

# Generation of Main Text and Supplementary Figures

Venelin Mitov

11 December 2017

=====

This script generates all figures in the main text as well as supplementary figures. The script contains chunks named as the names of the corresponding figures.

Before executing other chunks in the script, execute the following commands manually:

```
library(knitr)
library(rmarkdown)
# change the following path accordingly:
#setwd("~/SP")

# You can generate all figures at once by the following command:
#render('GenerateFigures.Rmd', output_format='pdf_document', clean=FALSE)
```

## Toy Model

Load rate functions from the ToyModelSetup script

Load h2statsAll table

Fig 2: Plot trait-value distribution

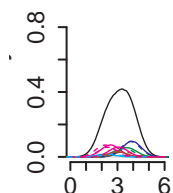
```
size <- 2e6
par(mfrow=c(1, 1))
setParCex()
par(mai=c(0.22, 0.3, 0.05, 0.05))

par(lwd=.7)

probstrain <- c(0.46, 0.43, 0.73, 0.12, 0.56, 0.24) #runif(6)
typestrain <- cbind(type=sample(1:2, size=size, replace=TRUE), strain=sample(1:6, size=size, replace=TRUE))
values <- as.data.table(cbind(typestrain, v=GEVs[typestrain[,1:2]]+rnorm(size, 0, .6)))

values[, plot(density(v), xlab=expression(z), ylab='Density', xlim=c(0,6), ylim=c(0, 0.8), main='', fra
cols <- c('magenta', 3:6, 'brown')

values[, {dens <- density(v); dens$y <- dens$y*length(v)/size; lines(dens, col=cols[unique(strain)], lty
```



```
plot(density(rnorm(size, 0, 0.6)), xlim=c(-3, 3), ylim=c(0, 0.8), xlab='e', main='', frame.plot=FALSE)
```

