Data exploration of the immune response of wild mice against parasite infections

Fay

2024-02-01

Data structure of field and laboratory data sets

Data originating from yearly field excursions in the House Mouse Hybrid Zone

GitHub repository: https://github.com/derele/Mouse_Eimeria_Field/tree/master

Infection intensity data

Eimeria spp. oocysts counting:

counter: character, Person who counted Feces_g: numeric, amount of feces used in flotation Date_count: date counted N_oocysts_sq1 ... sq8: numeric, individual count for each single square on the neubaure chamber (up to 8) mean_neubauer: numeric, mean of the 8 squares PBS_dil_in_mL: numeric, in which volume of PBS whas Ncells: number of neubauer "cells" (squares) counted OPG: oocysts per gram feces (calculated from the above) In some cases only OPG data might be available or rather data would need to be re-formatted to access all the raw values.

Eimeria spp. detection qPCR We perform relative qPCR for detection and quantification of Eimeria DNA in intestinal tissue. We amplify a locus in the nuclear genome of the house mouse and a locus in the mitochondrial genome (COI) of Eimeria. We then calculate a "delta" between the two ct values.

delta_ct_ilwe_MminusE: threshold cycle for mouse minus Eimeria in Ileum tissue. Only E. vermiformis is (at low pervalence in Ileum tissue) and we therefore don't obtain this data-type for all years.

delta_ct_cewe_MminusE: threshold cycle for mouse minus Eimeria in Caecume tissue. E. ferrisi and E. falcifromis are detected here. We should have this (as coprehensively as possible) for every year!

MC.Eimeria: TRUE/FALSE. This was established in 2018 as an improvement over the '> -5 delta ct rule' for identification of Eimeria -positive qPCRs. Melting curves have to show a drastic drop at XX°C to indicate melting of a proper Eimeria COI amplification product. It might be added where possible for per 2018 data post-hoc (if melting curves exist for a review of raw data).

Data structure - laboratory challenge infections with *Eimeria spp.*

Contains data for challenge (repeated) infections performed between 2017 and 2019. The data product is structure in the following columns:

Mouse_ID: the unique identifier of the mouse experiment: the experiment as numbered in the overview table mouse_strain: the strain (inbred or outbred) of the mouse primary_infection: The Eimeria strain used for the primary infection challenge_infection: The Eimeria strain used for the challenge infection infection_history:

The resulting infection history labels: the unique label of the fecal sample at a particular dpi weight: the weight of the mouse at this dpi weight_dpi0: the weight at the day of infection relative_weight: the weight of the mouse at this dpi relative to the weight at dpi0 feces_weight: the weight of the feces collected at this dpi dpi: days post infection at which samples and data in this row were taken infection: the infection (primary or challenge) this row/dpi corresponds to oocyst_sq1, oocyst_sq2, oocyst_sq3, oocyst_sq4: the raw values for squared during oocyst counting dilution: the amount of PBS the feces (with it's relative weight) was dissolved in OO4sq: the sum of oocysts in the four counting squares OOC: the overall number of oocysts in in the feces (of a particular weight) at this dpi OPG_O: Old way of counting opg (Emanuel ask me) infection_type: what kind of infection are we looking at (challenge or primary, homologous or heterologous immunization). This is differently coded to infection, as here UNI:E88 (first uninfected, then infected with E88) would count as a "primaryE88" infection The next values max_dpi until maximum_weight are calculated for the infection type in which the mice died

max_dpi = maximum dpi that the mouse that the mouse reached for each infection challenge or primary (group_by EH_ID and infection (primary/challenge) to get the value

maximum oocysts for each infection type (group_by EH_ID and infection (primary/challenge) to get the value)

maximum weight loss for each infection type (group_by EH_ID and infection (primary/challenge) to get the value)

