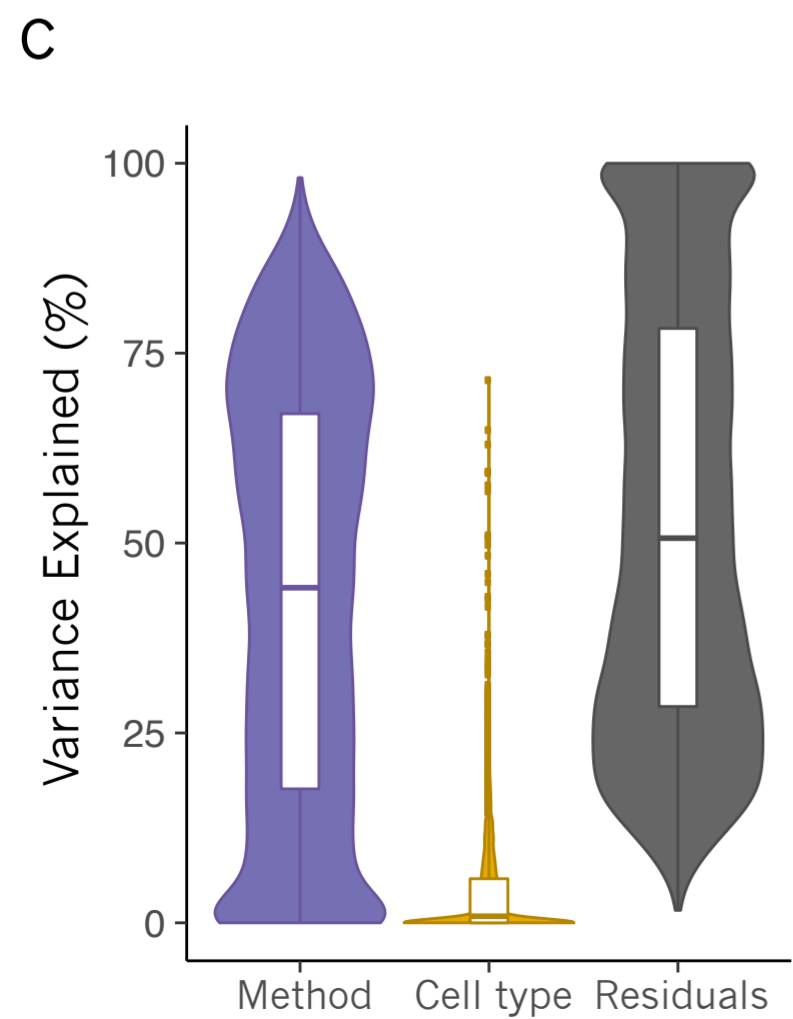
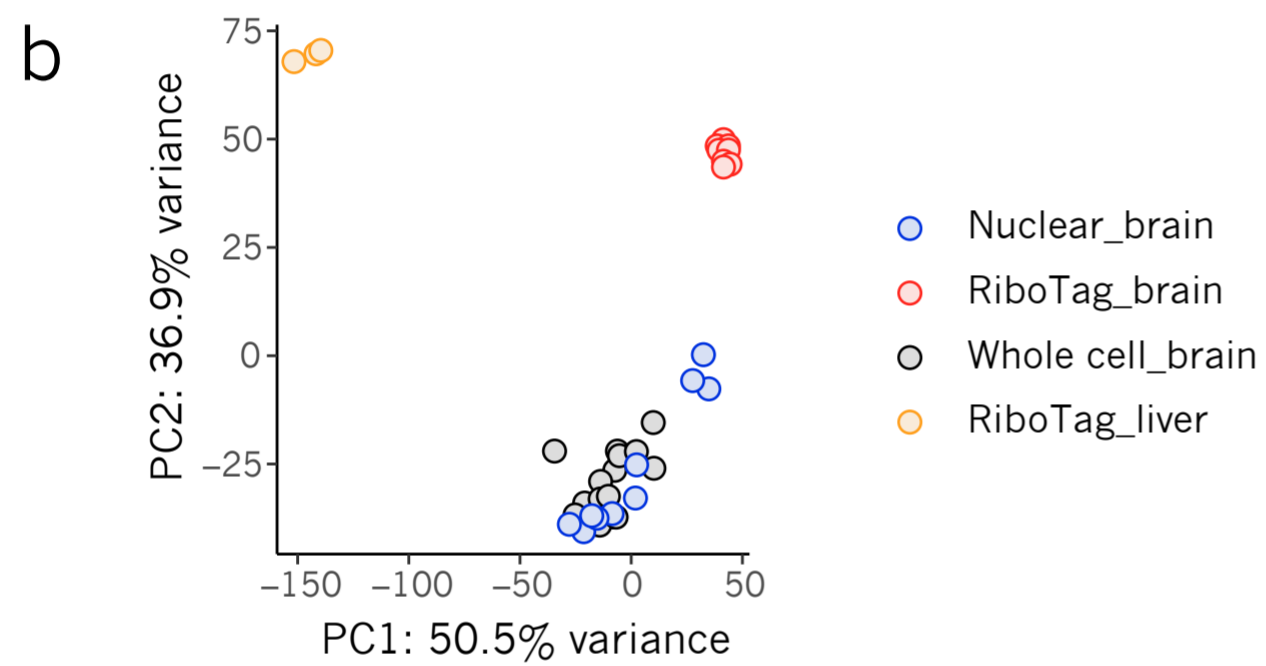
