Morphology Trajectories

Niklas Rindtorff

Loading packages

```
library(tidyverse)
library(here)
library(ggrastr)
library(cowplot)
library(scico)
library(ggridges)

umap_df <- read_rds(here::here(params$data))

umap_df_sample <- tibble(file = list.files(here::here(params$sample), full.names = TRUE)) %>% mutate(dr mutate(data = purrr::map(file, ~ readRDS(.x))) %>% dplyr::select(data) %>% unnest(data)

umap_tidy <- rbind(umap_df, umap_df_sample) # for pseudotime fit

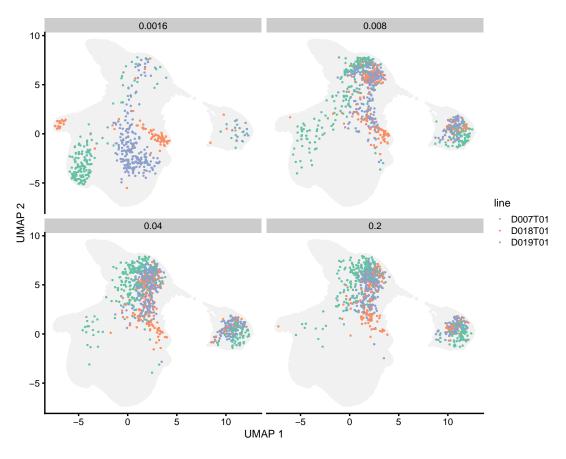
source(here("src/models/misc/pseudotime.R"))</pre>
```

Bortezomib Transition

```
set.seed(123)
gg_bortezomib <- umap_df %>%
  filter(partition %in% c(1,2)) %>%
  dplyr::select(-line, -concentration) %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(alpha = 1, size = 0.35, color = "#f1f1f1") +
  geom_point_rast(data = umap_df_sample %>%
        filter(grepl(drug, pattern = "Bortezomib")) %>%
       filter(line %in% c("D007T01", "D019T01", "D018T01")) %>%
        filter(concentration != "1.0") %>% # removing the maximum concentration as it is not adding fur
    group_by(concentration) %>%
    sample_n(10000, replace = TRUE),
  aes(color = line),alpha = 1, size = 0.35, shape=16) +
  scale_color_brewer(type = "qual", palette = "Set2") +
  facet_wrap( ~ concentration, ncol = 2) +
  labs(x = "UMAP 1",
```

```
y = "UMAP 2")+
theme(legend.position = "none") +
theme_cowplot(font_size = 8) +
coord_fixed()

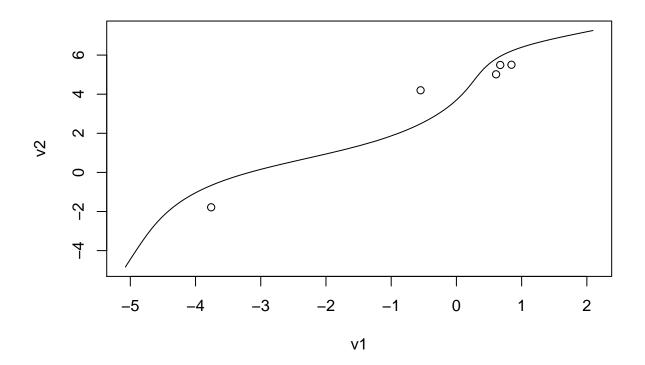
gg_bortezomib
```



