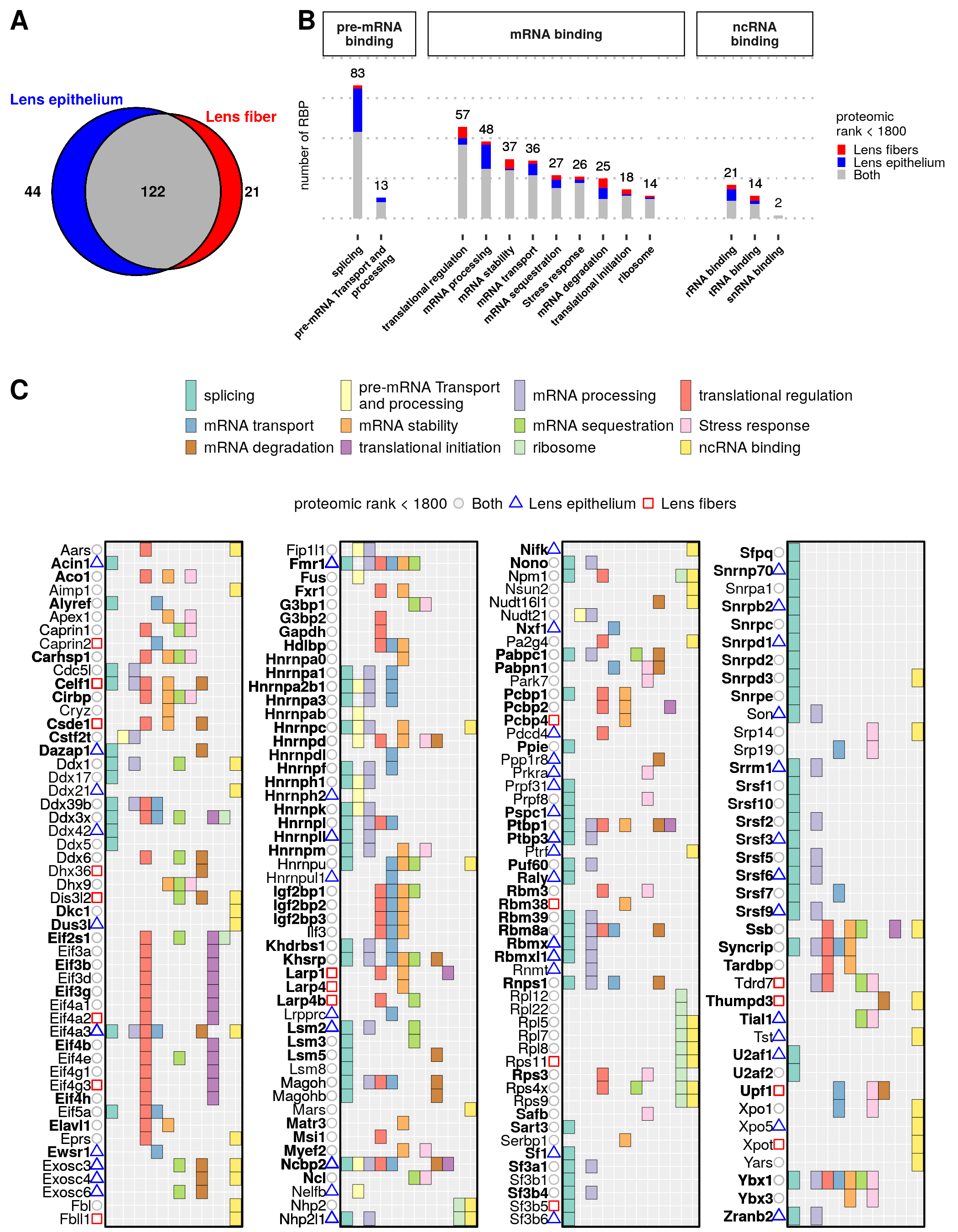
Systematic Analysis of Lens-Expressed RNA-Binding Proteins