death = challenge/primary (in which infection did the mouse die

Eim_MC = Melting curve for eimeria

delta = delta ct value

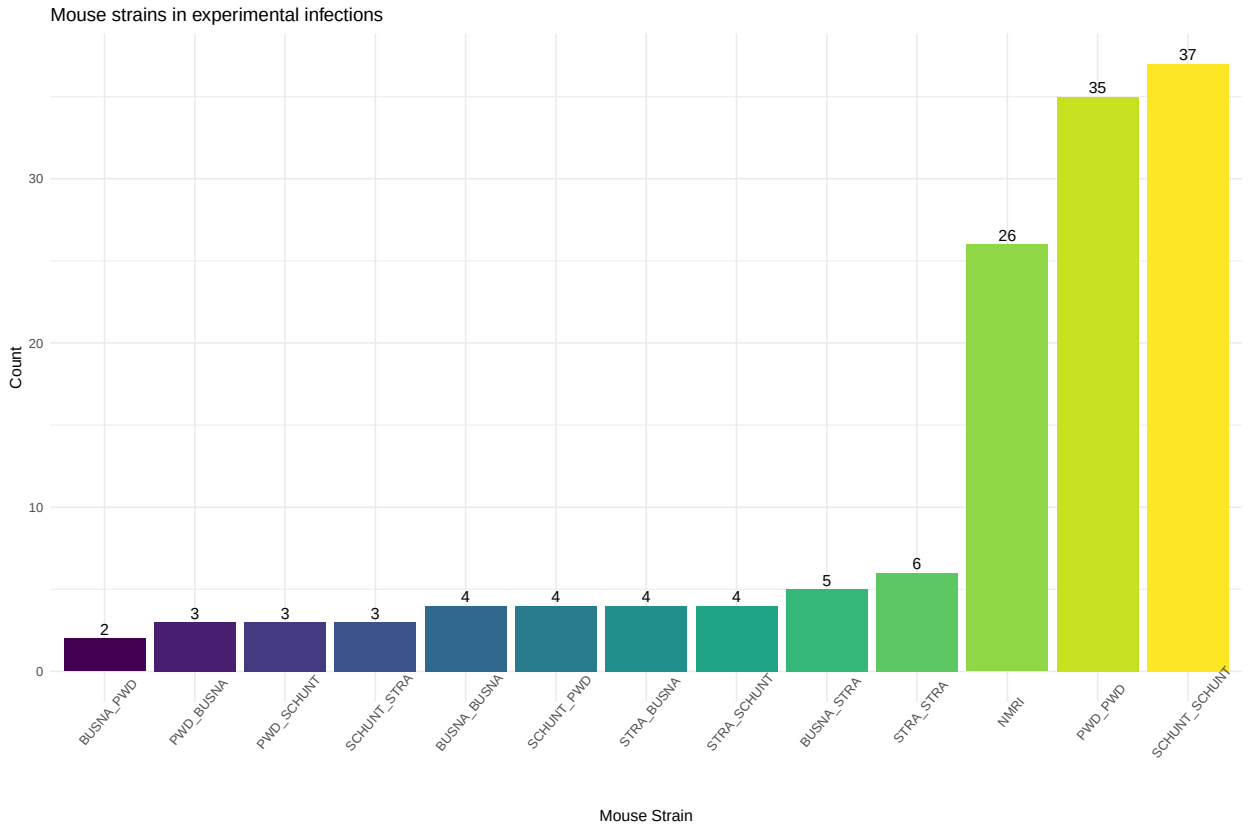
Experimental planning - Laboratory infections

Selected mouse strains

- 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).

	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

Numbers of each mouse strain



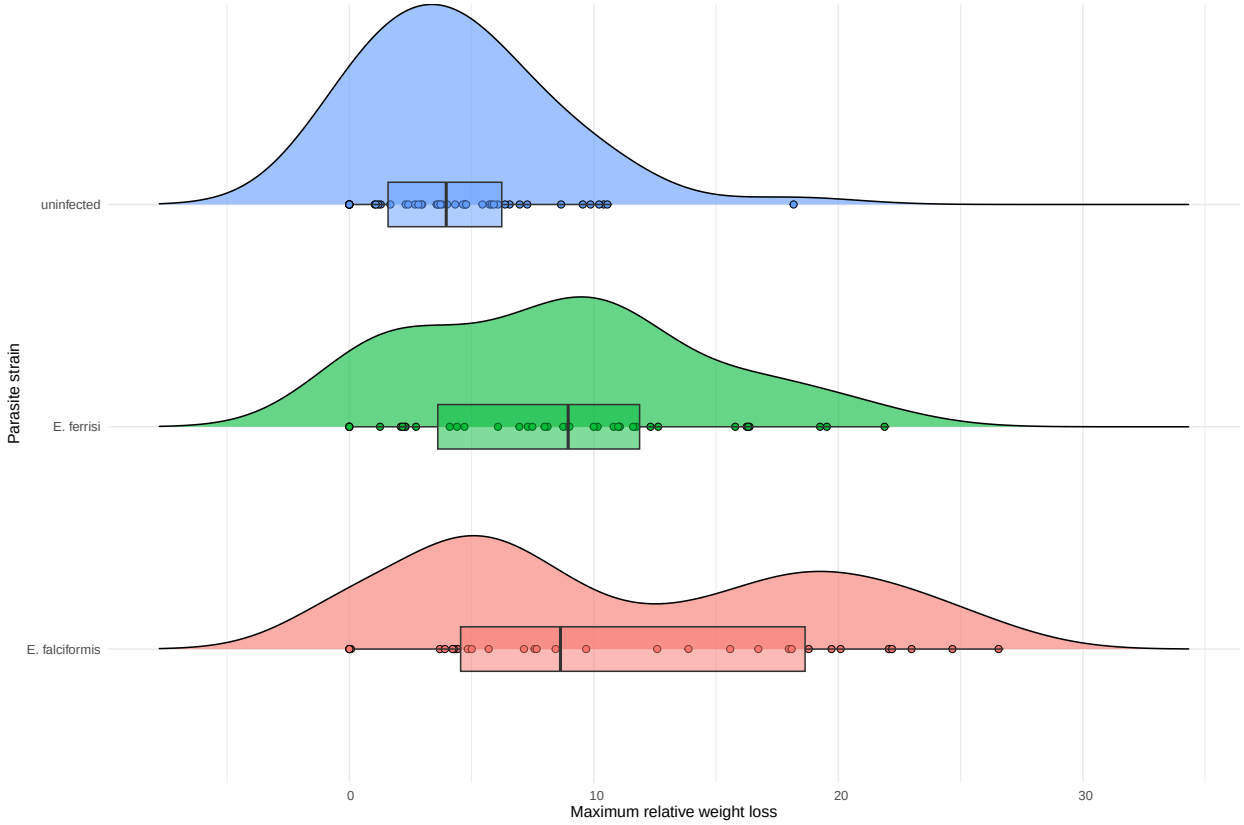
Selected parasite strains

Infections were initiated by oral administration of 150 sporulated Eimeria oocysts * Up to 16 species of Eimeria have been described from house mice * Overall prevalence in the wild 25.9% * Prevalence of *E. ferrisi* 14 % * Prevalence of *E. falciformis* 4%