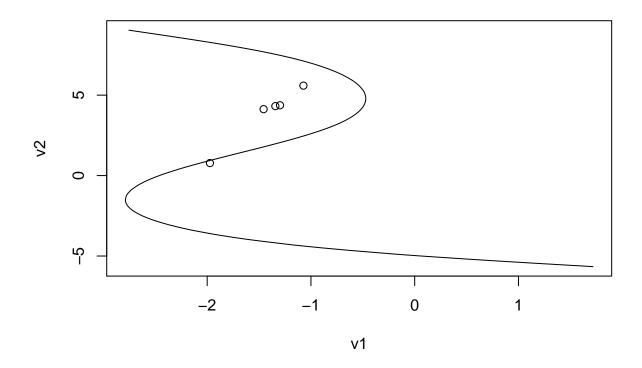
I am generating a pseudotime plot for Bortezomib, a toxic drug

```
## Starting curve---distance^2: 6923.961
## Iteration 1---distance^2: 3231.457
## Iteration 2---distance^2: 2838.282
## Iteration 3---distance^2: 2739.663
## Iteration 4---distance^2: 2697.724
## Iteration 5---distance^2: 2664.242
## Iteration 6---distance^2: 2635.131
## Iteration 7---distance^2: 2611.228
```

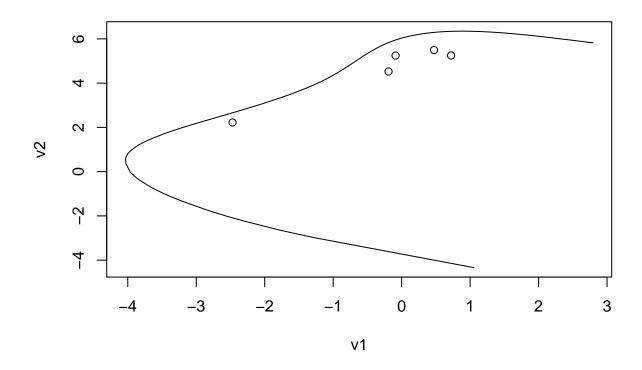
```
## Iteration 8---distance^2: 2591.754
## Iteration 9---distance^2: 2575.216
## Iteration 10---distance^2: 2561.11
## Iteration 11---distance^2: 2548.819
## Iteration 12---distance^2: 2538.777
## Iteration 13---distance^2: 2530.808
## Iteration 14---distance^2: 2524.718
## Iteration 15---distance^2: 2520.287
## Iteration 16---distance^2: 2517.269
## Iteration 17---distance^2: 2515.212
```



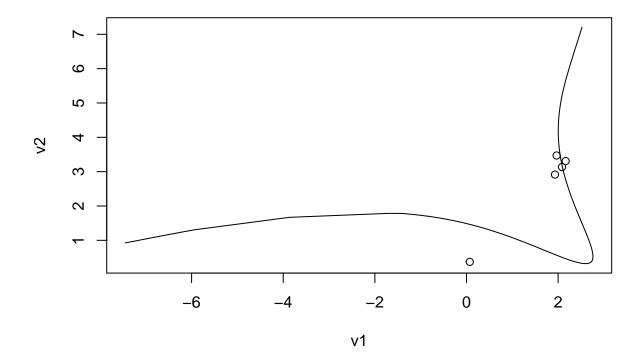
```
## Starting curve---distance^2: 22320.49
## Iteration 1---distance^2: 7311.392
## Iteration 2---distance^2: 4366.299
## Iteration 3---distance^2: 3281.056
## Iteration 4---distance^2: 3193.239
## Iteration 5---distance^2: 3042.677
## Iteration 6---distance^2: 2665.67
## Iteration 7---distance^2: 2560.964
## Iteration 8---distance^2: 2525.297
## Iteration 9---distance^2: 2509.164
## Iteration 10---distance^2: 2497.064
## Iteration 11---distance^2: 2497.064
## Iteration 12---distance^2: 2494.993
```



