

PAPER

Tolerance of hybrid hosts against infections, TAC Meeting December 2022

Fay Webster,^{1,2,*} Lubomír Bednář^{1,2,†} and Emanuel Heitlinger^{1,2,‡}

¹Department of Molecular Parasitology, Institute for Biology, Humboldt University Berlin (HU), Philippstraße 13, Berlin, 10115, Germany and ²Research Group Ecology and Evolution of Molecular Parasite-Host Interactions, Leibniz-Institut für Zoo und Wildlife Research (IZW), Alfred-Kowalke-Straße 17, Berlin, 10315, Germany

*Corresponding author. fay.webster@hu-berlin.de [†]Current email address: fay.webster@hu-berlin.de [‡]

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Parasites in hybrid zones can give insight into species barriers, as they are modulating the fitness of hybrid hosts. Recent findings have demonstrated lower infection intensities with parasites in hybrids in the European House Mouse Hybrid zone (HMHZ), indicating higher disease resistance. However, tolerance has not yet been addressed in depth, as it is impractical to measure in wild populations. In an attempt to predict and evaluate the health impact of parasite infections and extrapolate tolerance in the HMHZ, we use a machine learning method. A random forest model was trained on immune parameters measured in experimental lab infections with *Eimeria* and then applied to data obtained from field sampling. Our predictions revealed that these infections are more detrimental to hybrid male mice. This approach represents an initial step in assessing tolerance in field studies.

Key words: hybrids; parasites

1. Introduction

1.1. Introduce the topic

2. Methods

2.1. Statistical Analysis

2.1.1. Imputation of missing data

To make the most of our data collection, we aimed to resolve missingness. Missing data were imputed using multiple imputations by chained equations. We used the package MICE in R Van Buuren and Groothuis-Oudshoorn (2011), with five imputed data sets and five iterations. Data generated by FACS or the Gene Expression / Biomarker assay were regarded as missing if each mouse had measurements for some variables. For each continuous variable, we specified a predictive mean matching model. All the remaining variables were used as predictors in the imputation. To control the quality of our imputations, we evaluated the distribution plot of the existing data and the imputed data for all measurements. Further, we tested for convergence ???. We assume data is “missing completely at random” or “missing at random”. For both types of missingness, multiple imputation is a suggested method to impute missing variables Van Buuren (2018).

Questions / To-dos:

1. Should I log-transform the data prior to imputation?
2. Increasing produced data sets / iterations
3. sensitivity analyses using complete cases only

3. Results

4. Discussion

5. Conclusion

5.1. A subsection

6. Literature citations

7. Equations

An equation without a label for cross-referencing:

$$E = mc^2$$

An inline equation: $y = ax + b$

An equation with a label for cross-referencing:

$$\int_0^{r_2} F(r, \varphi) dr d\varphi = 1 \quad (1)$$

This equation can be referenced as follows: Eq. 1

8. Inserting R figures

The code below creates a figure. The code is included in the output because `echo=TRUE`.

```
plot(1:10,main="Some data",xlab="Distance (cm)",
     ylab="Time (hours)")
```

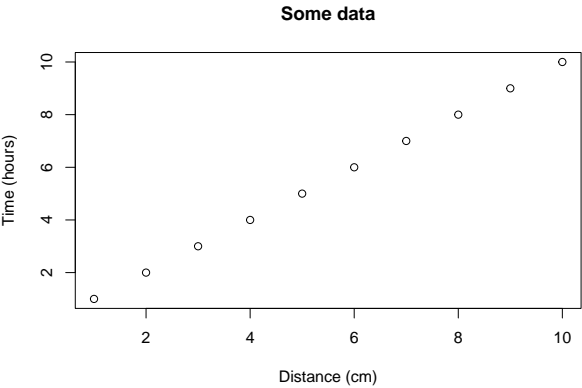


Fig. 1. This is the first figure.

You can reference this figure as follows: Fig. 1.

