

Data exploration of immune data in laboratory and field infections

Fay

2024-02-01

Data exploration of field and laboratory infection data

Laboratory data compilation of experimental infections

Experimental planning

Mice

- four wild-derived inbred mouse strains along with their respective F1 hybrids.
- Two of these strains, SCHUNT and STRA, represent *M. m. domesticus*.
- The strains BUSNA and PWD were derived from *M. m. musculus*
- Two intersubspecific hybrids (STRAxBUSNA and SCHUNTxPWD) and
- Two intrasubspecific hybrids (SCHUNTxSTRA and PWDxBUSNA).

```
mouse_strains <- c("SCHUNT", "STRA", "BUSNA", "PWD", "STRA BUSNA",  
                  "SCHUNT PWD", "SCHUNT STRA", "PWD BUSNA")  
hybrid_status <- c("M. m. domesticus", "M. m. domesticus", "M. m. musculus",  
                  "M. m. musculus", "intersubspecific hybrids",  
                  "intersubspecific hybrids", "intrasubspecific hybrids",  
                  "intrasubspecific hybrids")  
strains <- as.data.frame(hybrid_status, mouse_strains)  
strains
```

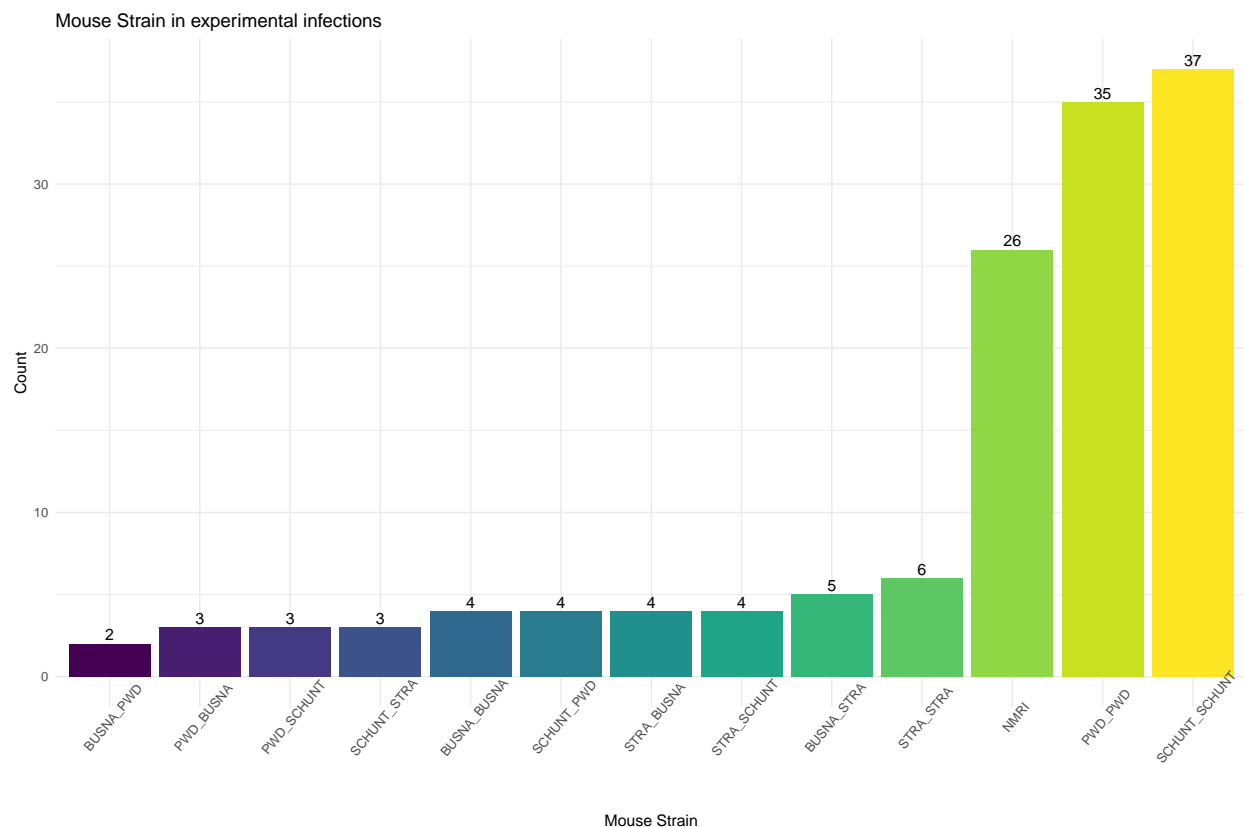
```
##                hybrid_status  
## SCHUNT          M. m. domesticus  
## STRA            M. m. domesticus  
## BUSNA           M. m. musculus  
## PWD             M. m. musculus  
## STRA BUSNA      intersubspecific hybrids  
## SCHUNT PWD      intersubspecific hybrids  
## SCHUNT STRA     intrasubspecific hybrids  
## PWD BUSNA       intrasubspecific hybrids
```

```
lab %>%  
  group_by(mouse_strain) %>%  
  # Summarize the data to get counts for each mouse strain  
  summarize(count = n()) %>%  
  # Reorder mouse_strain by count  
  mutate(mouse_strain = reorder(mouse_strain, count)) %>%  
  # Plotting
```

```

ggplot(aes(x = mouse_strain, y = count, fill = mouse_strain)) +
  geom_bar(stat = "identity") +
  # Specify stat = "identity" for pre-summarized data
  geom_text(aes(label = count), vjust = -0.3) +
  # Add count labels on top of bars
  scale_fill_viridis_d() +
  # Use a nice color scale, like Viridis
  theme_minimal() + # Apply a minimal theme for a cleaner look
  theme(axis.text.x = element_text(angle = 50)) +
  labs(title = "Mouse Strain in experimental infections", x = "Mouse Strain",
        y = "Count") + # Add label
  guides(fill = FALSE)