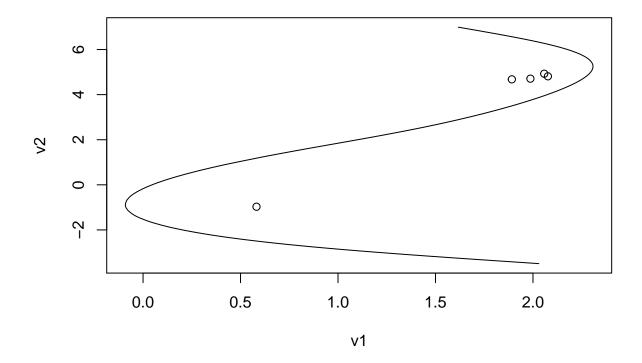
```
## Starting curve---distance^2: 9258.826
## Iteration 1---distance^2: 4286.348
## Iteration 2---distance^2: 3625.018
## Iteration 3---distance^2: 3287.906
## Iteration 4---distance^2: 2733.201
## Iteration 5---distance^2: 2486.007
## Iteration 6---distance^2: 2402.714
## Iteration 7---distance^2: 2350.676
## Iteration 8---distance^2: 2294.797
## Iteration 9---distance^2: 2250.427
## Iteration 10---distance^2: 2249.56
```



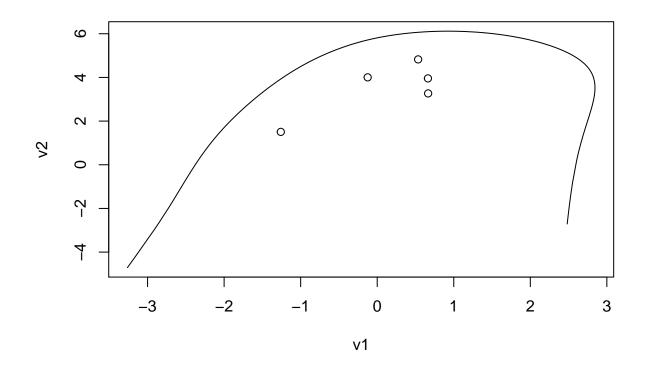
```
## Starting curve---distance^2: 17818.75
## Iteration 1---distance^2: 3879.162
## Iteration 2---distance^2: 1634.857
## Iteration 3---distance^2: 1587.207
## Iteration 4---distance^2: 1605.169
## Iteration 5---distance^2: 1633.437
## Iteration 6---distance^2: 1658.009
## Iteration 7---distance^2: 1677.896
## Iteration 8---distance^2: 1694.304
## Iteration 9---distance^2: 1710.728
## Iteration 10---distance^2: 1730.266
## Iteration 11---distance^2: 1745.779
## Iteration 12---distance^2: 1747.044
```

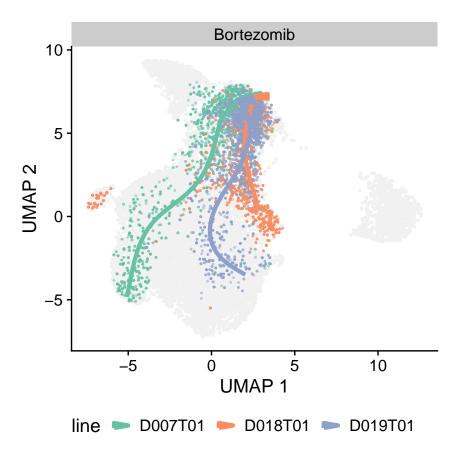


```
## Starting curve---distance^2: 4910.536
## Iteration 1---distance^2: 2387.613
## Iteration 2---distance^2: 2020.734
## Iteration 3---distance^2: 1914.154
## Iteration 4---distance^2: 1870.209
## Iteration 5---distance^2: 1842.44
## Iteration 6---distance^2: 1819.73
## Iteration 7---distance^2: 1800.767
## Iteration 8---distance^2: 1785.536
## Iteration 9---distance^2: 1773.615
## Iteration 10---distance^2: 1763.846
## Iteration 11---distance^2: 1755.139
## Iteration 12---distance^2: 1746.564
## Iteration 13---distance^2: 1737.674
## Iteration 14---distance^2: 1728.247
## Iteration 15---distance^2: 1718.655
## Iteration 16---distance^2: 1709.725
## Iteration 17---distance^2: 1703.35
## Iteration 18---distance^2: 1699.688
## Iteration 19---distance^2: 1697.722
## Iteration 20---distance^2: 1697.108
```