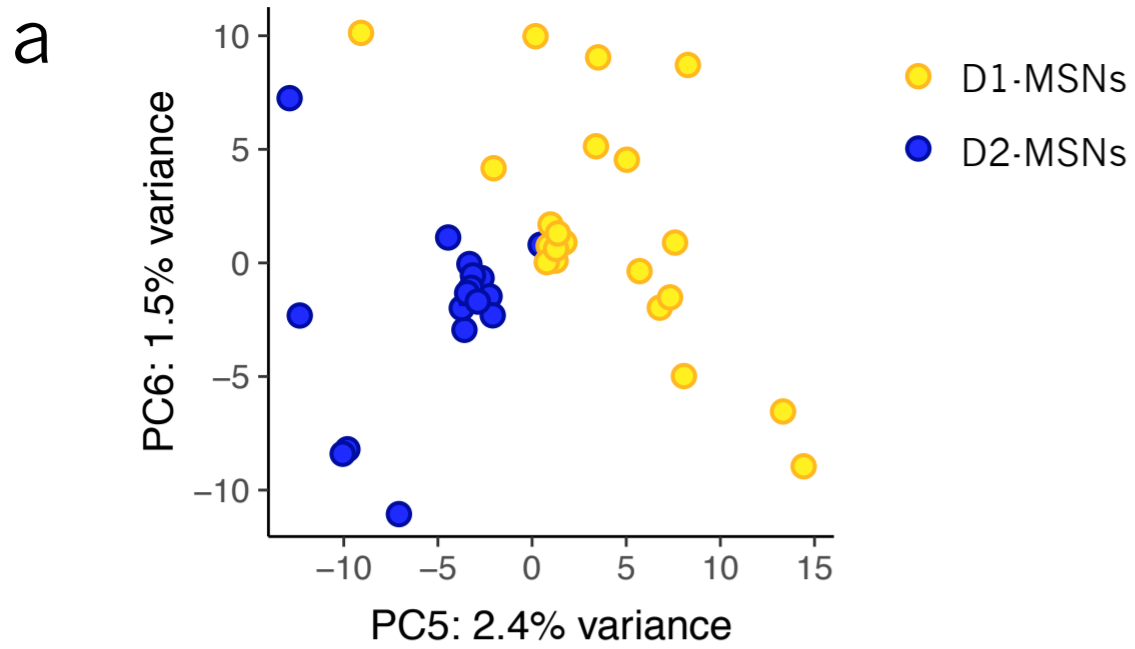
Have differential...  