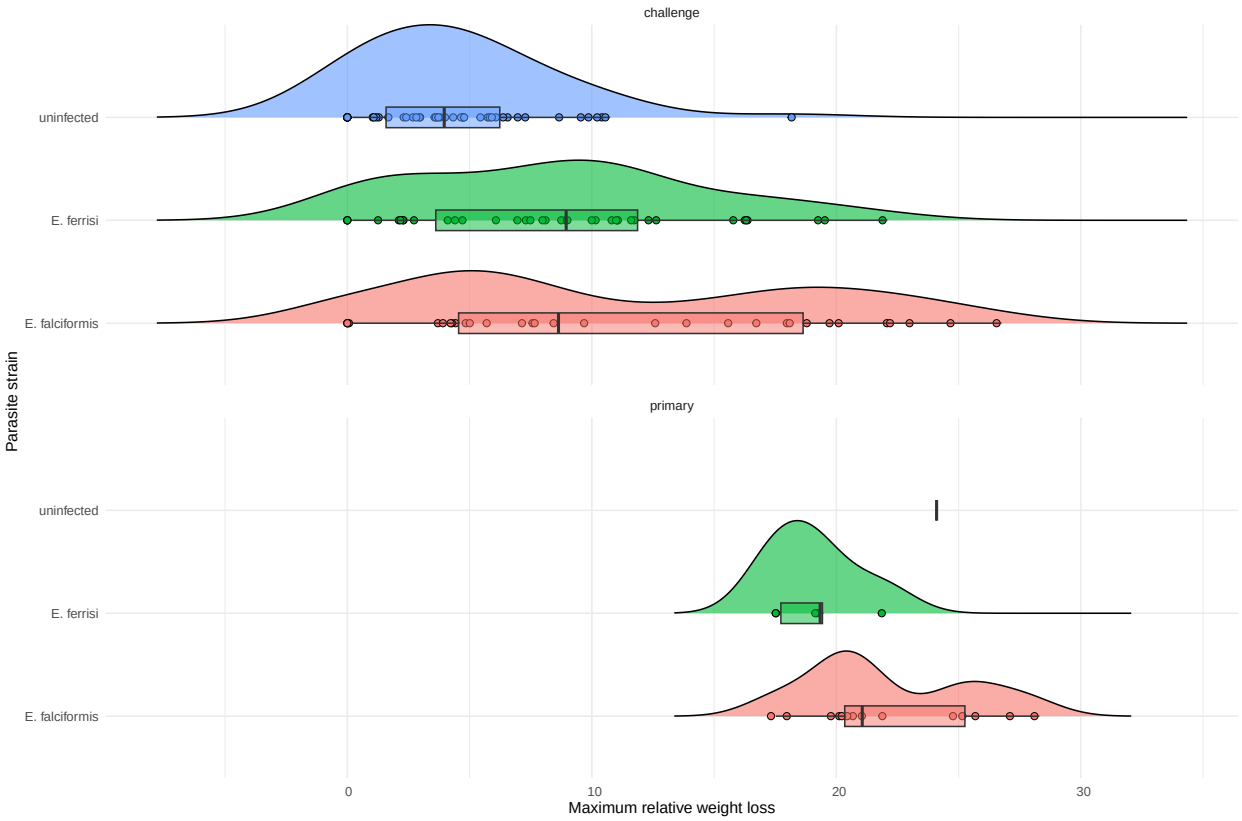
As a proxy for health we use the maximum relative weight lost during infection

Maximum weight loss = highest relative weight on any day of the experiment / starting weight

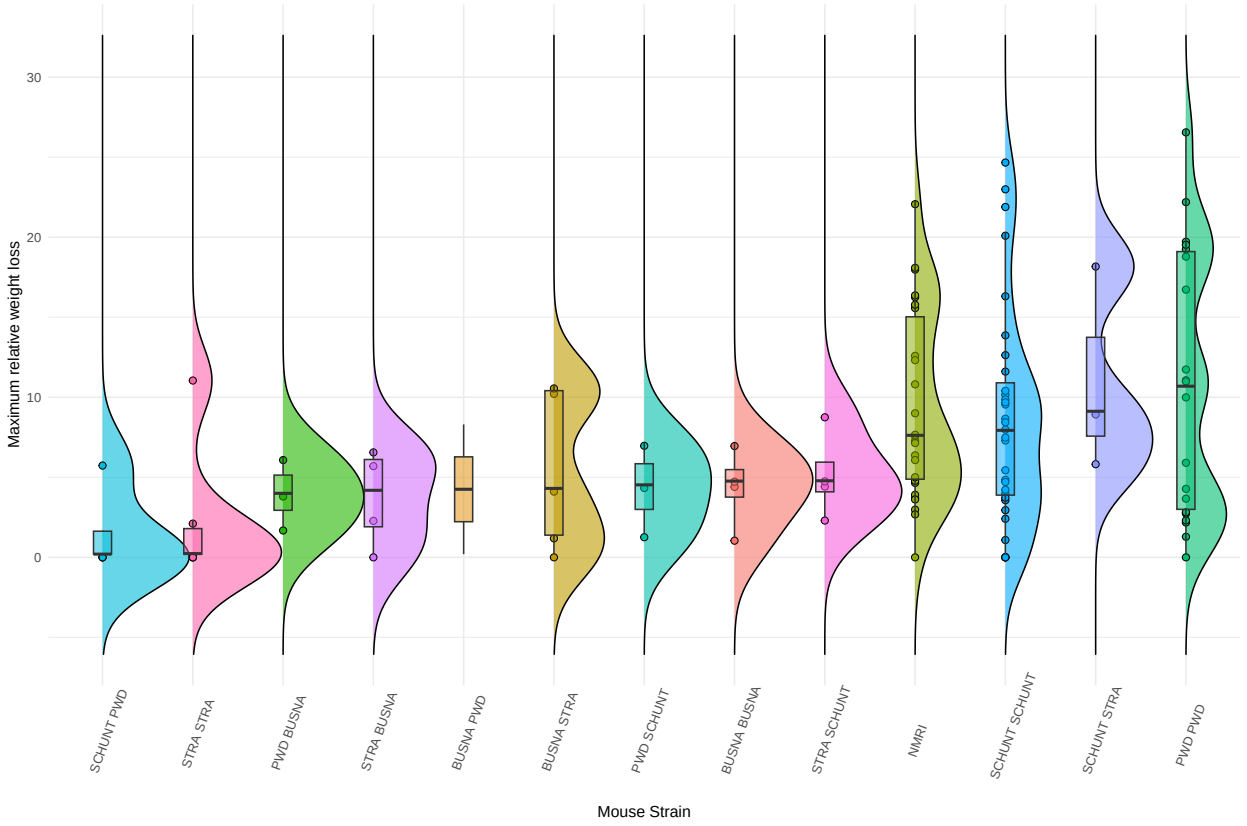
Maximum relative weight loss in each infection group, in challenge and primary infections combined.



