Joint Embedding - Pilot Analysis

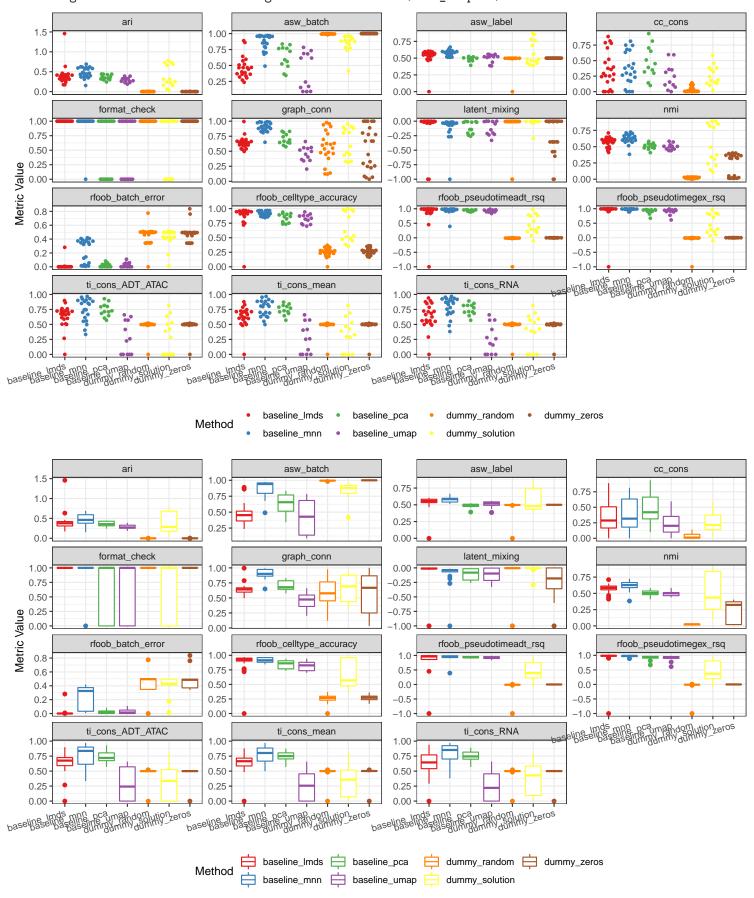
1

Visualise results

Colour by method.

Warning: Removed 22 rows containing missing values (position_quasirandom).

Warning: Removed 22 rows containing non-finite values (stat_boxplot).



Colour by dataset source.

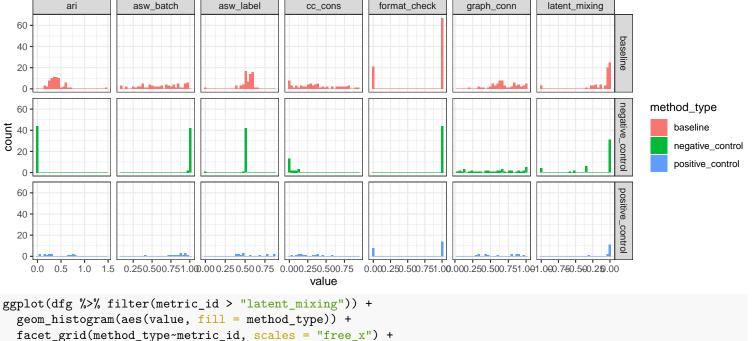
Warning: Removed 22 rows containing missing values (position_quasirandom).

Warning: Removed 22 rows containing non-finite values (stat_boxplot).



Comparing most interesting metrics.

```
# unique(dfg$metric_id)
patchwork::wrap_plots(
     ggplot(dfg %>% filter(metric_id %in% c("ari", "nmi", "asw_label", "ti_cons_mean", "cc_cons", "asw_batch", "gra
          geom_histogram(aes(value, fill = method_type)) +
          facet_grid(method_type~metric_id, scales = "free_x") +
         labs(title = "Metrics included in paper") +
          theme bw(),
     ggplot(dfg %>% filter(!metric_id %in% c("format_check", "ari", "nmi", "asw_label", "ti_cons_mean", "cc_cons",
          geom_histogram(aes(value, fill = method_type)) +
          facet_grid(method_type~metric_id, scales = "free_x") +
          labs(title = "Metrics not included in paper") +
          theme_bw(),
    ncol = 1
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## Warning: Removed 22 rows containing non-finite values (stat_bin).
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
          Metrics included in paper
                                              asw_batch
                                                                            asw_label
                                                                                                                                                                                                 ti_cons_mean
                                                                                                            cc_cons
                                                                                                                                       graph_conn
                                                                                                                                                                           nmi
                                                                                                                                                                                                                                       method_type
                                                                                                                                                                                                                                               baseline
count
                                                                                                                                                                                                                                               negative control
                                                                                                                                                                                                                                               positive_control
         0.0 \quad 0.5 \quad 1.0 \quad 1.5 \quad 0.250.500.751.0 \\ 0.000.250.500.75 \quad 0.000.250.500.75 \quad 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.500.751.0 \\ 0.000.250.750.0 \\ 0.000.250.500.751.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.0 \\ 0.000.250.
                                                                                                             value
          Metrics not included in paper
                                                                                                                                         rfoob_pseudotimeadt_rsq
                    latent mixing
                                                           rfoob batch error
                                                                                               rfoob_celltype_accuracy
                                                                                                                                                                                  rfoob_pseudotimegex_rsq
                                                                                                                                                                                                                                       method_type
                                                                                                                                                                                                                                                baseline
                                                                                                                                                                                                                                                negative_control
                                                                                                                                                                                                                                               positive_control
         -1.00-0.75-0.50-0.25 0.00 0.0 0.2 0.4 0.6 0.8 0.00 0.25 0.50 0.75 1.00-1.0 -0.5 0.0 0.5 1.0 -1.0 -0.5 0.0 0.5
ggplot(dfg %>% filter(metric_id <= "latent_mixing")) +</pre>
     geom_histogram(aes(value, fill = method_type)) +
     facet_grid(method_type~metric_id, scales = "free_x") +
     theme_bw()
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
## Warning: Removed 22 rows containing non-finite values (stat_bin).
```



```
facet_grid(method_type~metric_id, scales = "free_x") +
theme_bw()
```

'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.