**Transcription**  
**Post-transcriptional processes**  
**Translation**



c

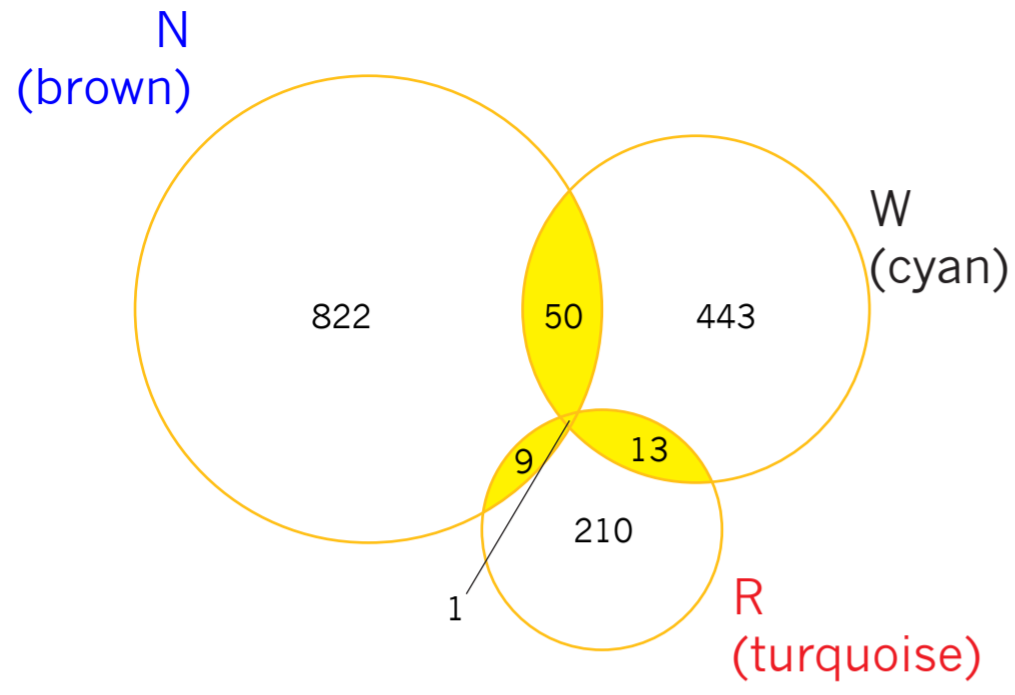
mmu-miR-9-5p	mmu-miR-301b-3p