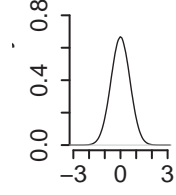


Fig 2: Plot toy-model dynamics

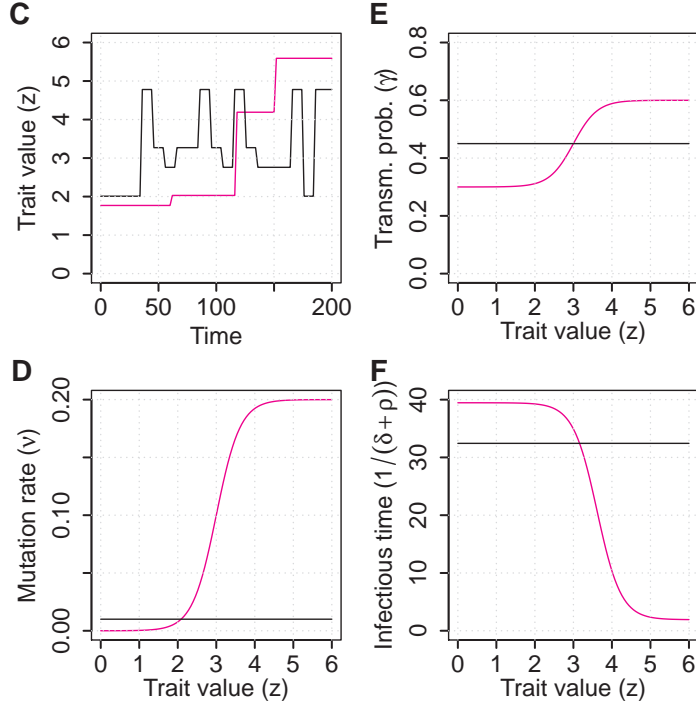
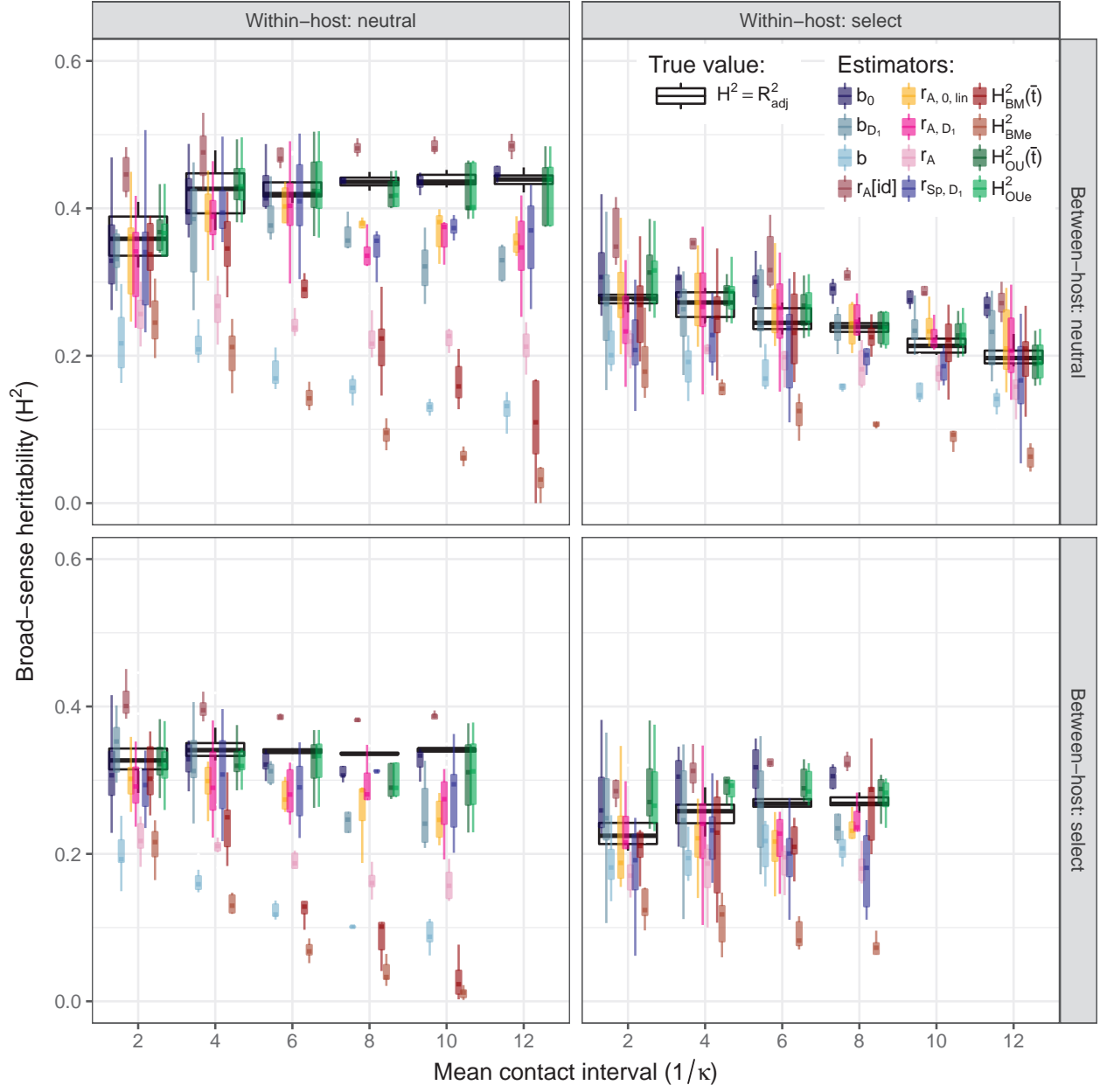


Table summaryToySim: Heritability estimate errors in toy model simulations

Table 1: SIR simulations of case 1: mean difference $\hat{H}^2 - H^2$									
process	N	$b[0]$	$b[D_1\tau]$	$b[\tau]$	$r_A[id]$	$r_A[PP : D_1\tau]$	$r_A[PP : \tau]$	$r_{Sp}[CPP : D_1\tau]$	$H_{EM}^2(t)$
neutral/neutral	50	-0.01(8.4e-03)	-0.07(1.1e-10)	-0.25(5.4e-32)	0.05(3.3e-29)	-0.05(7.1e-09)	-0.18(1.8e-31)	-0.05(3.6e-40)	-0.17(9.4e-31)
neutral/select	41	-0.02(3.1e-07)	-0.04(2e-05)	-0.2(1e-27)	0.05(2.9e-25)	-0.06(9.7e-10)	-0.15(1.8e-29)	-0.05(2.1e-31)	-0.17(6.2e-31)
select/neutral	47	0.05(3.1e-16)	0(5.2e-01)	-0.07(3e-25)	0.08(1.5e-28)	0(8.5e-01)	-0.06(1.6e-21)	-0.05(7.8e-27)	-0.01(4.6e-27)
select/select	37	0.04(2.4e-11)	-0.01(3.1e-01)	-0.06(2.1e-17)	0.06(3.4e-38)	-0.03(1e-03)	-0.08(8.1e-23)	-0.07(5.8e-21)	-0.04(1.9e-21)

