OOPS CL:NC and RNase +:- ratios

Here we want to visualise the CL vs NC (non-crosslinked) and RNase vs Ctrl SILAC experiments.

Below we load the required packages and set a plotting theme.

```
# load packages
library(tidyverse)
library(ggbeeswarm)

# set up standardised plotting scheme
theme_set(theme_bw(base_size = 20) +
            theme(panel.grid.major=element_blank(),
                  panel.grid.minor=element_blank(),
                  aspect.ratio=1))

cbPalette <- c("#E69F00", "#56B4E9", "#009E73", "#F0E442",
               "#0072B2", "#D55E00", "#CC79A7", "#999999")
```

We start by reading in the data. Our input here is the protein-level quantification for the CL vs NC and RNase + vs - experiments conducted for the OOPS NBT paper (https://www.nature.com/articles/s41587-018-0001-2). The peptide-level abundances have been aggregated to protein level abundance and center-median normalised. Proteins with missing values have been removed. Only proteins quantified in both "total" and "OOPS" samples are included.

The input data here is identical to supplementary table 1 from the above paper. Below we read in the data and extract the ratio between CL vs NC or RNase vs Ctrl from the respective datasheets. Where the protein was not quantified in the NC/Ctrl samples, the ratio is NA and the value in the CL/RNase sample is used instead. This represents a "pseudo value" for the ratio which could not be quantified.

```
glycoproteins <- read.delim('../raw/glycoproteins.tsv') %>% pull(protein)

cl_nc_protein_quant_raw <- readxl::read_excel('../raw/ncbi_30607034_OOPS_NBT_table_S1.xlsx',
                                              sheet=1, skip=1, n_max=2655, na='NA',
                                              col_types = c("text", "numeric", "text", "numeric", "numeric",
  filter(step==3, !master_protein %in% glycoproteins) %>%
  filter(is.finite(CL)) %>%
  mutate(pseudo_CL_NC_Ratio=ifelse(is.na(CL_NC_Ratio), CL, CL_NC_Ratio)) %>%
  select(master_protein, ratio=pseudo_CL_NC_Ratio) %>%
  mutate(exp='CL:NC')


RNase_ctrl_protein_quant_raw <- readxl::read_excel('../raw/ncbi_30607034_OOPS_NBT_table_S1.xlsx',
                                              sheet=3, skip=1, n_max=4307, na='NA',
                                              col_types = c("text", "text", "text", "numeric", "numeric", "nu
                                                            "numeric", "numeric", "numeric", "text")) %>%
  filter(!master_protein %in% glycoproteins, cell_line=='U2OS', Phase=='Org') %>%
  filter(is.finite(RNAse)) %>%
  mutate(pseudo_RNAse_NC_Ratio=ifelse(is.na(RNAse_NC_Ratio), RNAse, RNAse_NC_Ratio)) %>%
  select(master_protein, ratio=pseudo_RNAse_NC_Ratio) %>%
  mutate(exp='RNase:Ctrl')
```

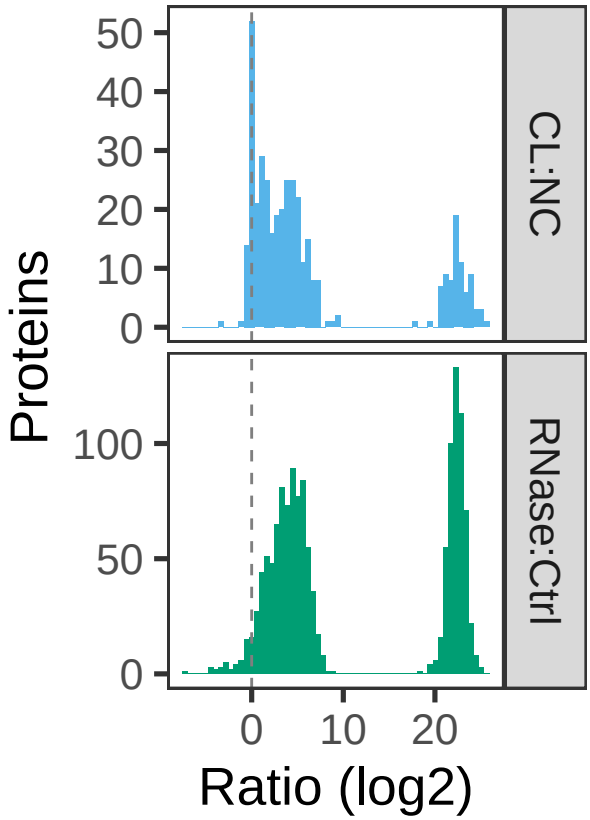
Combine the two experiments

```
combined_cl_rnase <- rbind(RNase_ctrl_protein_quant_raw, cl_nc_protein_quant_raw)
```