mmu-miR-7b-5p	mmu-miR-297a-5p
mmu-miR-7a-2-3p	mmu-miR-26a-5p
mmu-miR-7a-1-3p	mmu-miR-24-3p
mmu-miR-758-3p	mmu-miR-223-3p
mmu-miR-669d-5p	mmu-miR-211-5p
mmu-miR-541-5p	mmu-miR-19b-3p
mmu-miR-539-5p	mmu-miR-199a-3p
mmu-miR-490-3p	mmu-miR-188-5p
mmu-miR-425-5p	mmu-miR-181a-5p
mmu-miR-384-3p	mmu-miR-15a-5p
mmu-miR-340-5p	mmu-miR-149-5p
mmu-miR-30e-5p	mmu-let-7b-5p



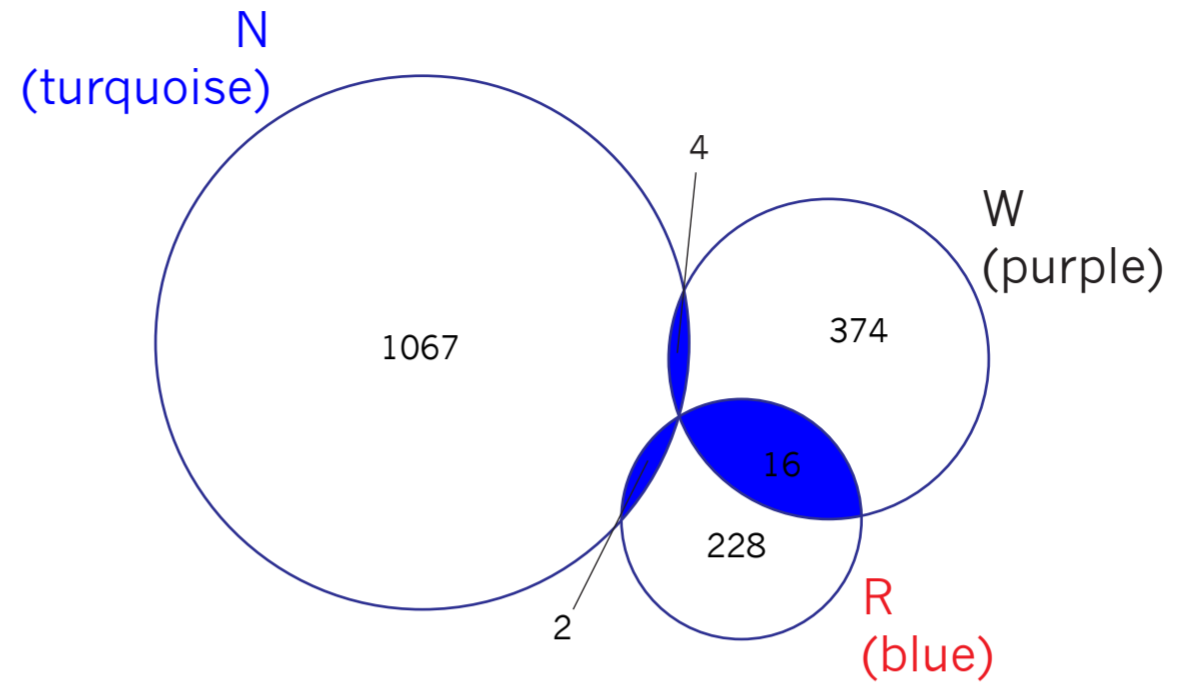


a

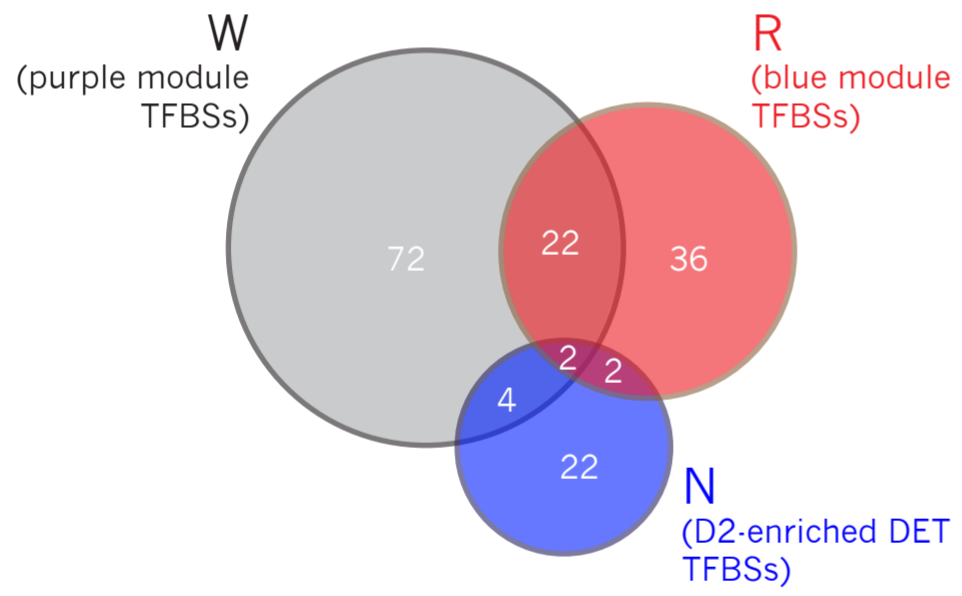
**D1**



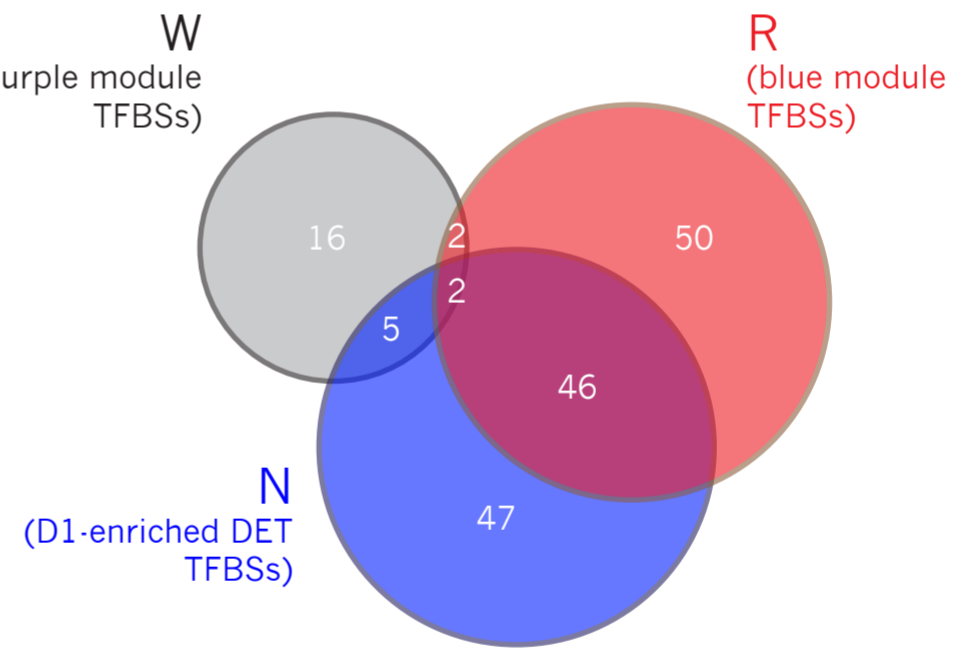