Paul Langard

# Introduction

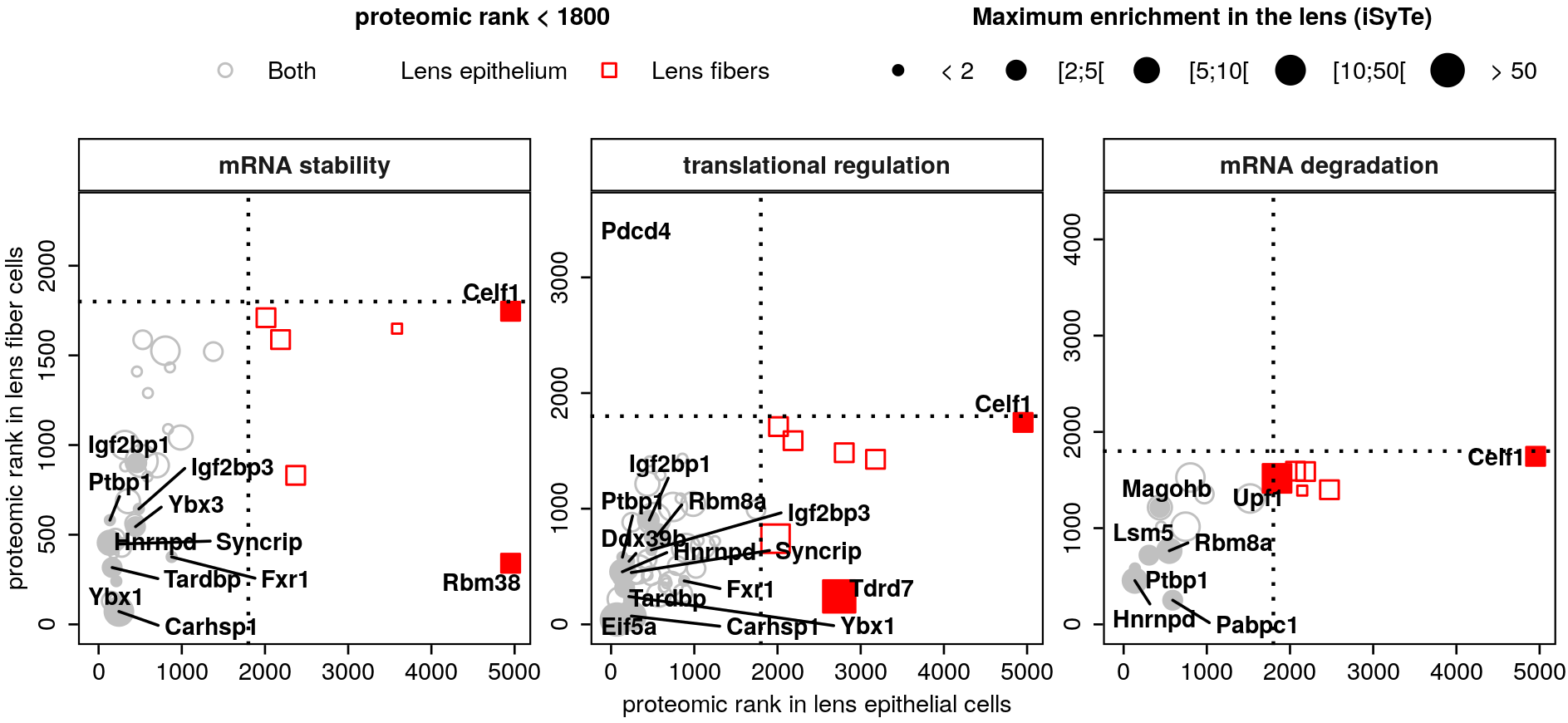
# Results

## Systematic Analysis of Lens-Expressed RNA-Binding Proteins



(A) Venn diagram of highly-expressed RBPs grouped in either lens epithelial or fiber cells or both. (B) Functional representativity of lens-expressed RBPs, divided with regard of target RNAs class. (C) Dotplot showing the functional classification of expressed RBPs in the lens. The tissue specificity of an RBP is indicated by a colored sign next to its name.

t\_ISYTE\_enrichment <- read\_tsv("../data/raw\_data/functionalanalysis/ISYTE\_enrichment.tsv") %>% dplyr::rename(gene=Symbol)  
#> Parsed with column specification:  
#> cols(  
#> .default = col\_character()  
#> )  
#> See spec(...) for full column specifications.  
  
t\_proteomicEnrichment <- t\_ISYTE\_enrichment %>%  
 gather(-gene, -Rank, key="stage", value="enrichement") %>%  
 mutate(stage=str\_extract(string = stage, "E[0-9]{1,2}|P[0-9]{1,2}"), #extract stage  
 enrichement = as.numeric(enrichement)) %>%  
 group\_by(gene) %>%  
 summarise(enrich\_category = max(enrichement, na.rm = T)) %>%  
 mutate(enrich\_category = case\_when(enrich\_category < 2 ~ "< 2",  
 enrich\_category >= 2 & enrich\_category < 5 ~ "[2;5[",  
 enrich\_category >= 5 & enrich\_category < 10 ~ "[5;10[",  
 enrich\_category >= 10 & enrich\_category < 50 ~ "[10;50[",  
 enrich\_category >= 50 ~ "> 50")) %>%  
 mutate(enrich\_category = factor(enrich\_category, levels = c("< 2", "[2;5[", "[5;10[","[10;50[", "> 50")))%>%  
 right\_join(t\_annotationsRBP\_figure, by="gene") %>%   
 filter(groups %in% c("translational regulation", "mRNA degradation", "mRNA stability")) %>%  
 mutate(groups = factor(groups, levels = c("mRNA stability", "translational regulation","mRNA degradation"))) %>%  
 replace\_na(list(enrich\_category="NA"))  
#> Warning: NAs introduits lors de la conversion automatique  
#> Warning in `[<-.factor`(`\*tmp\*`, !is\_complete(data[[var]]), value = "NA"):  
#> invalid factor level, NA generated  
  