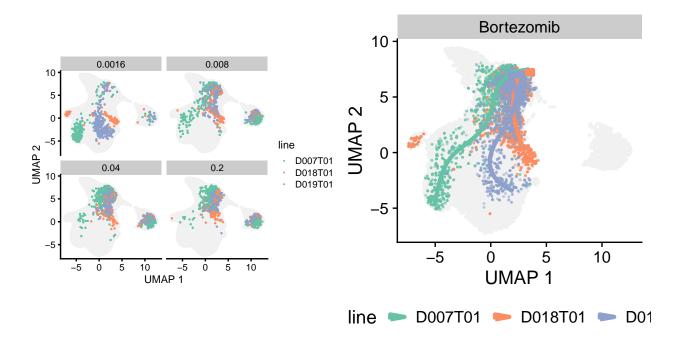
```
## Starting curve---distance^2: 12610.93
## Iteration 1---distance^2: 5254.146
## Iteration 2---distance^2: 3779.518
## Iteration 3---distance^2: 3471.257
## Iteration 4---distance^2: 3439.146
## Iteration 5---distance^2: 3566.97
## Iteration 6---distance^2: 3806.137
## Iteration 7---distance^2: 4077.759
## Iteration 8---distance^2: 4394.948
## Iteration 9---distance^2: 4777.317
## Iteration 10---distance^2: 5240.603
## Iteration 11---distance^2: 5210.229
## Iteration 12---distance^2: 3239.859
## Iteration 13---distance^2: 2959.096
## Iteration 14---distance^2: 2897.409
## Iteration 15---distance^2: 2882.383
## Iteration 16---distance^2: 2886.99
## Iteration 17---distance^2: 2889.456
```





e-7 good

A B



Irinotecan Transition

I am repeating a similar plot with SN-38, the active metabolite of Irinotecan.

```
set.seed(123)
gg_irinotecan <- umap_df %>%
  filter(partition %in% c(1,2)) %>%
  dplyr::select(-line, -concentration) %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(alpha = 1, size = 0.35, color = "#f1f1f1") +
  geom_point_rast(data = umap_tidy %>%
        filter(grepl(drug, pattern = "Irinotecan / SN-38")) %>%
filter(line %in% c("D007T01", "D019T01", "D018T01")) %>%
        filter(concentration != "1.0") %>% # removing the maximum concentration as it is not adding fur
    group_by(concentration) %>%
    sample_n(1000, replace = TRUE),
  aes(color = line),alpha = 1, size = 0.35, shape=16) +
  scale_color_brewer(type = "qual", palette = "Set2") +
  facet_wrap( ~ concentration, ncol = 2) +
  labs(x = "UMAP 1",
       y = "UMAP 2")+
  theme(legend.position = "none") +
  theme_cowplot(font_size = 8) +
```

```
coord_fixed()
gg_irinotecan
```