```



```

# Define colors
colors <- c("TRUE" = "firebrick3", "FALSE" = "steelblue")

# transform mouse strain into factor
lab$mouse_strain <- as.factor(lab$mouse_strain)

lab$mouse_strain <- gsub(pattern = "_", " ", lab$mouse_strain)

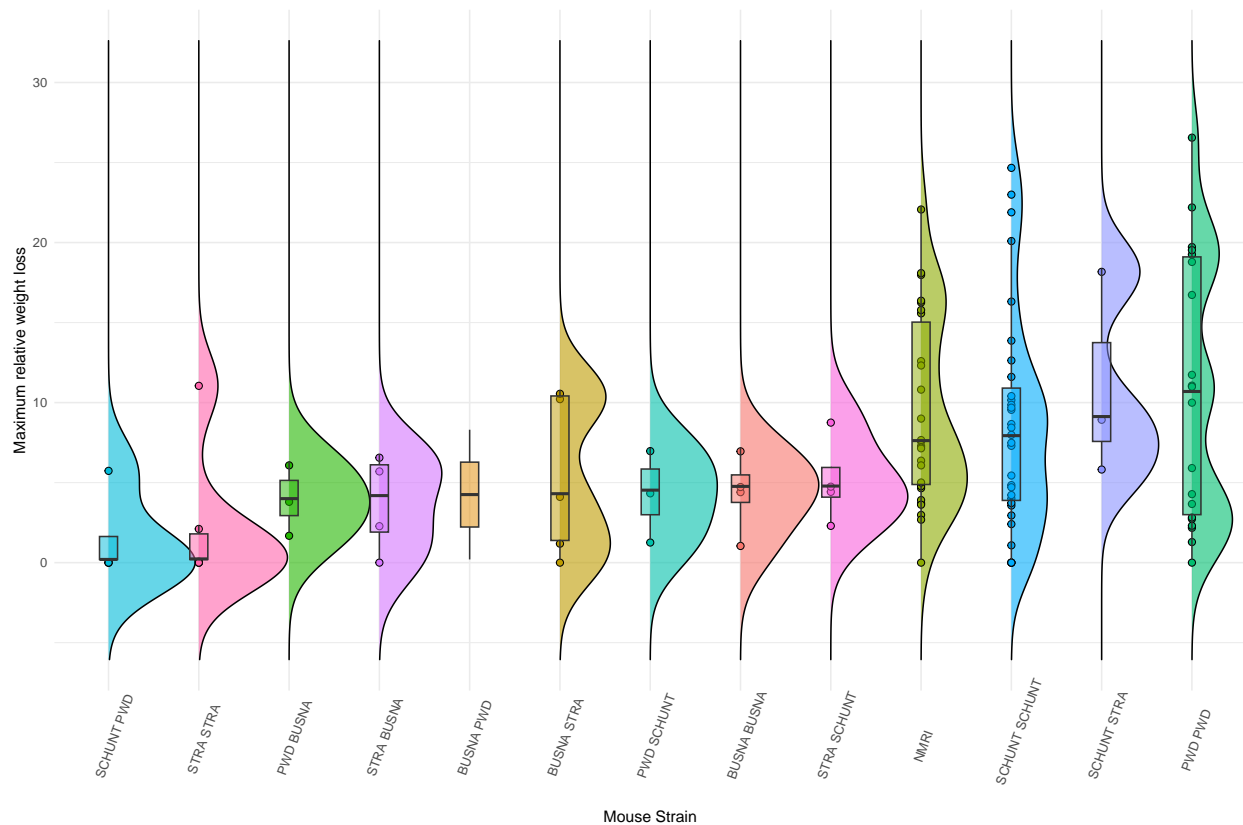
# order factor levels
lab$mouse_strain <- factor(lab$mouse_strain,
                           levels = names(
                             sort(tapply(lab$WL_max, lab$mouse_strain, median))))