8.1. Figures spanning two-columns

Figures can span two columns be setting `fig.env="figure"`.
Reference to second figure: Fig. 2

Table 2. This is a kable table.

ID	code
1	a
2	b
3	c

Table 1. This is a xtable table.

	ID	code
1	1	a
2	2	b
3	3	c

You can reference this table as follows: Table 1.

9. Tables

9.1. Generate a table using xtable

```
df = data.frame(ID=1:3,code=letters[1:3])

# Creates tables that follow OUP guidelines
# using xtable
library(xtable)

## Warning: package 'xtable' was built under R version 4.2.1

print(xtable(df,caption="This is a xtable table.",
              label="tab:tab1"),
      comment=FALSE,caption.placement="top")
```

9.2. Generate a table using kable

```
df = data.frame(ID=1:3,code=letters[1:3])

# kable can also be used for creating tables
knitr::kable(df,caption="This is a kable table.",
              booktabs=TRUE,label="tab2")
```

Some wide data

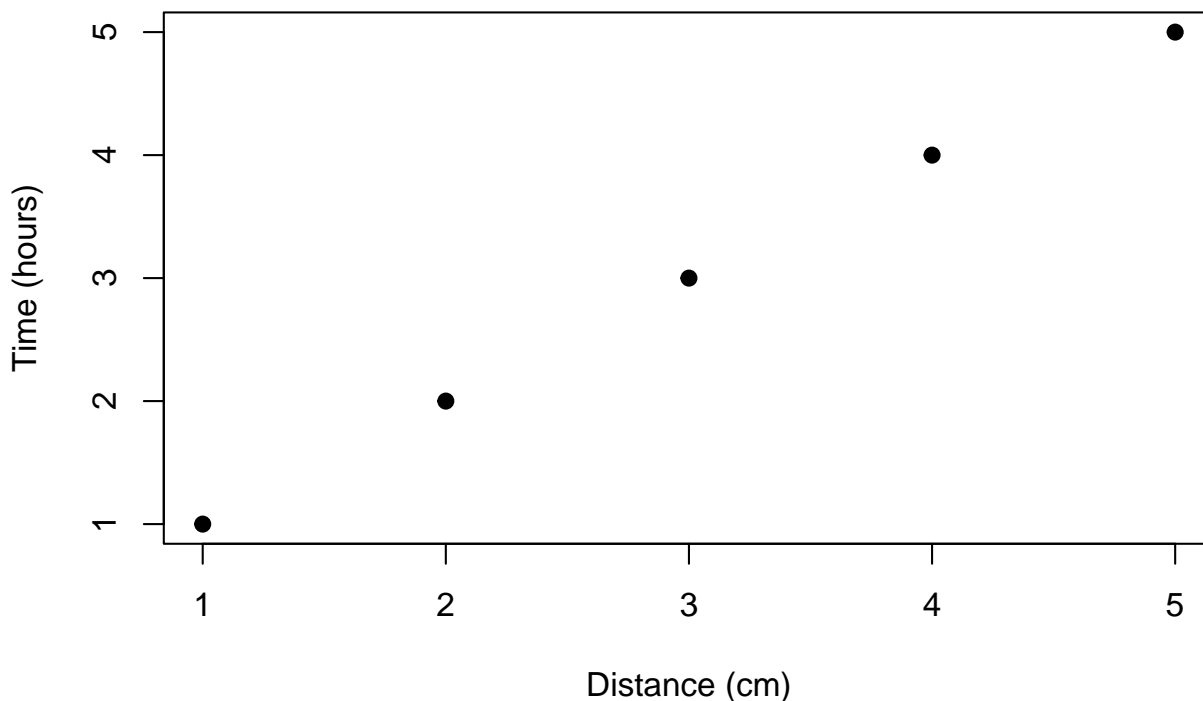


Fig. 2. This is a wide figure.

Table 3. This is a wide kable table.