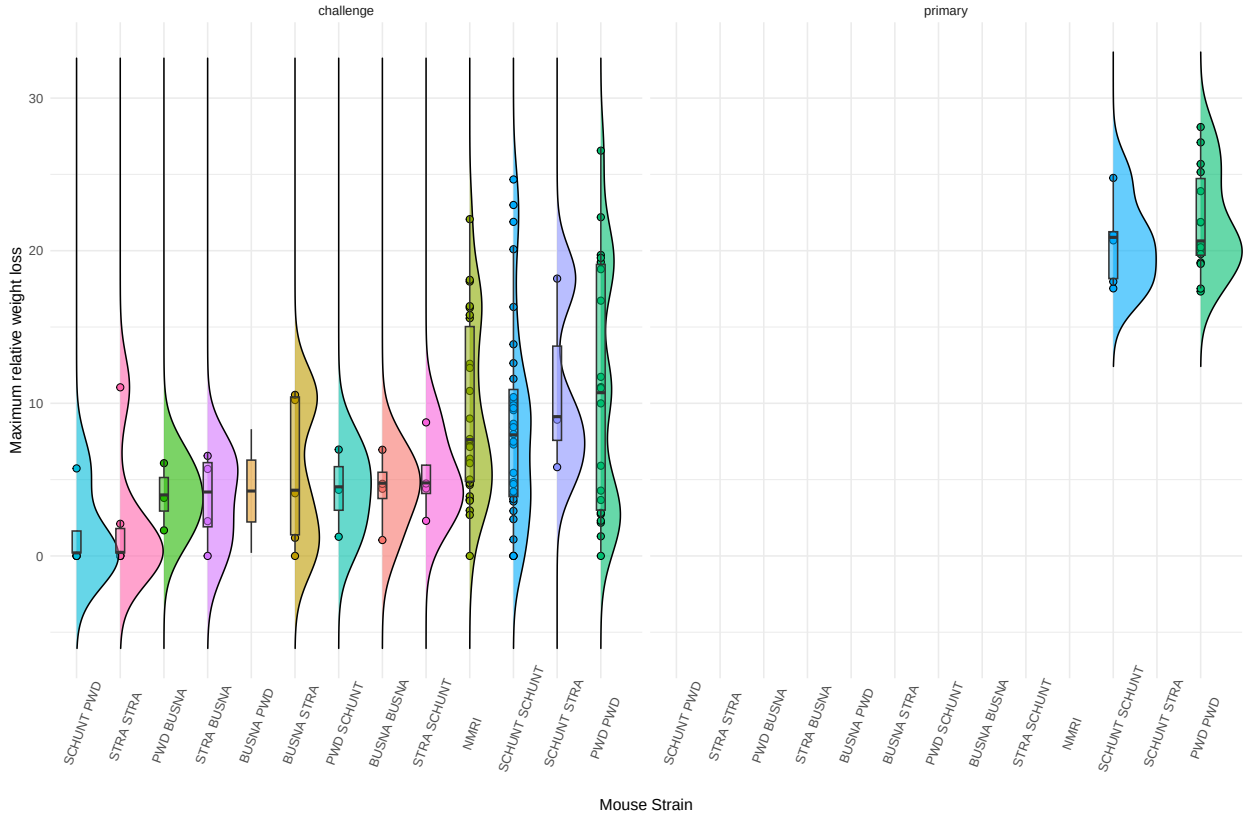
Most of the mice in this experiment, are mice that have been challenged (infected for a second time). This replicates more accurately what happens in the wild. A much higher weight loss is expected in the primary infections.


challenge
uninfected
E. ferrisi
E. falciformis
primary
uninfected
E. ferrisi
E. falciformis
Parasite strain
0
10
20
30
Maximum relative weight loss