Fig 4: Plot heritability estimates for each kappa



**Fig S3: tau in the toy-model**

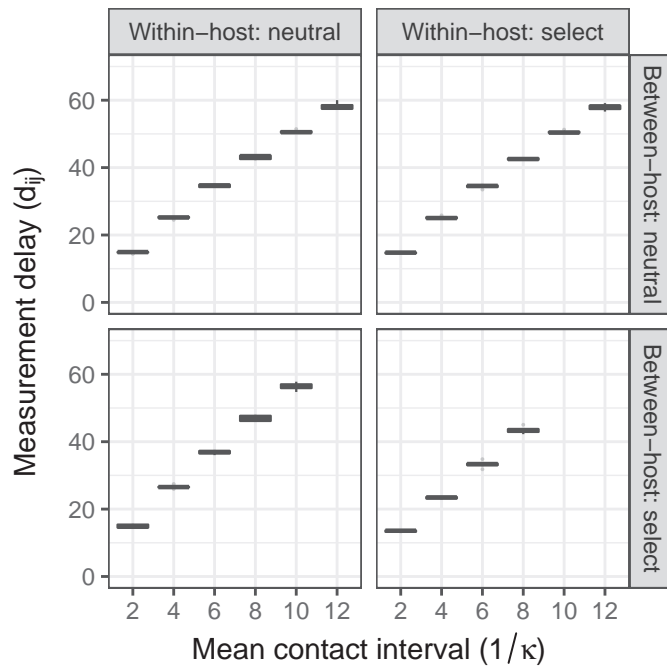