**D2**

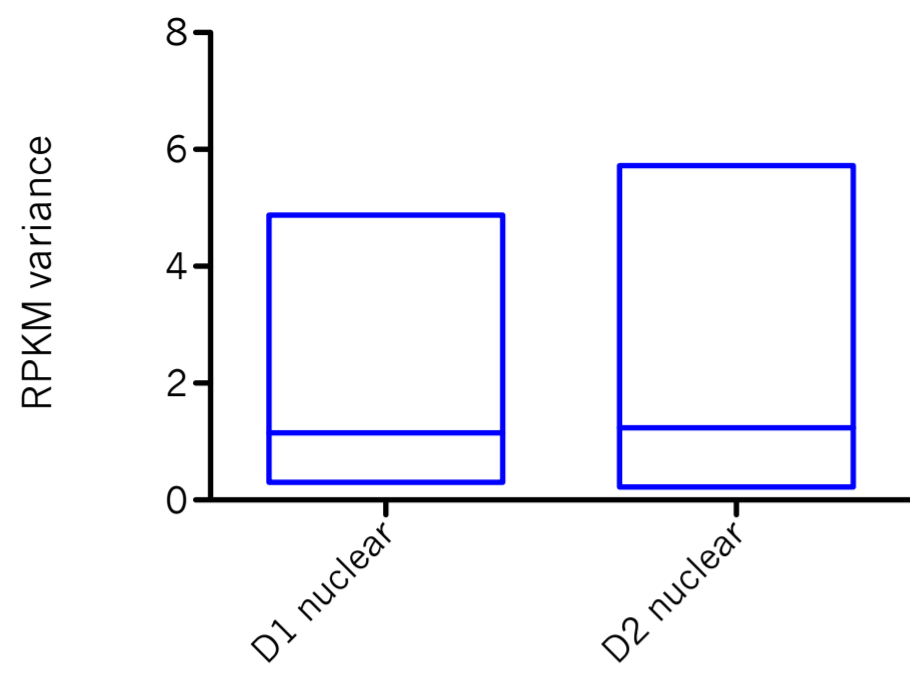
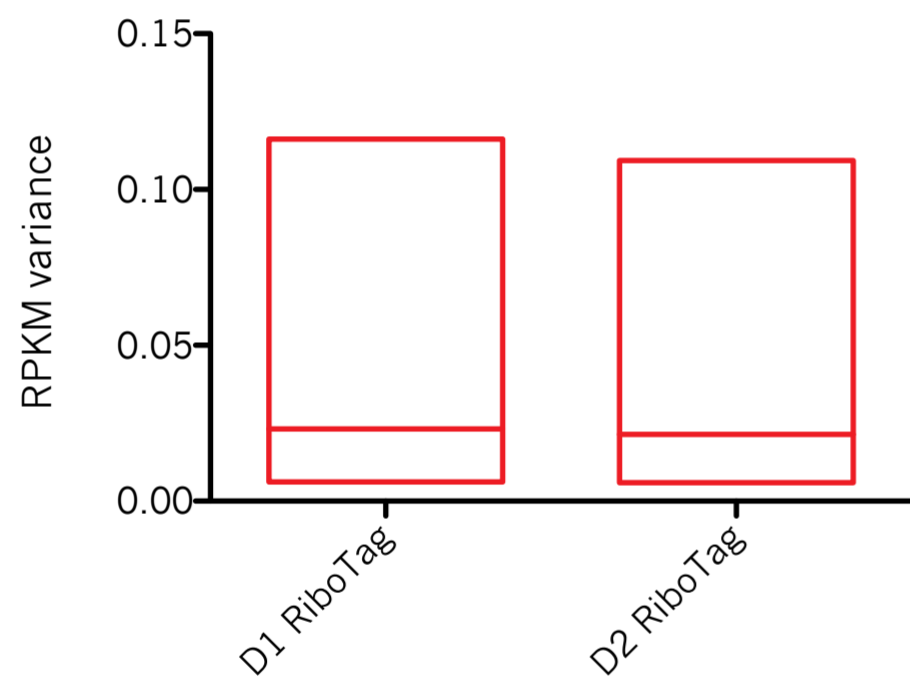
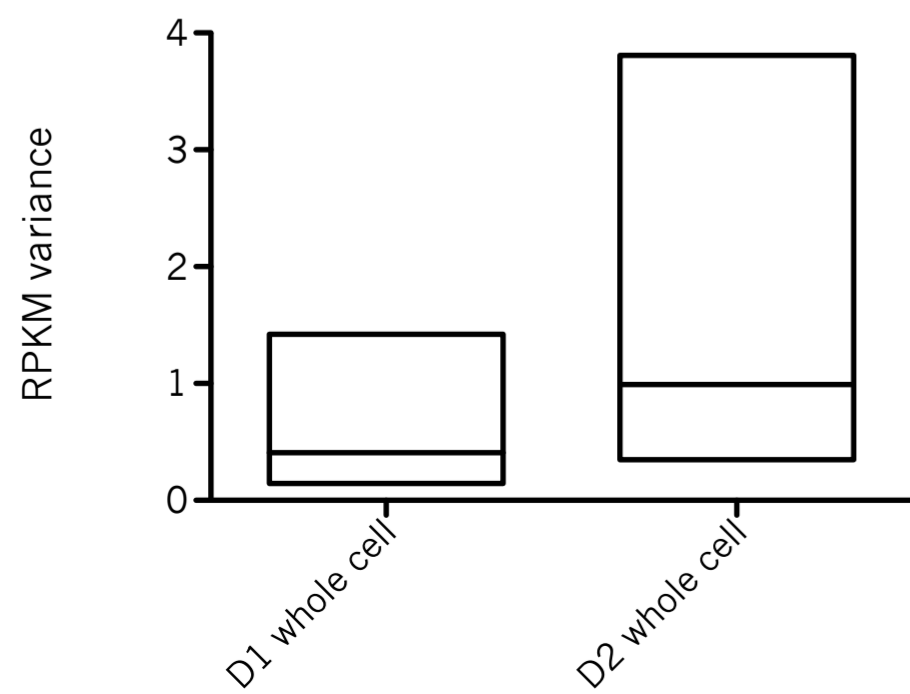


b