Plot

```
ratios_p1 <- combined_cl_rnase %>%
  mutate(exp=factor(exp, levels=c('CL:NC', 'RNase:Ctrl'))) %>%
  ggplot(aes(ratio, fill=exp)) +
  geom_histogram(bins=60) +
  facet_grid(exp~., scales='free_y') +
  xlab('Ratio (log2)') +
  ylab('Proteins') +
  geom_vline(xintercept=0, linetype=2, colour='grey50') +
  scale_fill_manual(values=cbPalette[2:3], guide=FALSE) +
  theme(panel.spacing = unit(0.25, "lines"))

print(ratios_p1)
```



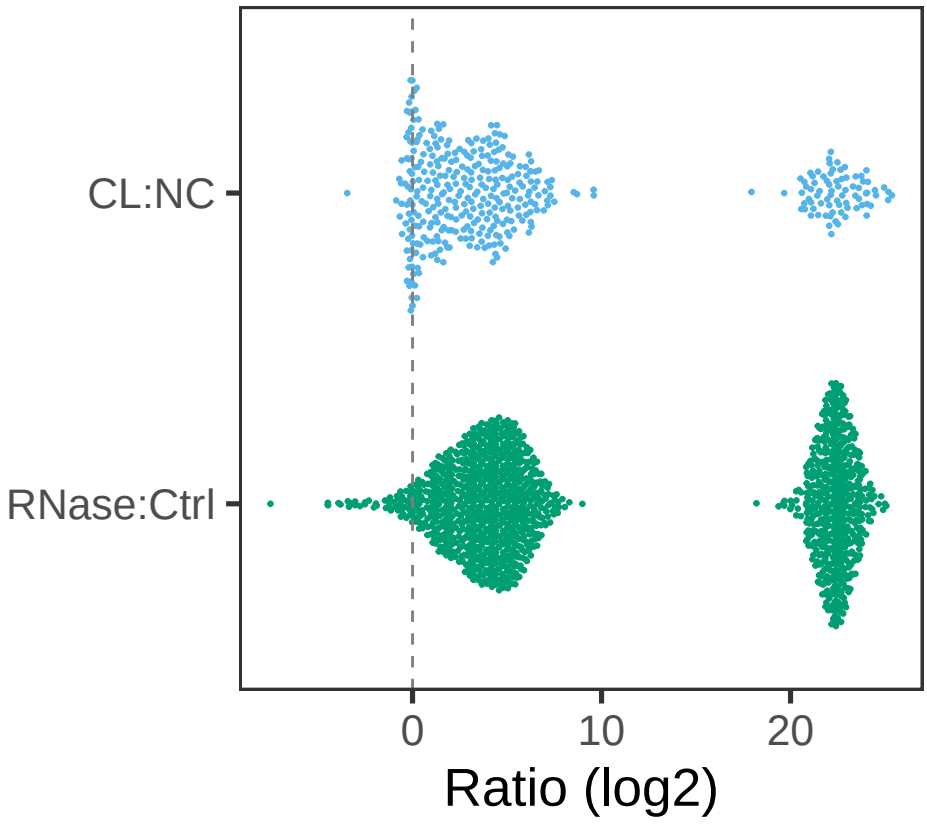

```
ggsave('../results/plots/ratios_p1.png')
```

```
## Saving 6.5 x 4.5 in image
```

```
ratios_p2 <- combined_cl_rnase %>%
  mutate(exp=factor(exp, levels=c('RNase:Ctrl', 'CL:NC'))) %>%
  ggplot(aes(exp, ratio, colour=exp)) +
  geom_quasirandom(size=0.5, bandwidth=0.25) +
  coord_flip() +
  xlab('') +
```

```
  ylab('Ratio (log2)') +
  geom_hline(yintercept=0, linetype=2, colour='grey50') +
  scale_colour_manual(values=cbPalette[c(3,2)], guide=FALSE)

print(ratios_p2)
```




```
ggsave('../results/plots/ratios_p2.png')
```

```
## Saving 6.5 x 4.5 in image
```