Fig S4: Genotypic variance in the toy-model

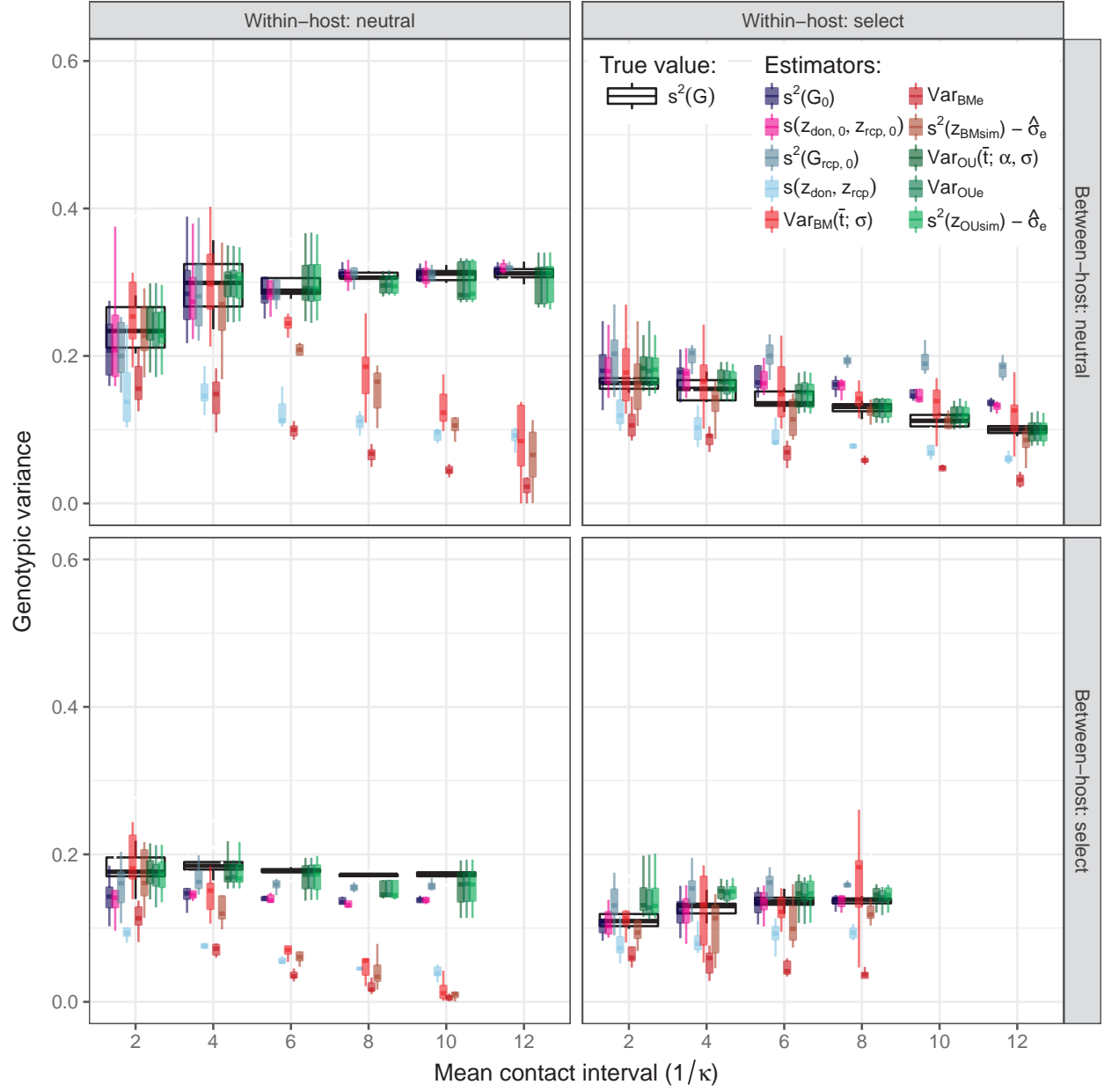