I am generating a pseudotime plot for Irinotecan, a toxic drug

5 major drug induced phenotypes

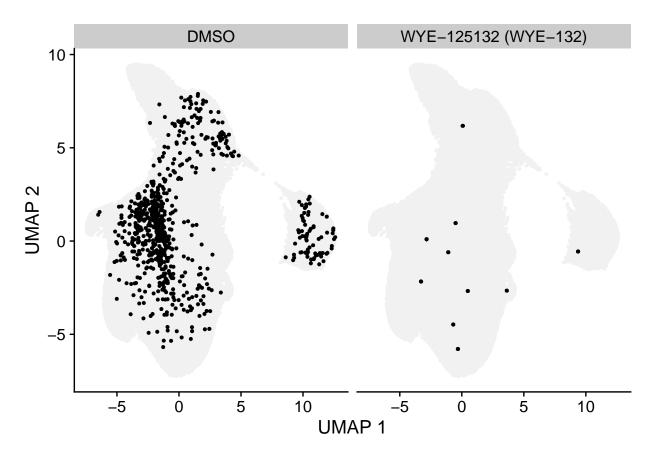
```
set.seed(123)
anno <- read_delim(here("references/layouts/Layouts.csv"), ";", escape_double = FALSE, trim_ws = TRUE)
 dplyr::rename(drug = product_name) %>%
 filter(library_id != 1) %>%
 dplyr::select(drug, target, pathway) %>%
 distinct()
anno = anno %>% mutate(group = case_when(grepl(pattern = "CDK", x= target) ~ "CDK",
                                        grepl(pattern = "mTOR", x= target) ~ "mTOR",
                                        grepl(pattern = "MEK", x= target) ~ "MEK",
                                        grepl(pattern = "EGFR", x= target) ~ "EGFR",
                                        grepl(pattern = "GSK", x= target) ~ "GSK",
                                        drug == "DMSO" ~ "CTRL",
                                        TRUE ~ "other"
                                        ))
activity = umap_tidy %>% distinct(line, drug) %>%
  nest(-line) %>%
  mutate(activity = purrr::map(line, ~ read_csv(here::here( "data/interim/FeatureAnalysis/drug_effects/
                                        janitor::clean_names() %>%
                                        dplyr::rename(drug = x1)
```

```
) %>%
  dplyr::select(-data) %>%
  unnest(activity)
gg_big5_density <- umap_df %>%
  filter(partition %in% c(1,2)) %>%
  dplyr::select(-line, -concentration) %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(alpha = 1, size = 0.35, color = "#f1f1f1") +
  stat_density_2d(geom = "polygon",
                  data = umap_tidy %>%
        semi_join(anno %>% filter(group != "other")) %>%
       left_join(anno) %>%
        semi_join(activity %>% filter(auc_mean >= 0.85 | drug == "DMSO")) %>%
       filter(line %in% c("D004T01", "D030T01")) %>% # "D007T01", "D030T01", "D027T01", "D046T01" # 4
        filter(concentration == "nan") %>% # keeping single dose drugs
        filter(group %in% c("DMSO", "MEK", "mTOR", "CDK")) %>%
    group_by(drug, line) %>%
    sample_n(1000, replace = TRUE),
  aes(alpha = ..level.., fill = group)) +
  scale_fill_brewer(type = "qual", palette = "Set2") +
  facet_grid(line ~ group) + # ~ line
  labs(x = "UMAP 1",
      y = "UMAP 2")+
  theme(legend.position = "none") +
  theme_cowplot(font_size = 8) +
  coord_fixed()
gg_mekgsk_density <- umap_df %>%
  filter(partition %in% c(1,2)) %>%
  dplyr::select(-line, -concentration) %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(alpha = 1, size = 0.35, color = "#f1f1f1") +
  stat_density_2d(geom = "polygon",
                  data = umap_tidy %>%
        semi_join(anno %>% filter(group != "other")) %>%
       left_join(anno) %>%
        semi_join(activity %>% filter(auc_mean >= 0.85 | drug == "DMSO")) %>%
       filter(line %in% c("D004T01", "D007T01", "D019T01")) %>% # "D007T01", "D030T01", "D027T01", "D0
       filter(concentration == "nan") %>% # keeping single dose drugs
        filter(group %in% c("DMSO", "MEK", "GSK")) %>%
    group_by(drug, line) %>%
    sample_n(1000, replace = TRUE),
  aes(alpha = ..level.., fill = group)) +
  scale_fill_brewer(type = "qual", palette = "Set2") +
  facet_grid(line ~ group) + # ~ line
  labs(x = "UMAP 1",
      y = "UMAP 2")+
  theme(legend.position = "none") +
  theme_cowplot(font_size = 8) +
```

```
coord_fixed()
```

