

Comparison allelic diversity DQB and DRB

Meinolf Ottensmann

```
library(magrittr)
library(ggplot2)
```

Comparison of nucleotide diversity using p-distances

```
dqb_exon <- read.table("Phylogeny/p-distances/ArGa-DQB-Exon.txt",
                      skip = 21, sep = ";", nrows = 153, header = T)[["Distance"]]
  as.numeric()
> numeric(0)

drb_exon <- read.table("Phylogeny/p-distances/ArGa-DRB-Exon.txt",
                      skip = 21, sep = ";", nrows = 36, header = T)[["Distance"]]

t.test(dqb_exon, drb_exon)
>
> Welch Two Sample t-test
>
> data: dqb_exon and drb_exon
> t = 6.3855, df = 72.758, p-value = 1.419e-08
> alternative hypothesis: true difference in means is not equal to 0
> 95 percent confidence interval:
> 0.02408221 0.04593739
> sample estimates:
> mean of x mean of y
> 0.10639869 0.07138889

dqb_PBR <- read.table("Phylogeny/p-distances/ArGa-DQB-PBR.txt",
                      skip = 21, sep = ";", nrows = 153, header = T)[["Distance"]]

drb_PBR <- read.table("Phylogeny/p-distances/ArGa-DRB-PBR.txt",
                      skip = 21, sep = ";", nrows = 36, header = T)[["Distance"]]

t.test(dqb_PBR, drb_PBR)
>
> Welch Two Sample t-test
>
> data: dqb_PBR and drb_PBR
> t = 4.3462, df = 71.327, p-value = 4.512e-05
> alternative hypothesis: true difference in means is not equal to 0
> 95 percent confidence interval:
> 0.04040014 0.10888091
> sample estimates:
> mean of x mean of y
> 0.2727516 0.1981111

dqb_nonPBR <- read.table("Phylogeny/p-distances/ArGa-DQB-non-PBR.txt",
                        skip = 21, sep = ";", nrows = 153, header = T)[["Distance"]]
```

```

drb_nonPBR <- read.table("Phylogeny/p-distances/ArGa-DRB-non-PBR.txt",
                        skip = 21, sep = ";", nrow = 36, header = T)[["Distance"]]

t.test(dqb_nonPBR, drb_nonPBR)
>
> Welch Two Sample t-test
>
> data: dqb_nonPBR and drb_nonPBR
> t = 7.2637, df = 72.582, p-value = 3.463e-10
> alternative hypothesis: true difference in means is not equal to 0
> 95 percent confidence interval:
> 0.02420070 0.04250518
> sample estimates:
> mean of x mean of y
> 0.06057516 0.02722222

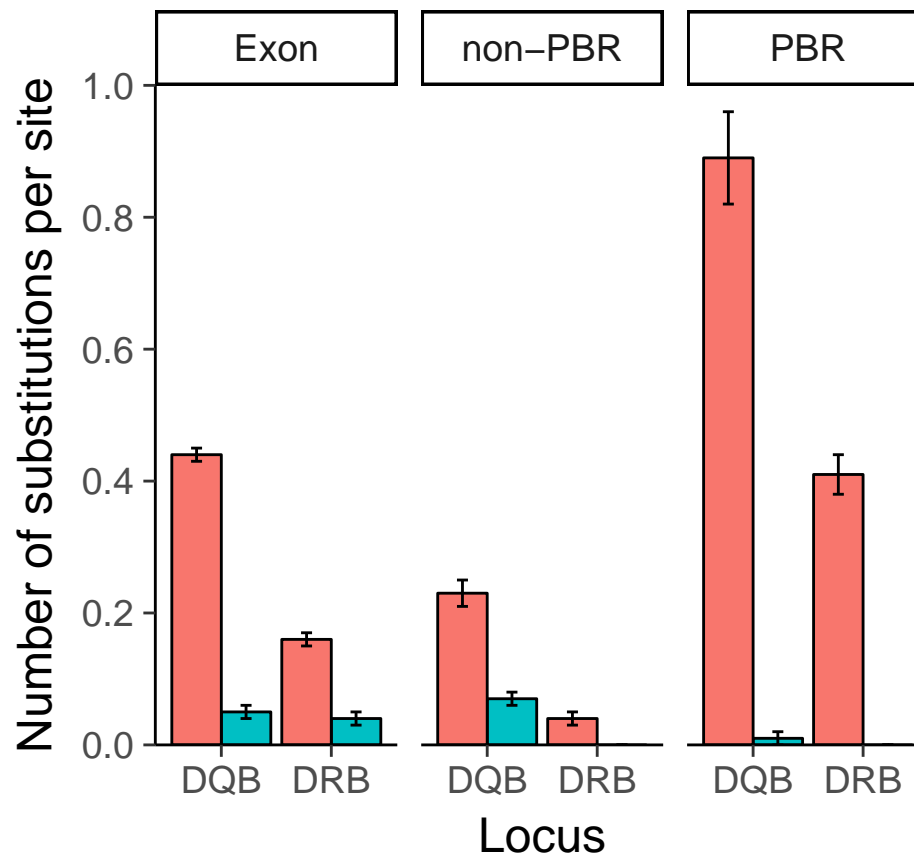
```

```

###
df <- data.frame(
  locus = rep(c("DQB", "DRB"), each = 6),
  sub_rate = c(0.44, 0.89, 0.23, 0.05, 0.01, 0.07,
               0.16, 0.41, 0.04, 0.04, 0, 0),
  se = c(0.01, 0.07, 0.02, 0.01, 0.01, 0.01,
          0.01, 0.03, 0.01, 0.01, 0, 0),
  unit = rep(c("Exon", "PBR", "non-PBR"), 4),
  sub_type = rep(rep(c("dN", "dS"), each = 3), 2)
)

library(ggplot2)
p <- ggplot(df, aes(x = locus, y = sub_rate, fill = sub_type)) +
  geom_bar(stat = "identity", color = "black",
           position = position_dodge()) +
  geom_errorbar(aes(ymin = sub_rate - se, ymax = sub_rate + se), width = .2,
                position = position_dodge(.9)) +
  theme_classic(base_size = 18) +
  theme(legend.title = element_blank(),
        legend.position = "n") +
  scale_y_continuous(expand = c(0,0),
                     breaks = seq(0,1,0.2),
                     limits = c(0,1)) +
  labs(title = "", x = "Locus", y = "Number of substitutions per site") +
  facet_wrap(~unit)
p

```



```

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      datasets  methods
> [7] base
>
> other attached packages:
> [1] ggplot2_2.2.1 magrittr_1.5  knitr_1.17
>
> loaded via a namespace (and not attached):

```

```
> [1] Rcpp_0.12.14      digest_0.6.13      rprojroot_1.3-1
> [4] plyr_1.8.4        grid_3.4.3         gtable_0.2.0
> [7] backports_1.1.2   evaluate_0.10.1    scales_0.5.0
> [10] rlang_0.1.4       stringi_1.1.6      lazyeval_0.2.1
> [13] rmarkdown_1.8     tools_3.4.3        stringr_1.2.0
> [16] munsell_0.4.3     yaml_2.1.16        compiler_3.4.3
> [19] colorspace_1.3-2  htmltools_0.3.6    tibble_1.3.4
```