Fig S5 phenotypic variance in the toy model

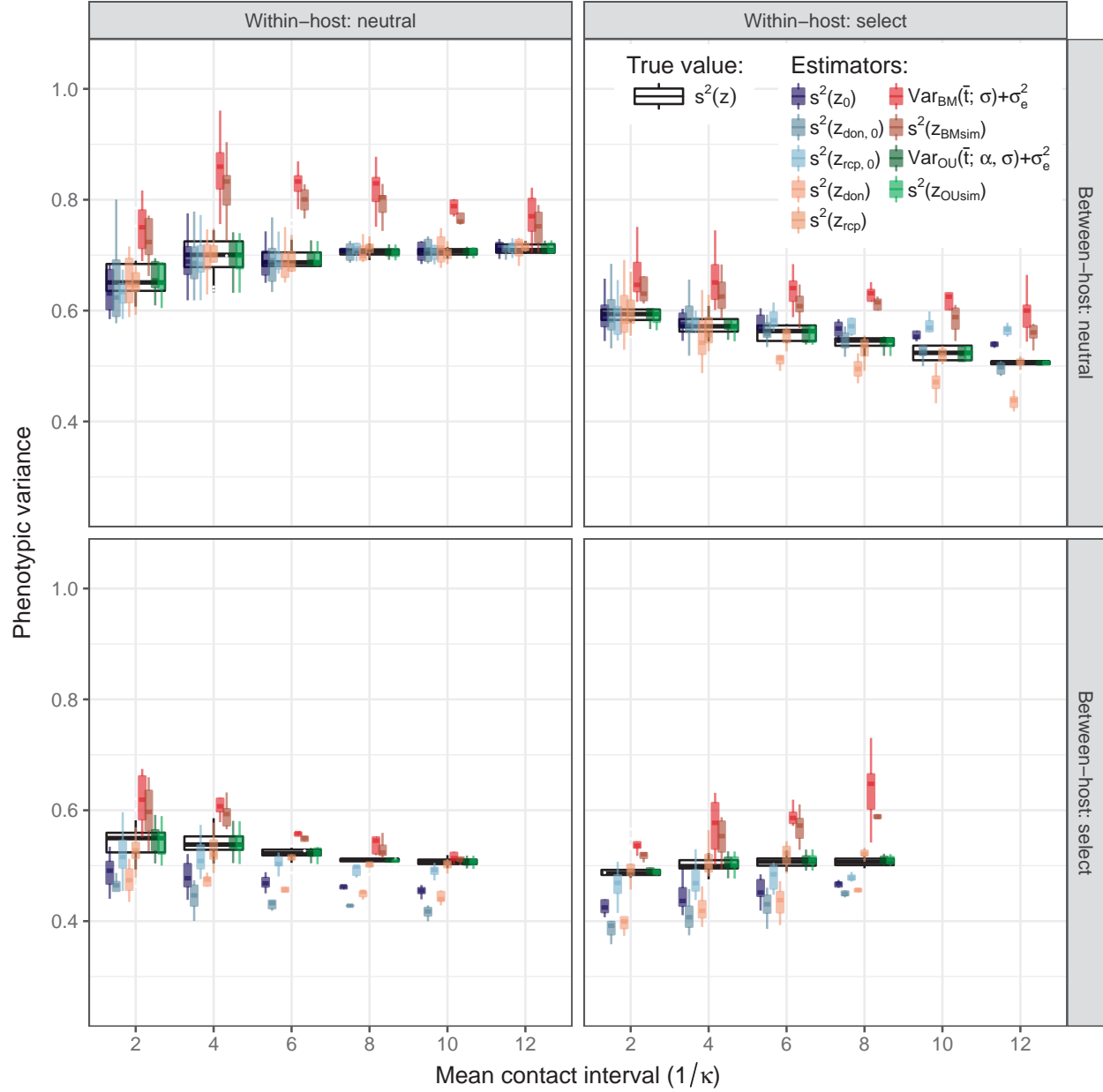
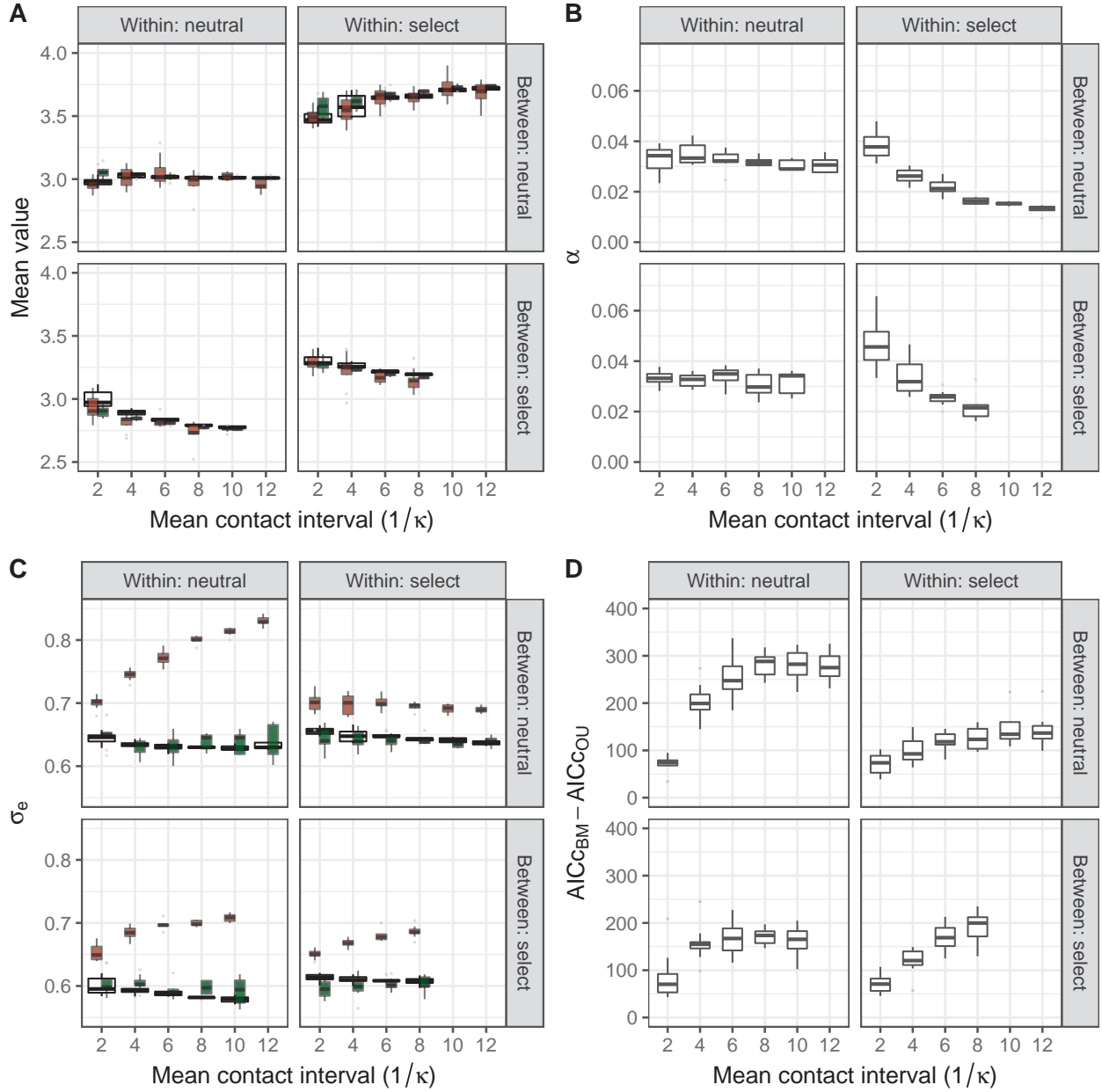