Unlabeled

```
gg_mtor_all <- umap_df %>% rbind(umap_df_sample) %>%
  filter(partition %in% c(1,2)) %>%
  filter(drug == "DMSO" | grepl(drug, pattern = "WYE-132")) %>%
  #filter(line %in% c("D030T01", "D027T01", "D013T01")) %>%
  filter(line %in% c("D027T01")) %>%
  group_by(drug) %>%
  sample_n(700, replace = TRUE) %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(data = umap_df %>% dplyr::select(-drug), alpha = 1, size = 0.35, color = "#f1f1f1") +
  geom_point_rast(aes(color = drug),alpha = 1, size = 1, shape=16, color = "black") +
  facet_wrap(~ drug) +
  theme_cowplot() +
  labs(x = "UMAP 1",
      y = "UMAP 2")+
  theme(legend.position = "nothing")
gg_mtor_all
```



```
set.seed(123)
drug_order <- c("CHIR-98014", "DMSO")</pre>
center_df <- umap_tidy %>%
   filter(drug %in% drug_order) %>%
 filter(partition %in% c(1,2)) %>%
  #filter(line %in% c("D030T01", "D027T01", "D013T01")) %>%
  group_by(drug, line, concentration) %>%
  sample_n(500, replace = TRUE) %>% ungroup()
gg_gsk_all <- umap_tidy %>%
   filter(partition %in% c(1,2)) %>%
  filter(drug == "DMSO" | grepl(drug, pattern = drug_order[1])) %>%
  #filter(line %in% c("D030T01", "D027T01", "D013T01")) %>%
  group_by(drug) %>%
  sample_n(500) %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(data = umap_df %>% dplyr::select(-drug), alpha = 1, size = 0.35, color = "#f1f1f1") +
 geom_point_rast(aes(color = drug),alpha = 1, size = 1, shape=16, color = "black") +
```

```
facet_wrap(~ drug) +

theme_cowplot() +
labs(x = "UMAP 1",
    y = "UMAP 2")+

theme(legend.position = "nothing")
```

```
set.seed(123)
drug_order <- c("Ulixertinib (BVD-523, VRT752271)",</pre>
                "TAK-733",
                "Trametinib (GSK1120212)",
                "Cobimetinib (GDC-0973, RG7420)",
                "MEK162 (ARRY-162, ARRY-438162)",
                "PD0325901",
                "Refametinib (RDEA119, Bay 86-9766)",
                "AZD8330",
                "Pimasertib (AS-703026)",
                "Selumetinib (AZD6244)",
                "PD318088",
                "VX-11e",
                "DMSO")
center_df <- umap_tidy %>% filter(concentration == "nan") %>%
   filter(drug %in% drug_order) %>%
 filter(partition %in% c(1,2)) %>%
  group_by(drug, line, concentration) %>%
  sample_n(500, replace = TRUE) %>%
  summarise(v1 = mean(v1),
            v2 = mean(v2)) \%
  ungroup()
center_df %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(data = umap_df %>% dplyr::select(-drug, -line), alpha = 1, size = 0.35, color = "#f1f
  geom_point_rast(aes(color = drug),alpha = 1, size = 1, shape=16) +
 facet_wrap(~ line) +
 theme_cowplot() +
 labs(x = "UMAP 1",
       y = "UMAP 2")+
 theme(legend.position = "nothing")
```