Weight loss per mouse strain



Weight loss per mouse strain, challenge infections vs primary infections



Preliminary data analysis immune data from laboratory infection experiments

For how many mice do we have immune data?

```
length(lab$Mouse_ID)
```

```
## [1] 136
```

How many mice in primary and how many in challenge infections?

```
lab %>%
    group_by(infection) %>%
    summarize(n())
```

```
## # A tibble: 2 x 2
##   infection `n()`
##   <chr>     <int>
## 1 challenge   116
## 2 primary      20
```

How many mice are there in each infection group?

```
lab %>%
    group_by(infection, current_infection) %>%
    summarize(n())
```

```
## # A tibble: 6 x 3
## # Groups:   infection [2]
##   infection current_infection `n()`
##   <chr>     <chr>             <int>
## 1 challenge E. falciformis       31
## 2 challenge E. ferrisi           39
## 3 challenge uninfected           46
## 4 primary   E. falciformis       14
## 5 primary   E. ferrisi            5
## 6 primary   uninfected            1
```

For how many mice do we have FACS data?

```
## [1] 80
```

For how many mice do we have immune gene expression data?

```
## [1] 136
```

How many mice have immune gene expression AND FACS data?

```
length(intersect(FACS_M$Mouse_ID, GENE_M$Mouse_ID))
```

```
## [1] 80
```

For the complete FACS Data set, immune gene expression is complete too.

Field infections - FACS immune cell data

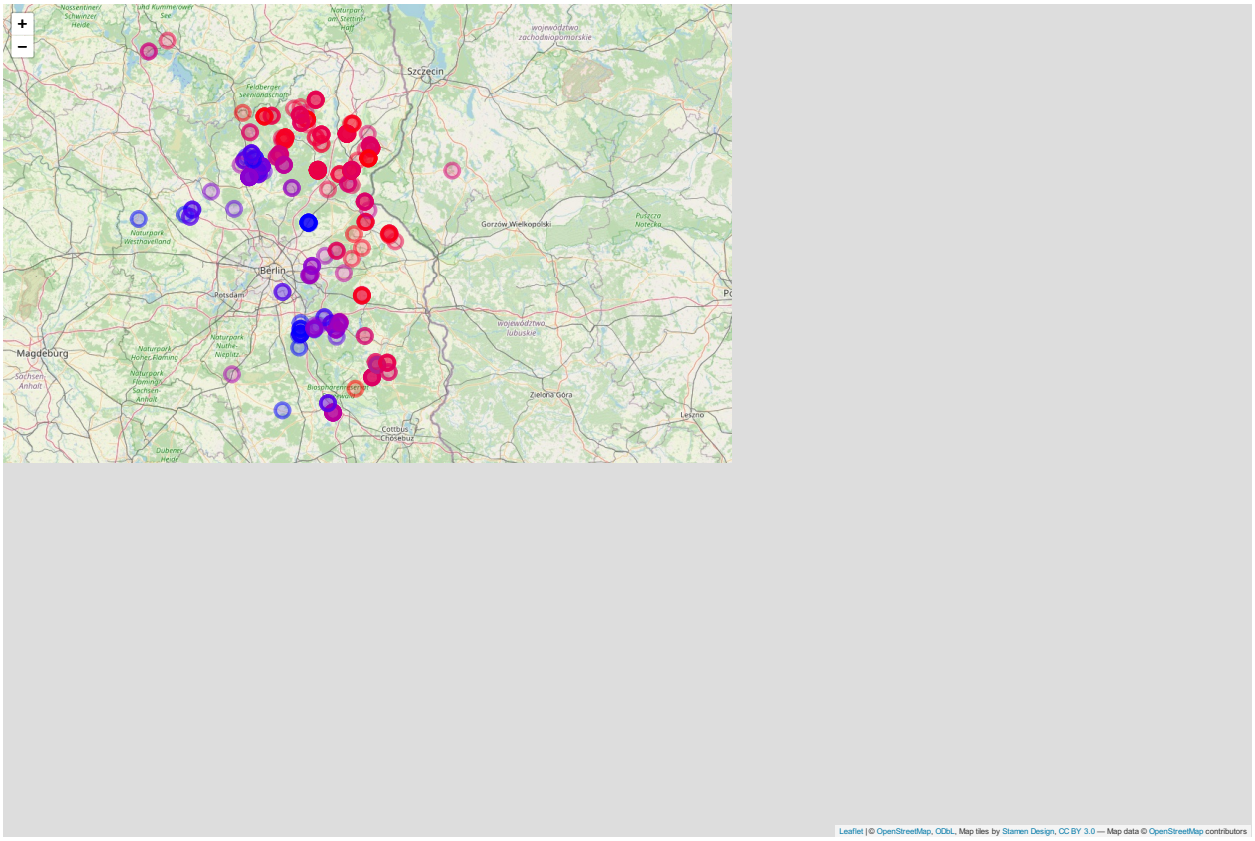
Number of mice with FACS data

```
## [1] 94
```