```

```
ggplot(lab %>%
  filter(infection == "challenge"),
  aes(x = WL_max, y = mouse_strain, fill = mouse_strain)) +
  geom_density_ridges(jittered_points = TRUE, position = position_points_jitter(height = 0),
    scale = 0.9, alpha = 0.6, point_shape = 21, point_size = 2, point_alpha = 1) +
  geom_boxplot(width = 0.2, outlier.shape = NA, alpha = 0.5, position = position_nudge(x = 0.2)) +
  coord_flip() +
  theme_minimal() +
  theme(legend.position = "none",
    axis.text.x = element_text(angle = 70, vjust = 0.5, hjust=0.3)) +
  xlab("Maximum relative weight loss") +
  ylab("Mouse Strain") -> strains_weight
```

strains_weight

Picking joint bandwidth of 2.02



```
ggsave(filename = "figures/strains_weight.jpeg", plot = strains_weight,
  width = 8, height = 6, dpi = 1000)
```

Picking joint bandwidth of 2.02

```
ggplot(lab, aes(x = WL_max, y = mouse_strain, fill = mouse_strain)) +
  geom_density_ridges(jittered_points = TRUE, position = position_points_jitter(height = 0),
```



```

color_mapping <- c("E_falciformis" = "salmon",
                  "E_ferrisi" = "forestgreen",
                  "uninfected" = "blue")

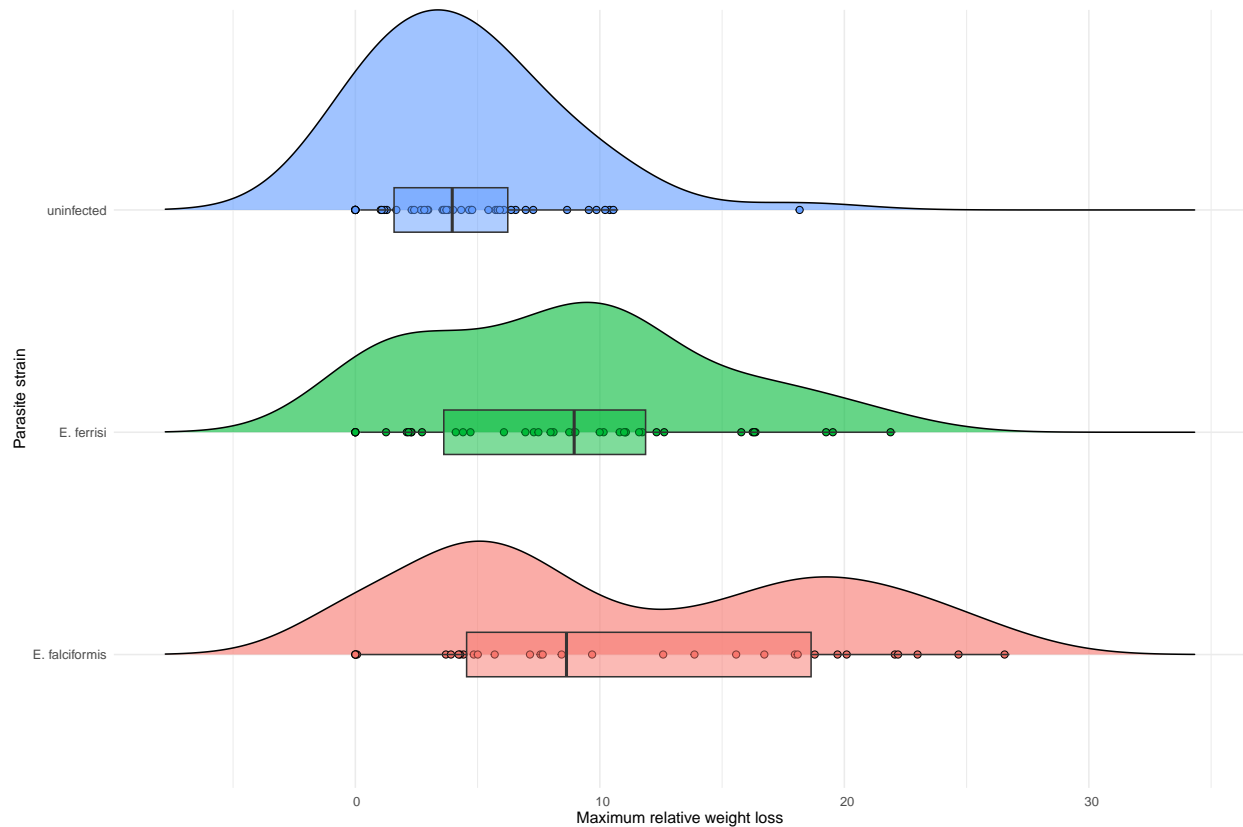
lab$current_infection <- gsub(pattern = "_", replacement = ". ", lab$current_infection)

ggplot(lab %>%
  filter(infection == "challenge"),
  aes(x = WL_max, y = current_infection, fill = current_infection)) +
  geom_density_ridges(jittered_points = TRUE, position = position_points_jitter(height = 0),
                    scale = 0.9, alpha = 0.6, point_shape = 21, point_size = 2, point_alpha = 1) +
  geom_boxplot(width = 0.2, outlier.shape = NA, alpha = 0.5, position = position_nudge(x = 0.2)) +
  # coord_flip() +
  theme_minimal() +
  scale_color_manual(values = color_mapping) +
  theme(legend.position = "none",
        axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.3)) +
  xlab("Maximum relative weight loss") +
  ylab("Parasite strain") -> eimeria_weight

eimeria_weight

