Barcoding primers for pooling multiple individuals

Meinolf Ottensmann

Preface

This document provides details on barcoding individual samples using indexed primer pairs. Details on the development of target-specific primer pairs are given in files **01_Primer_Design**.

```r
library(plyr)
library(dplyr)
library(seqinr)
library(rex)
library(ShortRead)
```

Assing 82 samples and 6 replicates to 11 rows of a 96-well rack

All 82 samples (41 mother-pup pairs) are randomly assigned to a location on 96-well rack. Additionally, six replicates are included.

```r
# read sample names
names <- read.csv("data/mhc_sample_names.csv", header = T,
    sep = ";")
names <- names[, c("ID", "Pair")]
names <- unique(names)
row.names(names) <- as.character(1:nrow(names))

# Pseudo-random order of samples Note, this order
# will change when execute on another computer
set.seed(99)
coins <- sample(1:nrow(names), size = nrow(names),
    replace = F)

# create rack locations
R <- 1:11
C <- LETTERS[1:8]
p <- character()
for (r in 1:length(R)) {
    for (c in 1:length(C)) {
        p <- c(p, paste0(C[c], as.character(R[r])))
    }
}

df <- data.frame(id = names[["ID"]][coins], pair = names[["Pair"]][coins],
    no = coins, pos = p[1:82])

## Last column are replicates
replicates <- c("W8914mum", "W8928pup", "W8933pup",
    "W8935mum", "W8915pup", "AGF11007")
coins2 <- unlist(lapply(X = replicates, FUN = function(a) which(df[["id"]] ==
    a)))
```

```r
df2 <- df[coins2, 1:3]
df2[["pos"]] <- p[83:88]

df <- rbind(df, df2)
df[["age"]] <- unlist(lapply(X = df[["pair"]], FUN = function(a) substring(a,
    1, 1)))
df[["pair_id"]] <- unlist(lapply(X = df[["pair"]],
    FUN = function(a) substring(a, 2)))
```

Add barcodes

Index sequences of the Illumina Nextera_XT kit (Illumina 2017) are used as barcodes. 12 indices may be assigned to each of the rack´s rows, while 8 indices are used for each of the columns. Note, DRB and DQB use slightly different combinations.

```r
# read barcodes and assing to forward and reverse
barcodes <- read.csv2("data/mhc_barcodes.csv", skip = 1)
f <- barcodes[barcodes[["primer"]] == "Forward", 1:2]
r <- barcodes[barcodes[["primer"]] == "Reverse", 1:2]

dqb <- data.frame(Index_F = rep(f$Index[1:11], each = 8),
    Index_R = rep(r$Index, 11), Sequence_F = rep(f$Sequence[1:11],
        each = 8), Sequence_R = rep(r$Sequence, 11))

drb <- data.frame(Index_F = rep(f$Index[c(1, 2, 3,
    4, 5, 6, 7, 12, 9, 10, 11)], each = 8), Index_R = rep(r$Index,
    11), Sequence_F = rep(f$Sequence[c(1, 2, 3, 4,
    5, 6, 7, 12, 9, 10, 11)], each = 8), Sequence_R = rep(r$Sequence,
    11))

df <- df[, c("id", "age", "pair_id", "pos")]

dqb <- cbind(df, dqb)
drb <- cbind(df, drb)

write.csv(dqb, file = "documents/dqb_table.csv", row.names = F)
write.csv(dqb, file = "documents/dqa_table.csv", row.names = F)
write.csv(dqb, file = "documents/dra_table.csv", row.names = F)
write.csv(drb, file = "documents/drb_table.csv", row.names = F)
```

```r
sessionInfo()
> R version 3.4.3 (2017-11-30)
> Platform: x86_64-w64-mingw32/x64 (64-bit)
> Running under: Windows 10 x64 (build 16299)
>
> Matrix products: default
>
> locale:
> [1] LC_COLLATE=English_United Kingdom.1252
> [2] LC_CTYPE=English_United Kingdom.1252
> [3] LC_MONETARY=English_United Kingdom.1252
> [4] LC_NUMERIC=C
```

```
> [5] LC_TIME=English_United Kingdom.1252
>
> attached base packages:
> [1] stats     graphics  grDevices utils
> [5] datasets  methods   base
>
> other attached packages:
> [1] knitr_1.17
>
> loaded via a namespace (and not attached):
>  [1] compiler_3.4.3  backports_1.1.2
>  [3] magrittr_1.5    rprojroot_1.3-1
>  [5] formatR_1.5     tools_3.4.3
>  [7] htmltools_0.3.6 yaml_2.1.16
>  [9] Rcpp_0.12.14    stringi_1.1.6
> [11] rmarkdown_1.8   stringr_1.2.0
> [13] digest_0.6.13   evaluate_0.10.1
```

References

Illumina, Inc. 2017. "Oligonucleotide Sequences © 2017 Illumina, Inc. All Rights Reserved. Derivative Works Created by Illumina Customers Are Authorized for Use with Illumina Instruments and Products Only . All Other Uses Are Strictly Prohibited. Document # 1000000002694 V03." https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/experiment-design/illumina-adapter-sequences-1000000002694-03.pdf.