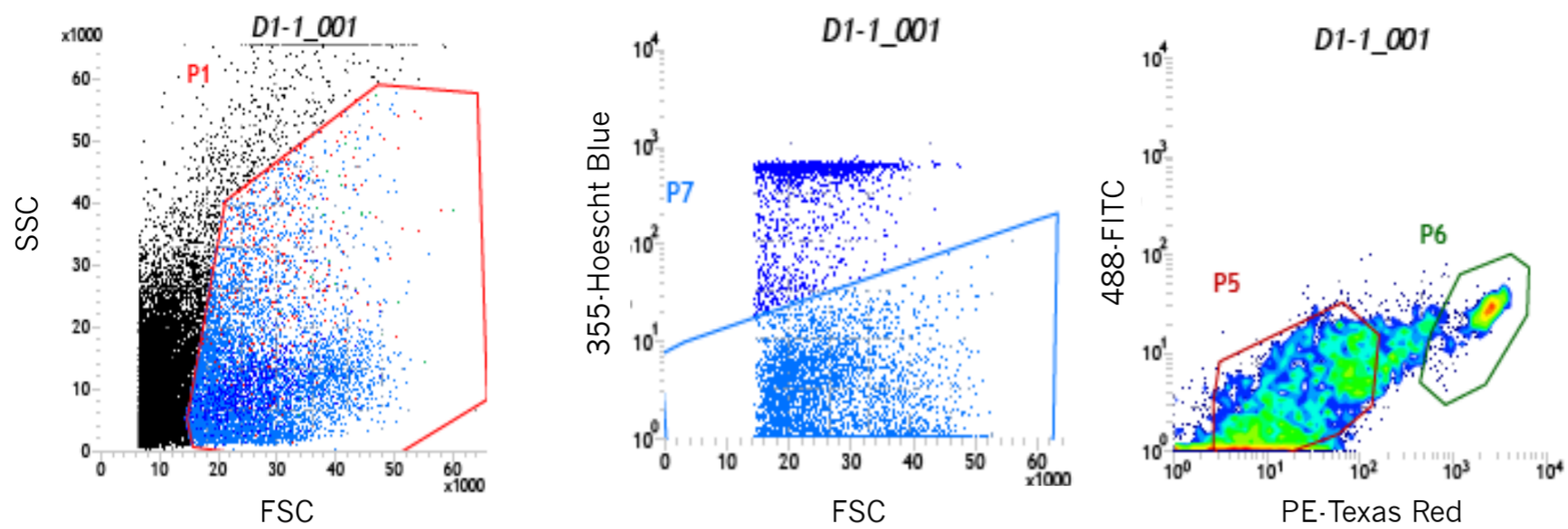
c



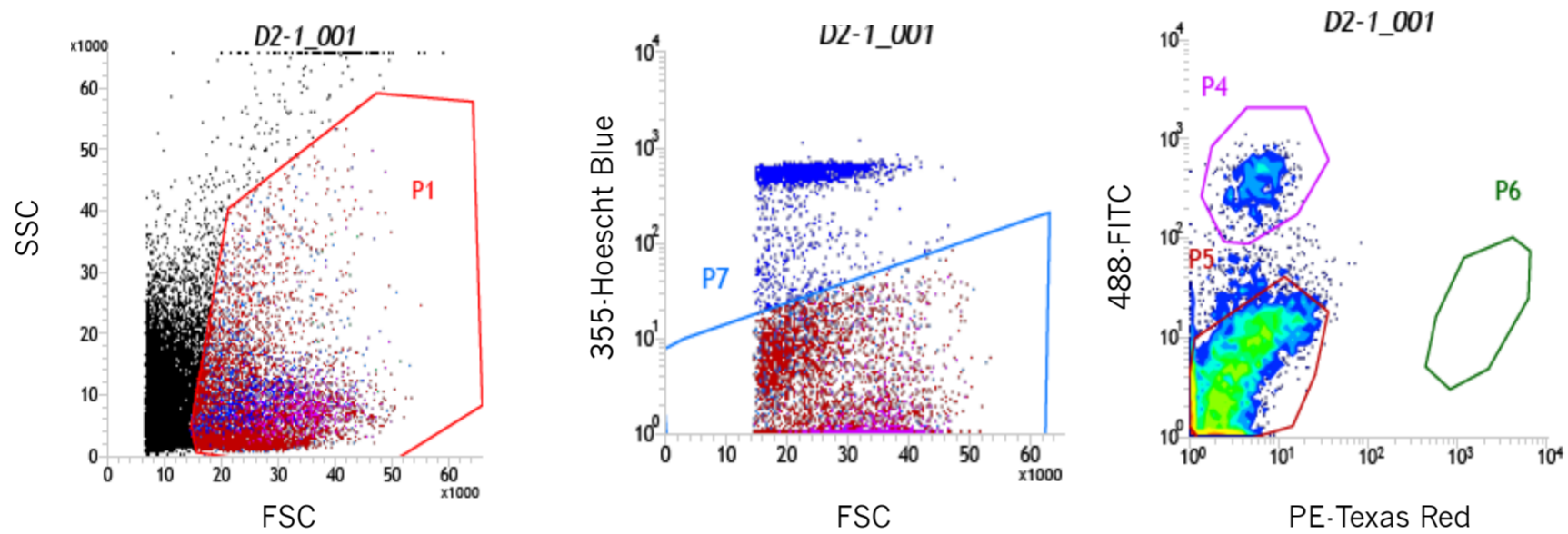
**a****b****c**



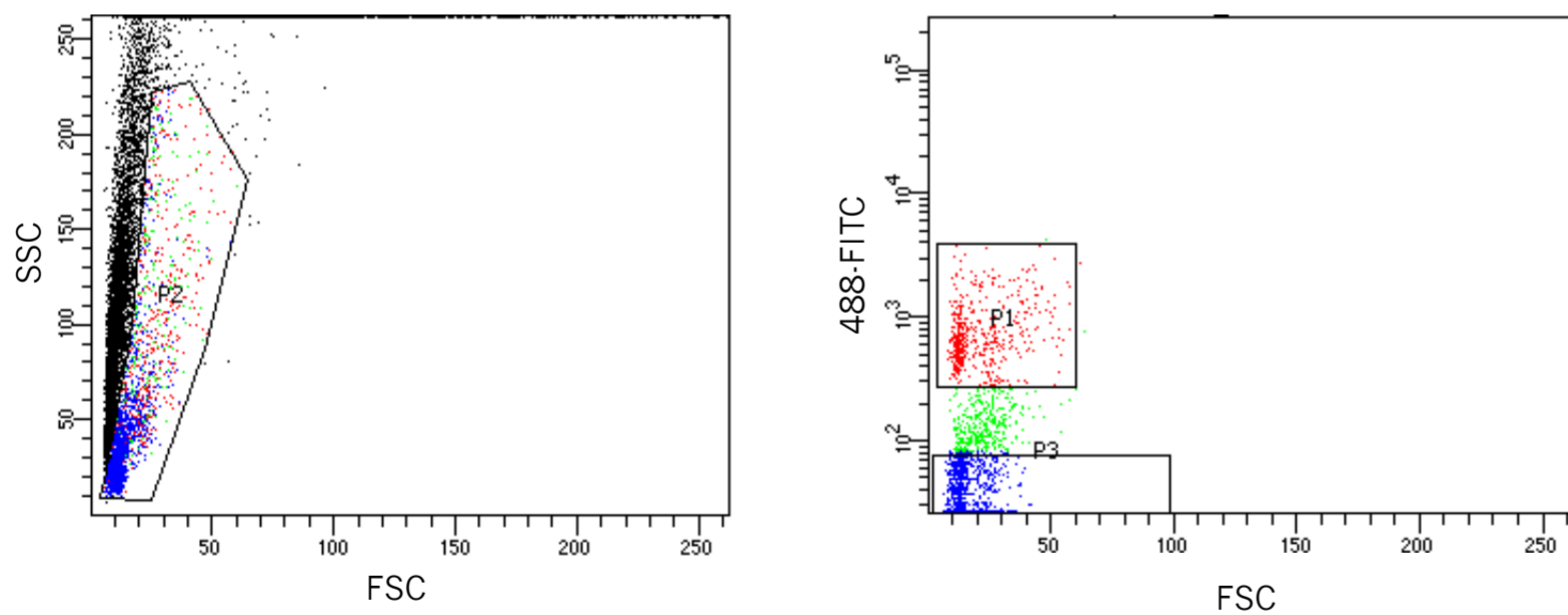
a



b



c



d