ID	code1	code2	code3	code4	code5
1	a	d	g	j	m
2	b	e	h	k	n
3	c	f	i	l	o

You can reference this table as follows: Table 2.

9.3. Table spanning two columns

Tables can span two columns by setting `table.envir = "table"` in `knitr::kable`.

```
df = data.frame(ID=1:3,code1=letters[1:3],
               code2=letters[4:6],
               code3=letters[7:9],
               code4=letters[10:12],
               code5=letters[13:15])

# kable can also be used for creating tables
knitr::kable(df,caption="This is a wide kable table.",
             #format="latex",
             table.envir="table*",
             booktabs=TRUE,label="tab3")
```

10. Cross-referencing sections

You can cross-reference sections and subsections as follows: Section 6 and Section 5.1.

Note: the last section in the document will be used as the section title for the bibliography.

For more portable and flexible referencing of sections, equations, figures and tables, use `bookdown::pdf_document2` with YAML header option `base_format: rticles::oup_article`.

Appendices

A. Section title of first appendix

blabla

A.1. Subsection title of first appendix

and so on. . .

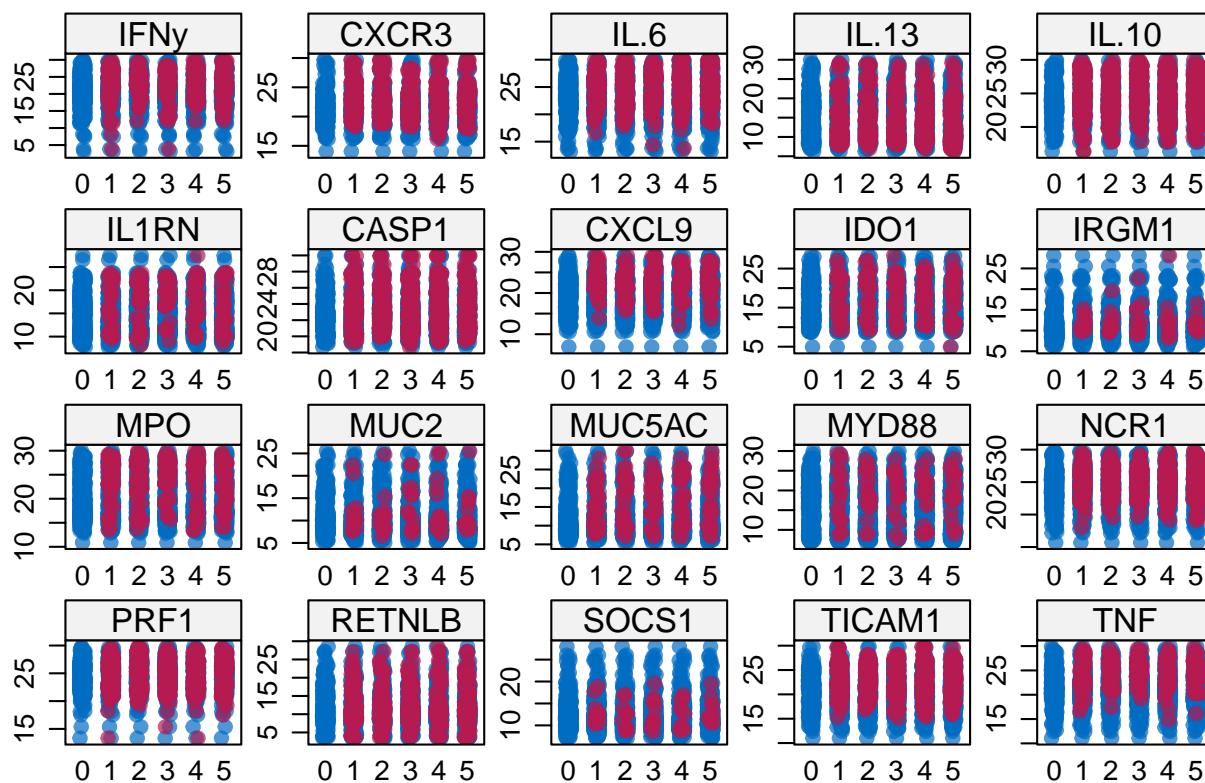


Fig. 3. Stripplot of observed and imputed data

11. References

12. Supplementary material

```
densityplot(igf, height = 1000, width = 800)
```

```
stripplot(igf, pch = 20, cex = 1.2)
```