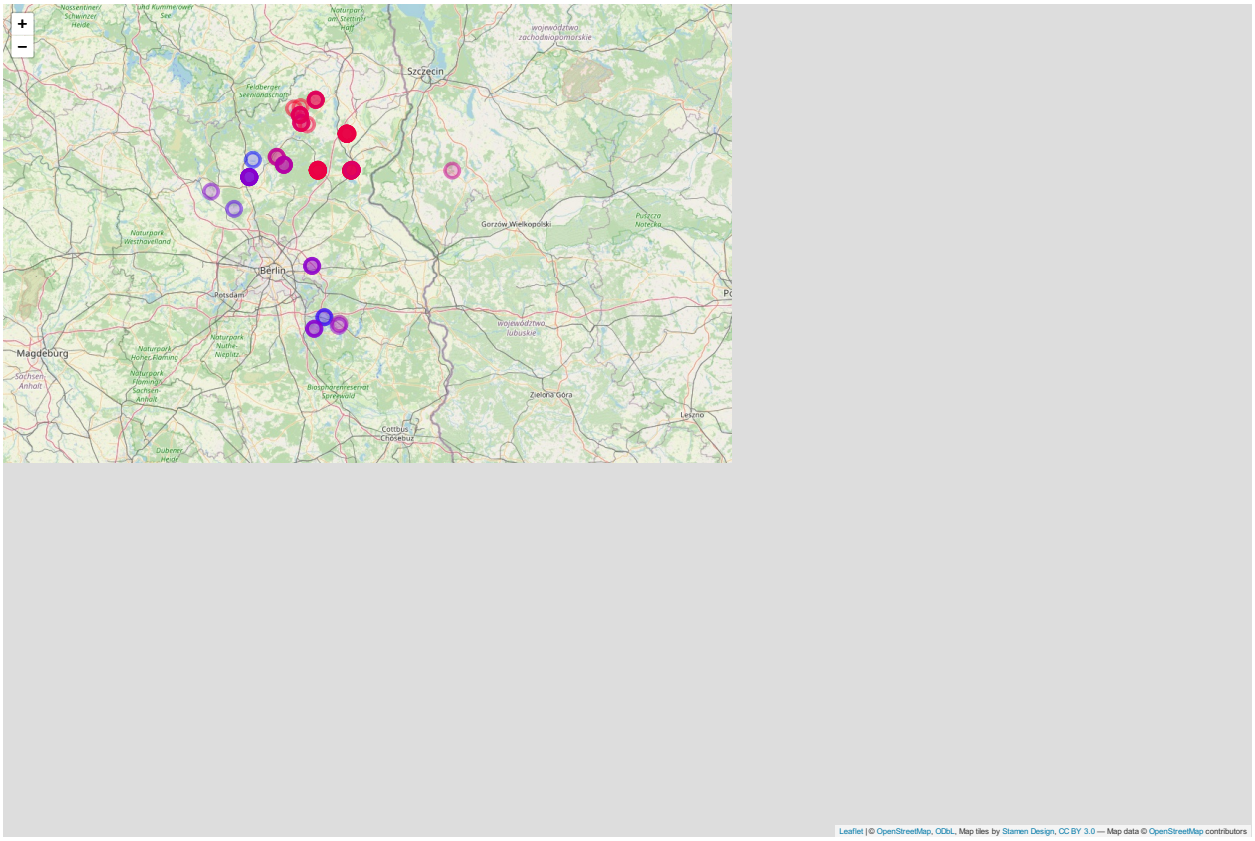
Number of mice with gene expression data

```
## [1] 336
```