Differential drug activity

I am focusing on a set of two lines, D019T01 and D004T01

```
i_dist_in = 0.05
j_nn_in = 30
k_res_in = .1
```

```
loi = "D019T01"
obj <- readRDS(here::here(paste0("data/processed/PhenotypeSpectrum/drug_effects/lm_umap", loi, "_", i_d
umap_tidy_drug <- reducedDims(obj)$UMAP %>% cbind(colData(obj),
                                                  cluster = clusters(obj),
                                                  partition = partitions(obj)) %>% as_tibble() %>% jani
                                                    mutate(line = loi)
loi = "D004T01"
obj <- readRDS(here::here(paste0("data/processed/PhenotypeSpectrum/drug_effects/lm_umap", loi, "_", i_d
umap_tidy_drug <- rbind(umap_tidy_drug, reducedDims(obj)$UMAP %>% cbind(colData(obj),
                                                  cluster = clusters(obj),
                                                  partition = partitions(obj)) %>% as_tibble() %>% jani
anno <- read_delim(here("references/layouts/Layouts.csv"), ";", escape_double = FALSE, trim_ws = TRUE)
  dplyr::rename(drug = product_name) %>%
  filter(library_id != 1) %>%
  dplyr::select(drug, target, pathway) %>%
  distinct()
diff_drug <- umap_tidy_drug %>%
 left_join(anno) %>%
  drop na() %>%
  mutate(target = case_when(grepl(pattern = "src", target, ignore.case = TRUE) == TRUE &
                           grepl(pattern = "Aurora", target, ignore.case = TRUE) == FALSE ~ "Src",
                           TRUE ~ target))
gg_target <- diff_drug %>%
  #filter(auc_mean > 0.85 | drug == "DMSO") %>%
  ggplot(aes(v1, v2, text = target)) +
  geom_point(color = "grey") +
  geom_point(data = df %>% filter(target %in% c("PLK", "mTOR", "Src")), aes(color = target), size = 2)
  ggforce::geom_mark_ellipse(data = diff_drug %>% filter(target %in% c("mTOR", "PLK"), cluster %in% c(5
  ggforce::geom_mark_ellipse(data = diff_drug %% filter(target %in% c("mTOR", "Src"), cluster %in% c(2
  scale_color_brewer(type = "qual", palette = "Set2") +
  theme_cowplot() +
  facet_wrap(~ line) +
  labs(x = "UMAP1",
      y = "UMAP2",
       caption = "active drugs plotted by separating plane, selected drugs are highlighted")
gg_target
gg_target_density <- umap_df %>%
  filter(partition %in% c(1,2)) %>%
  dplyr::select(-line, -concentration) %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(alpha = 1, size = 0.35, color = "#f1f1f1") +
  stat_density_2d(geom = "polygon",
                  data = umap_tidy %>%
```

```
left_join(diff_drug %>% dplyr::select(drug, target)) %>%
       filter(line %in% c("D004T01", "D019T01")) %>%
        filter(target %in% c("mTOR", "PLK", "Src") | drug == "DMSO") %>%
          # removing reference for clarity
          filter(target %in% c("mTOR", "PLK", "Src")) %>%
        mutate(target = ifelse(drug == "DMSO", "CTRL", target)) %>%
        filter(concentration == "nan") %>% # keeping single dose drugs
    group by(drug) %>%
    sample_n(500, replace = TRUE),
  aes(alpha = ..level.., fill = target)) +
  scale_fill_brewer(type = "qual", palette = "Set2") +
  facet_wrap( ~ line, ncol = 2) +
  labs(x = "UMAP 1",
      y = "UMAP 2")+
  theme(legend.position = "none") +
  theme_cowplot(font_size = 8) +
  coord_fixed()
gg_target_density
plot_grid(gg_target, gg_target_density,
     labels = c('A', 'B'), label_size = 12, ncol = 1) +
  ggsave(here::here("reports/enrichment_drugtarget.pdf"), width = 4, height = 4)
```