Fig S6: Plot tau, alpha, sigmae and AICc from PMM and POUMM

```
## Loading required package: cowplot
##
## Attaching package: 'cowplot'
## The following object is masked from 'package:ggplot2':
##
##     gggsave
```



**Fig S7 Plotting correlation decay as a function of tau**

The following chunk needs large data.table objects to be loaded in memory : simulEpidemicSIRNN, simulEpidemicSIRSN, simulEpidemicSIRNS, simulEpidemicSIRSS. To load these objects, the corresponding .RData files should either be generated by simulating the ToyModel (read file SimulateToyModel.Rmd) or downloaded. DATA/ToyModelSIR/NN.RData DATA/ToyModelSIR/SN.RData DATA/ToyModelSIR/NS.RData DATA/ToyModelSIR/SS.RData

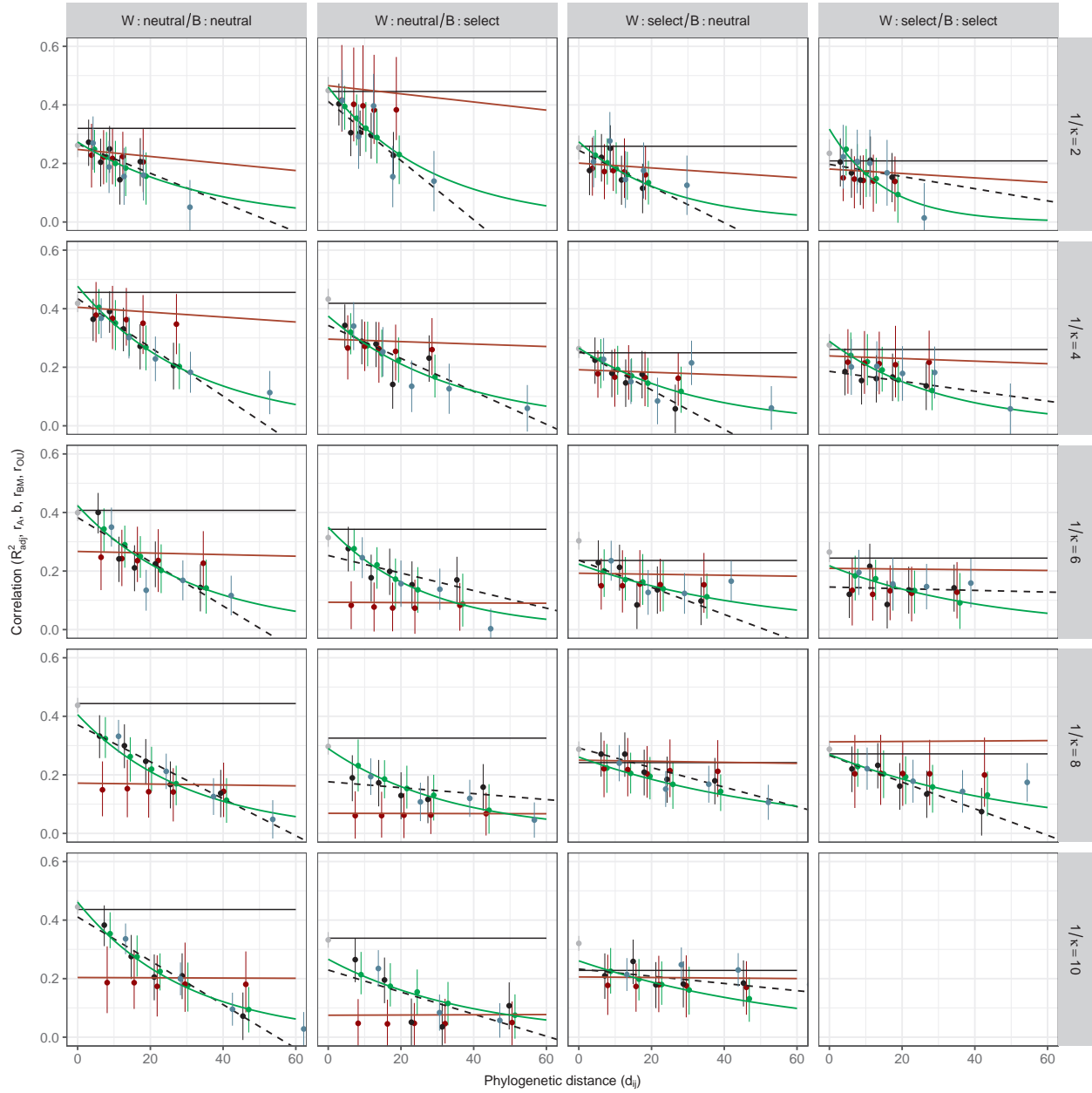
It is also possible to generate the above data-files by executing all toy-model simulations -> see the file "SimulateToyModel.Rmd"

```
GGally::ggmatrix(listPlots, nrow=5, ncol = 4, byrow = FALSE,
  yAxisLabels = c("1/kappa==2", "1/kappa==4", "1/kappa==6", "1/kappa==8", "1/kappa==10"),
  xAxisLabels = c("W: neutral / B: neutral", "W: neutral / B: select", "W: select / B: neutral", "W: select / B: select"))
```

```

labeller = "label_parsed",
xlab = expression(paste("Phylogenetic distance (", d[ij], ")")),
ylab = expression(paste("Correlation (", R[adj]^2, ", ", r[A], ", ", b, ", ", r[BM], "
theme(strip.text = element_text(size=9), axis.title = element_text(size=9))