Capture locations of mice with gene expression data

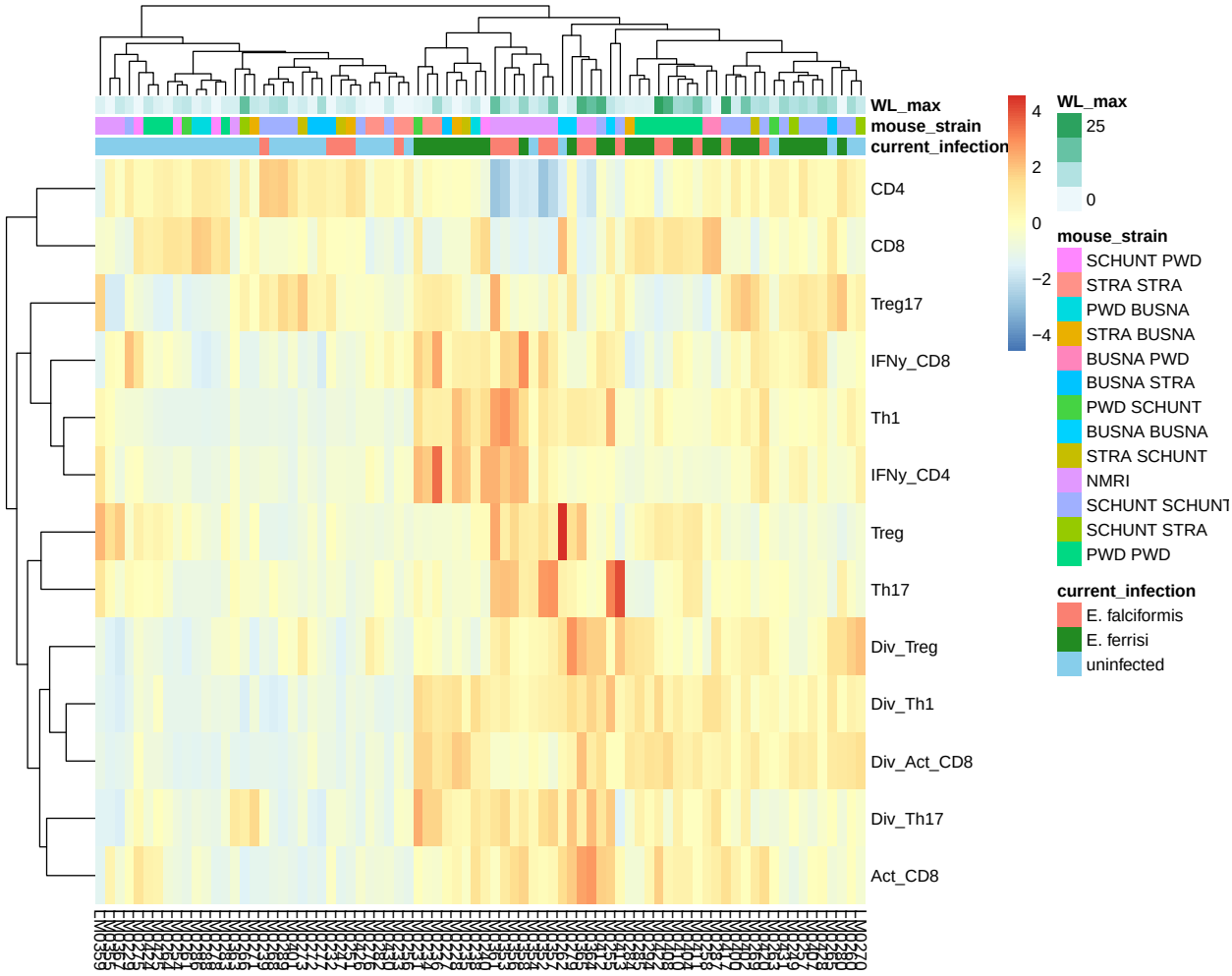


Capture locations of mice with immune cell data



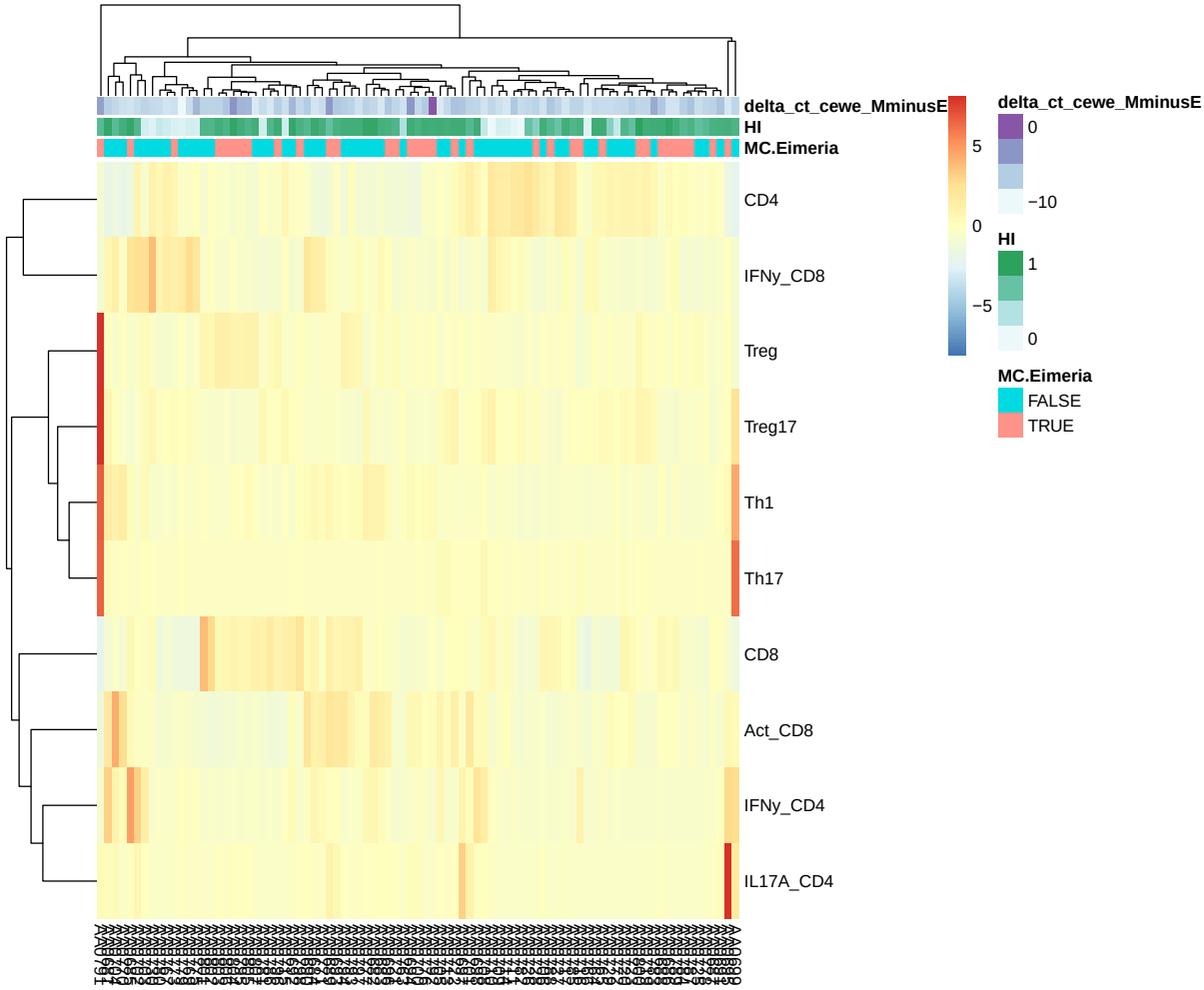
Immune cells

Laboratory infections - heatmap of immune cells



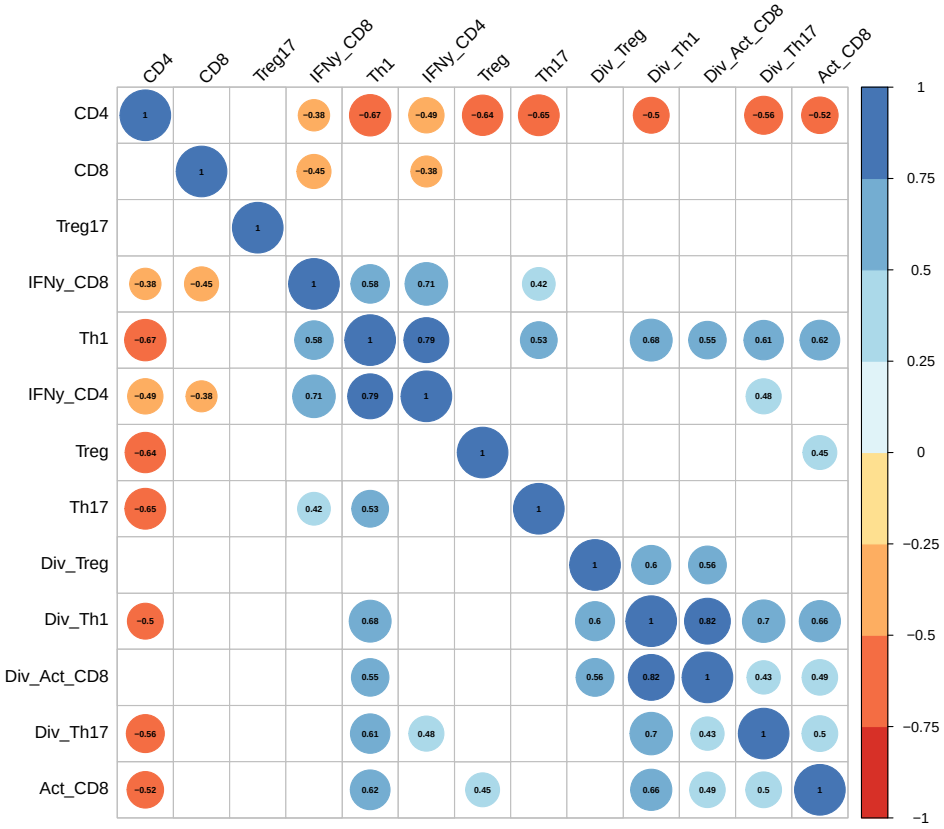
Visual difference between infected and uninfected mice

