Quantification methods benchmarking

Source

- Hyun-Hwan Jeong's google drive, need to have the access permission

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Conflicts with tidy packages ----------------------------------------------
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```

Summary of data

```
## Parsed with column specification:
## cols(
##   data_plot_name = col_character(),
##   library = col_character(),
##   length_gRNA = col_integer(),
##   num_gRNAs = col_number(),
##   data_db_id = col_character(),
##   data_name = col_character(),
##   total = col_number()
## )
```

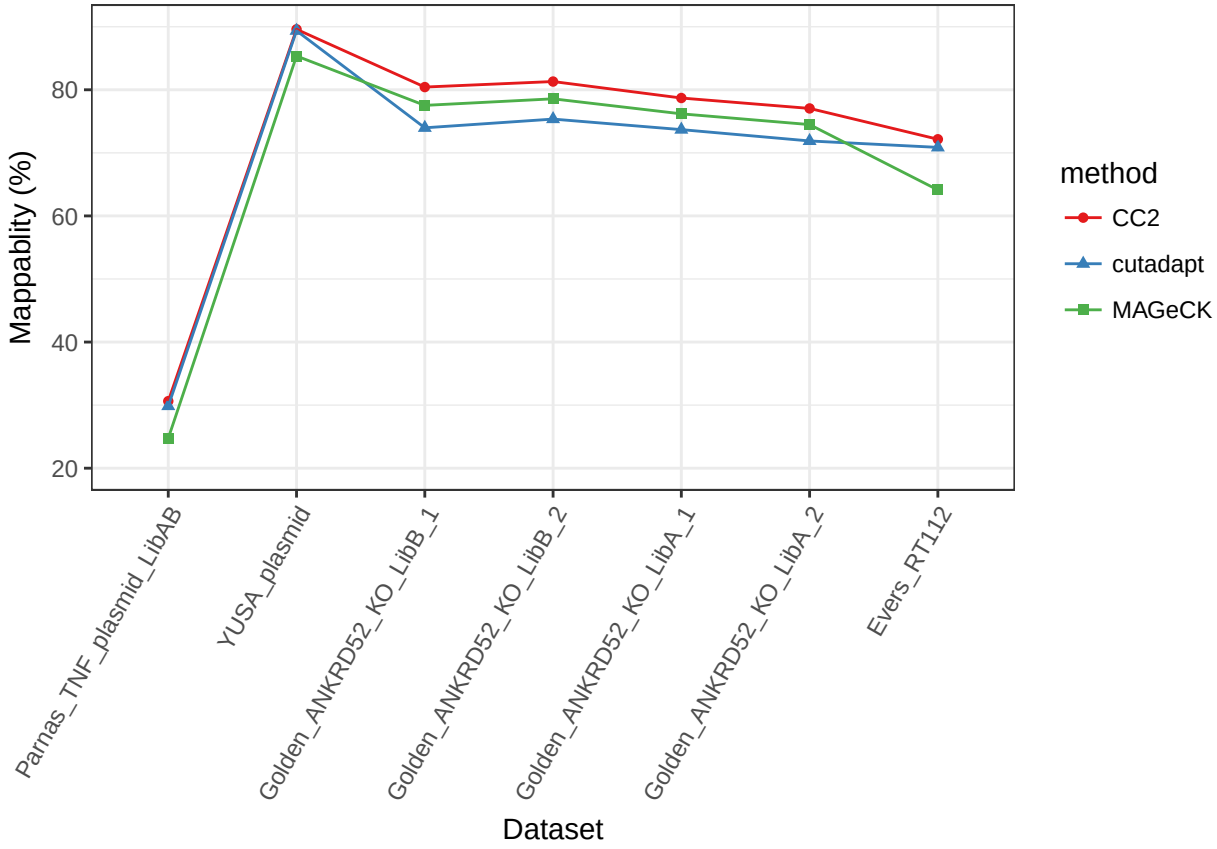
```
## # A tibble: 7 x 5
##             data_plot_name length_gRNA num_gRNAs         data_name
##                      <chr>       <int>     <dbl>             <chr>
## 1 Parnas_TNF_plasmid_LibAB          20    130209     plasmid_LibAB
## 2             YUSA_plasmid          19     87437           plasmid
## 3 Golden_ANKRD52_KO_LibB_1          20     58028 ANKRD52_KO_LibB_1
## 4 Golden_ANKRD52_KO_LibB_2          20     58028 ANKRD52_KO_LibB_2
## 5 Golden_ANKRD52_KO_LibA_1          20     65383 ANKRD52_KO_LibA_1
## 6 Golden_ANKRD52_KO_LibA_2          20     65383 ANKRD52_KO_LibA_2
## 7              Evers_RT112          20       961      CRISPR_RT112
## # ... with 1 more variables: total <dbl>
```

Mappability

```
read_delim("mappability.tsv", "\t") %>%
  gather(method, mappability, -data_plot_name) %>%
  ggplot(aes(x=data_plot_name, y=mappability*100)) +
  xlim(df.summary$data_plot_name) +
  #geom_bar(aes(fill=method), stat = "identity", position="dodge") +
  geom_point(aes(colour=method, shape=method)) +
```

```
  geom_line(aes(colour=method, group=method)) +
  scale_colour_brewer(palette = "Set1") +
  theme_bw() +
  theme(axis.text.x = element_text(angle=60, hjust=1)) + ylim(20,90) +
  ylab("Mappablity (%)") + xlab("Dataset")
```

```
## Parsed with column specification:
## cols(
##   data_plot_name = col_character(),
##   CC2 = col_double(),
##   MAGeCK = col_double(),
##   cutadapt = col_double()
## )
```



Run Time

```
read_delim("runtime.tsv", "\t") %>%
  gather(method, runtime, -data_plot_name) %>%
  ggplot(aes(x=data_plot_name, y=runtime)) +
  xlim(df.summary$data_plot_name) +
  #geom_bar(aes(fill=method), stat = "identity", position="dodge") +
  geom_point(aes(colour=method, shape=method)) +
  geom_line(aes(colour=method, group=method)) +
  scale_color_brewer(palette = "Set1") +
  theme_bw() +
```

```
  theme(axis.text.x = element_text(angle=60, hjust=1)) +
  ylab("Run time (seconds)") + xlab("Dataset")
```

```
## Parsed with column specification:
## cols(
##   data_plot_name = col_character(),
##   CC2 = col_double(),
##   MAGeCK = col_double(),
##   cutadapt = col_double()
## )
```