listRBPinterest <- "Caprin2,Carhsp1,Celf1,Ddx39b,Fxr1,Igf2bp1,Igf2bp3,Hnrnpd,Rbm8a,Lsm5,Magohb,Pabpc1,Ptbp1,Tardbp,Rbm38,Syncrip,Tdrd7,Ybx1,Ybx3,Lsm2,Eif5a,Pdcd4,Upf1" %>%   
 str\_split(pattern = ",") %>%   
 unlist()  
  
t\_proteomicEnrichment\_interest <- t\_proteomicEnrichment %>%   
 filter(gene %in% listRBPinterest)   
  
t\_proteomicEnrichment %>%  
 ggplot(aes(x=rankEpith, y=rankFibers, color=tissue, size=enrich\_category, shape=tissue)) +  
 geom\_point() +   
 geom\_point(data = filter(t\_proteomicEnrichment\_interest, tissue=="both"),   
 mapping = aes(x=rankEpith, y=rankFibers, size=enrich\_category),   
 color="#c0c0c0", shape=16, inherit.aes = F, show.legend = F) +  
 geom\_point(data=filter(t\_proteomicEnrichment\_interest, tissue=="epi"),   
 mapping = aes(x=rankEpith, y=rankFibers, size=enrich\_category),   
 color="#0000ff", shape=17, inherit.aes = F, show.legend = F) +  
 geom\_point(data=filter(t\_proteomicEnrichment\_interest, tissue=="fibers"),   
 mapping = aes(x=rankEpith, y=rankFibers, size=enrich\_category),   
 color="#ff0000", shape=15, inherit.aes = F, show.legend = F) +  
 scale\_size\_discrete(range= c(1, 4)) +  
 xlim(c(0,NA))+ ylim(c(0,NA))+  
 scale\_shape\_manual(values=c(1, 2, 0),   
 labels=c("Both", "Lens epithelium", "Lens fibers")) +  
 geom\_text\_repel(data = t\_proteomicEnrichment\_interest,   
 mapping = aes(label=gene, x=rankEpith, y=rankFibers),   
 inherit.aes = F,  
 fontface = 2,   
 color = "black",   
 segment.color = "black",  
 size = 2.4,   
 box.padding = unit(.2, "lines"),   
 segment.size = .4,   
 min.segment.length = unit(1, "mm"),  
 show.legend = F,   
 max.iter = 10000,   
 seed = 22,   
 force = 3) +   
 scale\_color\_manual(values = c("epi"="#0000ff",   
 "fibers" = "#ff0000",  
 "both" = "#c0c0c0"),   
 labels=c("Both","Lens epithelium", "Lens fibers")) +  
 geom\_hline(yintercept=1800, color="black", lty="dotted") +   
 geom\_vline(xintercept=1800, lty="dotted")+   
 guides(size = guide\_legend(title = "Maximum enrichment in the lens (iSyTe)", title.position = "top", title.hjust = .5),   
 color=guide\_legend(title="proteomic rank < 1800", title.position = "top", title.hjust = .5),  
 shape=guide\_legend(title="proteomic rank < 1800", title.position = "top", title.hjust = .5))+ theme\_gray() +   
 ylab("proteomic rank in lens fiber cells")+   
 xlab("proteomic rank in lens epithelial cells") +  
 theme(panel.border = element\_rect(color="black", fill = "transparent"),   
 axis.text.x = element\_text(size=7, color="black"),  
 axis.text.y = element\_text(size=7, color="black", angle=90, vjust=1, hjust=.5),  
 axis.title = element\_text(size=7),  
 axis.ticks = element\_line(size=.3, color="black"),  
 strip.background = element\_rect(color="black", fill = "white"),  
 panel.background = element\_rect(fill="transparent"),  
 panel.grid = element\_blank(),  
 axis.text = element\_text(size=8, color="black"),  
 strip.text = element\_text(face="bold", size=7),   
 legend.position = "top",  
 legend.margin = margin(0,0,0,10, "mm"),  
 legend.spacing = unit(0, "cm"),  
 legend.direction = "horizontal",  
 legend.box = "horizontal",  
 legend.title = element\_text(size=7, face = "bold"),   
 legend.box.margin = margin(0,0,0,0,"mm"),  
 legend.key = element\_blank(),  
 legend.text = element\_text(size=7)) +   
 facet\_wrap(~groups, scales = "free") +   
 theme(plot.margin = margin(.05,.12,.05,.05,"cm"))  
#> Warning: Using size for a discrete variable is not advised.  
#> Warning: Removed 12 rows containing missing values (geom\_point).



## RNA-binding predictions

# Discussion