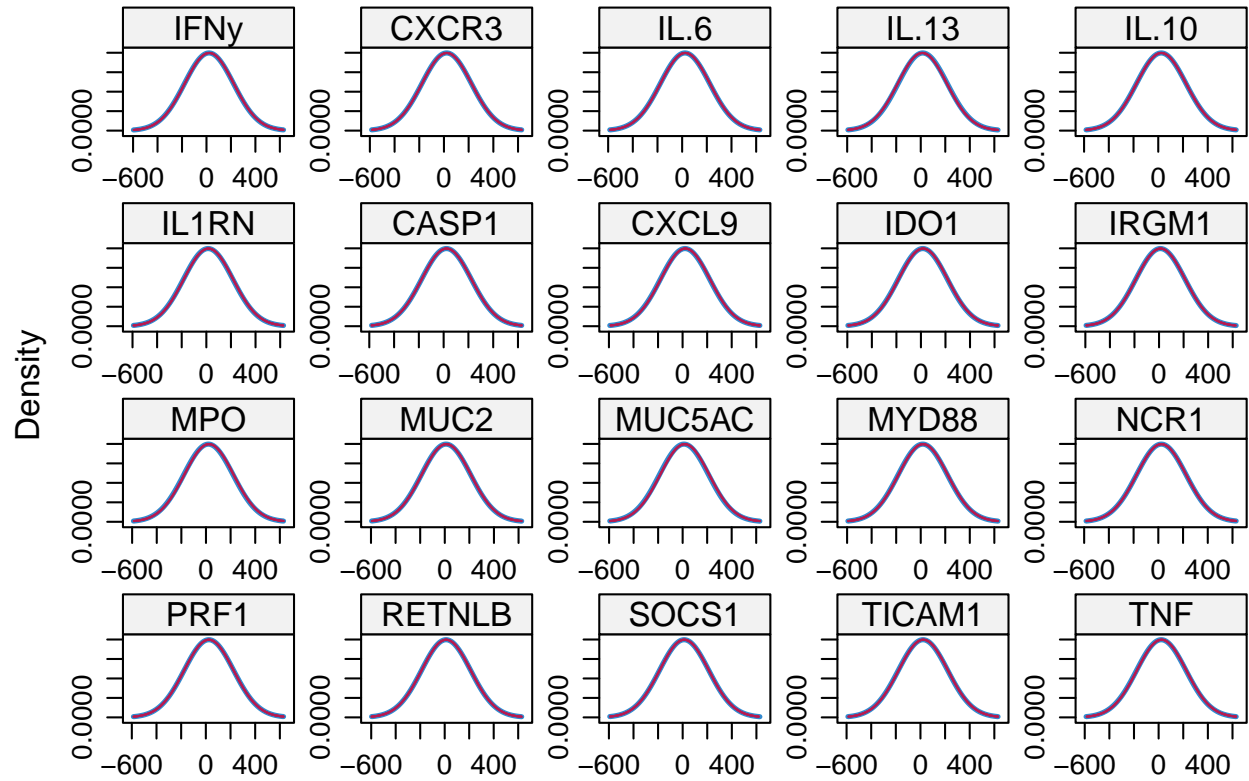


Fig. 4. Stripplot of observed and imputed data

13. Competing interests

There are no competing interest.

14. Author contributions statement

To be worked on

References

S. Van Buuren. *Flexible imputation of missing data*. CRC press, 2018.

S. Van Buuren and K. Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in r. *Journal of statistical software*, 45:1–67, 2011.

Fay Webster jkjb