Field infections immune cells - heatmap

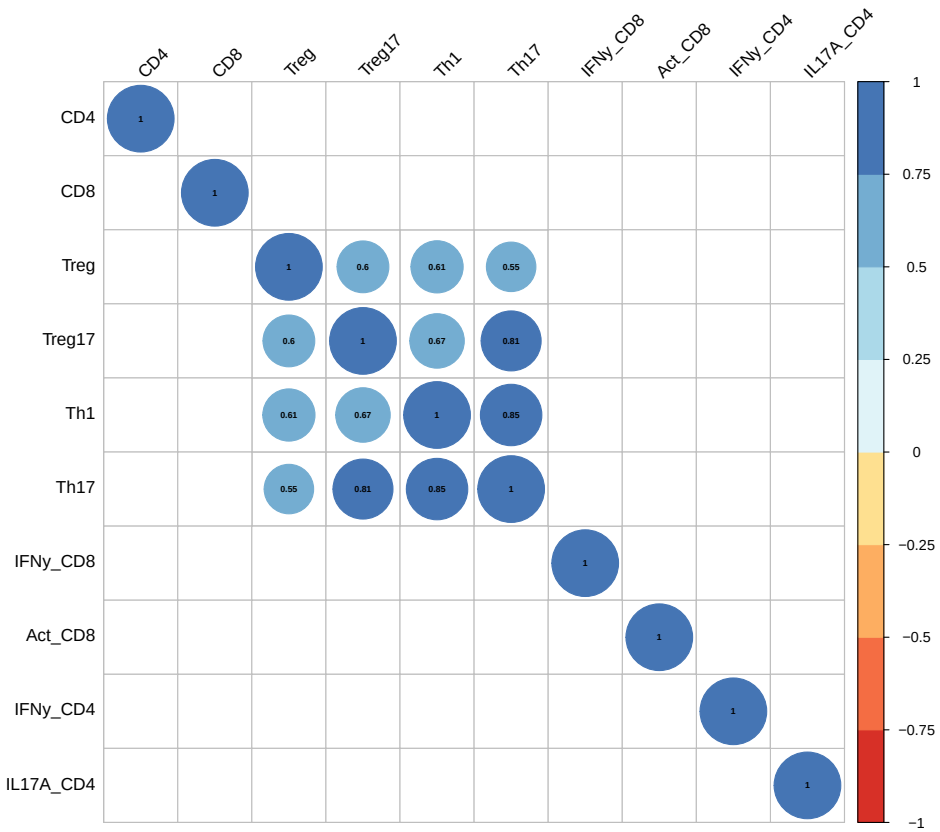


Nothing to gain from this

Correlation between cells in laboratory challenge infections



Correlations between cells in field mice



Immune cell percentages in response to infection group, laboratory infections

- Missing: completing the data set of genotyping parasites in field samples