```

Picking joint bandwidth of 2.59



```
ggsave(filename = "figures/eimeria_strains_weight.jpeg", plot = eimeria_weight,
        width = 8, height = 6, dpi = 1000)
```

```
## Picking joint bandwidth of 2.59
```

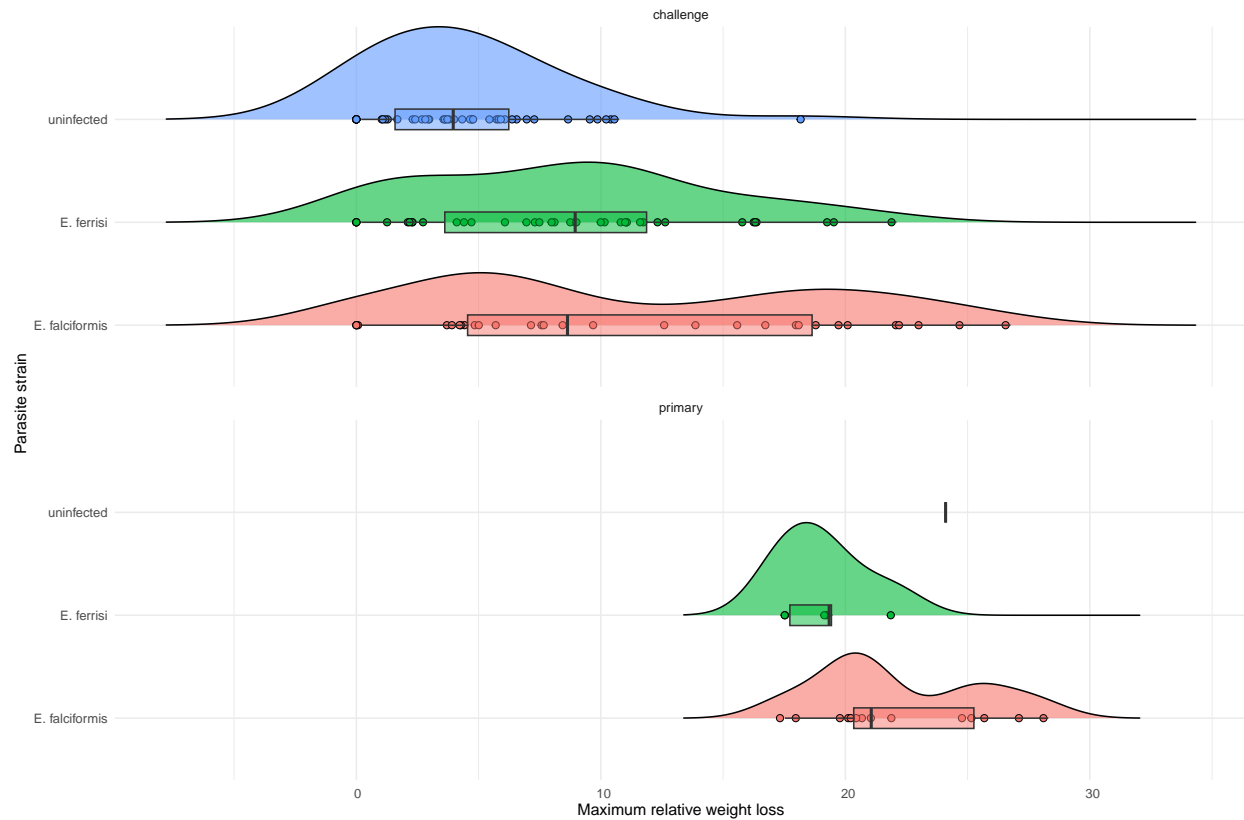
```
# prim vs challenge
lab$current_infection <- gsub(pattern = "_", replacement = ". ", lab$current_infection)

ggplot(lab,
        aes(x = WL_max, y = current_infection, fill = current_infection)) +
  geom_density_ridges(jittered_points = TRUE, position = position_points_jitter(height = 0),
                      scale = 0.9, alpha = 0.6, point_shape = 21, point_size = 2, point_alpha = 1) +
  geom_boxplot(width = 0.2, outlier.shape = NA, alpha = 0.5, position = position_nudge(x = 0.2)) +
  # coord_flip() +
  theme_minimal() +
  theme(legend.position = "none",
        axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.3)) +
  xlab("Maximum relative weight loss") +
  ylab("Parasite strain") +
  scale_color_manual(values = color_mapping) +
  facet_wrap(~ infection, nrow = 2)-> eimeria_weight_challenge

eimeria_weight_challenge
```

```
## Picking joint bandwidth of 2.59
```

```
## Picking joint bandwidth of 1.31
```



```
ggsave(filename = "figures/eimeria_strains_weight_challenge.jpeg", plot = eimeria_weight_challenge,
        width = 8, height = 10, dpi = 1000)
```

```
## Picking joint bandwidth of 2.59
## Picking joint bandwidth of 1.31
```

```
##### is the difference between parasite strains significant?
parasite_WL <- lm(formula = WL_max ~ current_infection, data = lab)
summary(parasite_WL)
```

```
##
## Call:
## lm(formula = WL_max ~ current_infection, data = lab)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.6407  -4.7977  -0.1096   5.2067  19.1025
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      14.641      1.018  14.384 < 2e-16 ***