MEK inhibitor Pseudotimes

```
set.seed(1323)
plot_df_raw <- create_plot_df(drug_order = c("Binimetinib", "Trametinib"),</pre>
                          loi = c("D027T01", "D019T01"),
                          n_{sample} = 500,
                          v1_cut = 5)
plot_df <- plot_df_raw %>%
  mutate(plot_trace = case_when(TRUE ~ purrr::map(plot, ~ .x$fit)))
gg_mek_traj_sel <- draw_trajectory_bulk(plot_df) + coord_fixed()</pre>
gg_mek_traj_sel
traj_df <- plot_df %>%
  mutate(proj = purrr::map2(data, plot_trace, ~ project_to_curve(.x %>% dplyr::select(v1, v2) %>% as.ma
                                                                   .y %>% dplyr::select(v1, v2) %>% as.ma
         threshold = purrr::map(proj, ~ median(.x$dist_ind) + 0*sd(.x$dist_ind)),
         df = purrr::map(proj, ~ cbind(dist = .x$dist_ind, ord = .x$ord, lambda = .x$lambda) %>% as_tib
  )
traj_df <- traj_df %>% dplyr::select(-plot, -plot_trace, -proj) %>% unnest(df, threshold, data)
```

```
traj_df %>%
  ggplot(aes(v1, v2, color = dist <= threshold)) +</pre>
  geom_point_rast() +
  scale color viridis d() +
 theme_cowplot() +
 facet_grid(line ~ drug) +
  coord fixed() +
  labs(x = "UMAP 1",
       y = "UMAP 2")
doi = "Trametinib"
loi = "D027T01"
gg_traj_incl <- traj_df %>%
  filter(drug == doi & line == loi) %>%
  #filter(lambda >= 5) %>% # hyperparameter tuned to avoid undershooting treatment effects
 filter(dist <= threshold) %>%
  ggplot(aes(v1, v2)) +
  geom_point_rast(data = umap_df %>% dplyr::select(-line, -concentration), alpha = 1, size = 0.35, color
  geom_point(aes(color = lambda)) +
  scale_color_viridis_c() +
  theme cowplot() +
  facet_grid(~ line) +
  theme(legend.position = "bottom") +
  coord_fixed() +
  labs(x = "UMAP 1",
       y = "UMAP 2")
gg_size_pc <- traj_df %>%
  filter(drug == doi & line == loi) %>%
 #filter(lambda >= 5) %>% # hyperparameter tuned to avoid undershooting treatment effects
 ggplot(aes(lambda, size)) +
  #geom_point() +
  #scale_color_viridis_d() +
  theme_cowplot() +
  geom_smooth(color = "black") +
  geom_vline(xintercept = 8, linetype= "dashed") +
  geom_vline(xintercept = 9.5, linetype= "dashed") +
  labs(x = "lambda",
       y = "size [pixels]")
gg_conc_pc <- traj_df %>%
  filter(drug == doi & line == loi) %>%
 # filter(lambda >= 5) %>% # hyperparameter tuned to avoid undershooting treatment effects
  ggplot(aes(lambda, fill = concentration)) +
  geom_density(alpha = 0.2, position = "fill") +
  theme_cowplot() +
  scale_fill_viridis_d() +
  theme(legend.position = "bottom") +
  facet grid(~ drug) +
  geom_vline(xintercept = 8, linetype= "dashed") +
```

Plot Export

```
gg_bortezomib + ggsave(here::here("reports/figures/gg_bortezomib.pdf"), width = 2, height = 8)

gg_mtor_all + ggsave(here::here("reports/figures/gg_mtor.pdf"), width = 6, height = 3)

gg_mtor_all + ggsave(here::here("reports/figures/gg_gsk.pdf"), width = 6, height = 3)

gg_mek_traj_sel + ggsave(here::here("reports/figures/gg_mek_traj_sel.pdf"), width = 4, height = 4)

gg_bini_traj_sel + ggsave(here::here("reports/figures/gg_bini_traj_sel.pdf"), width = 4, height = 4)

gg_tram_traj + ggsave(here::here("reports/figures/gg_tram_traj.pdf"), width = 4, height = 4)

gg_big5_density + ggsave(here::here("reports/figures/gg_big5_density.pdf"), width = 16, height = 8)

gg_mekgsk_density + ggsave(here::here("reports/figures/gg_mekgsk_density.pdf"), width = 12, height = 12
```

Supplement

Here I collect pieces of code that did not make it into the final analysis but can be run in theory. In order to access these peaces of code, you have to open the *.RMD* file.

```
knitr::knit_exit()
```