## current_infectionE. ferrisi  -4.876      1.448  -3.368  0.00099 ***
## current_infectionuninfected  -9.843      1.424  -6.912  1.8e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 6.828 on 133 degrees of freedom
## Multiple R-squared:  0.2643, Adjusted R-squared:  0.2533
## F-statistic: 23.9 on 2 and 133 DF,  p-value: 1.361e-09
```

```
##### is the difference between mouse strains significant?
```

```
mouse_WL <- lm(WL_max ~ mouse_strain, data = lab)
```

```
summary(mouse_WL)
```

```
##
## Call:
## lm(formula = WL_max ~ mouse_strain, data = lab)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15.4049  -4.4243  -0.4518   4.3301  14.7010
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4337     3.4788   0.412 0.680967
## mouse_strainSTRA STRA      0.7698     4.4911   0.171 0.864185
## mouse_strainPWD  BUSNA     2.4180     5.3140   0.455 0.649893
## mouse_strainSTRA  BUSNA     2.1998     4.9198   0.447 0.655563
## mouse_strainBUSNA PWD      2.6190     6.0255   0.435 0.664575
## mouse_strainBUSNA STRA     3.7786     4.6673   0.810 0.419739
## mouse_strainPWD  SCHUNT     2.7512     5.3140   0.518 0.605575
## mouse_strainBUSNA BUSNA     2.8474     4.9198   0.579 0.563809
## mouse_strainSTRA  SCHUNT     3.6208     4.9198   0.736 0.463151
## mouse_strainNMRI              7.8072     3.7368   2.089 0.038746 *
## mouse_strainSCHUNT SCHUNT     8.6346     3.6620   2.358 0.019957 *
## mouse_strainSCHUNT STRA     9.5383     5.3140   1.795 0.075117 .
## mouse_strainPWD  PWD      13.9712     3.6722   3.805 0.000223 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.958 on 123 degrees of freedom
## Multiple R-squared:  0.2935, Adjusted R-squared:  0.2246
## F-statistic: 4.259 on 12 and 123 DF,  p-value: 1.298e-05
```