```





## Analysis of HIV data

Load hivdata

Fig 5: Trait-values along phylogeny

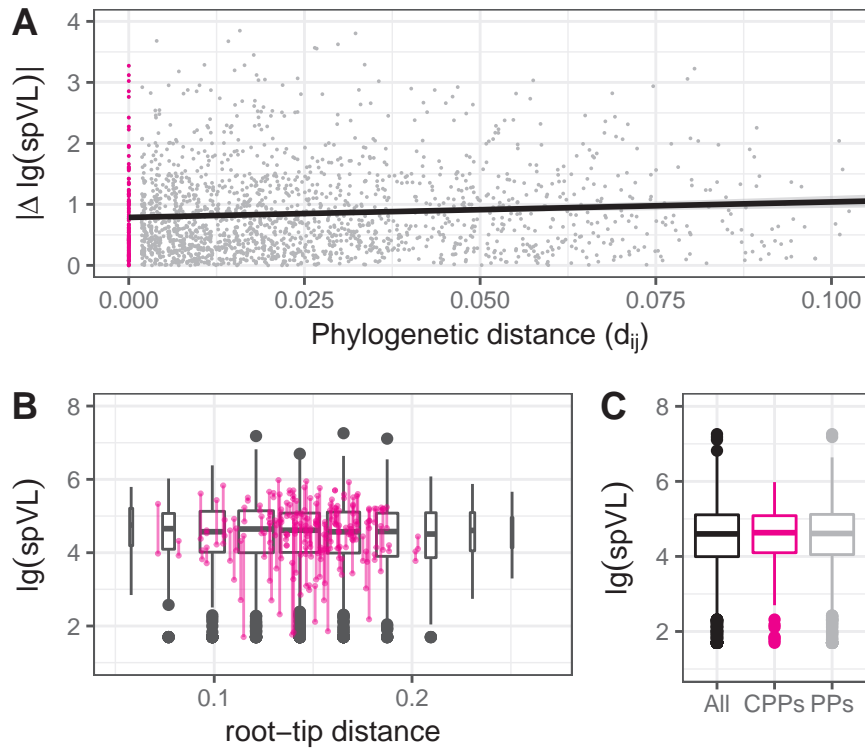


Fig 3: Correlation as a function of phylogenetic distance

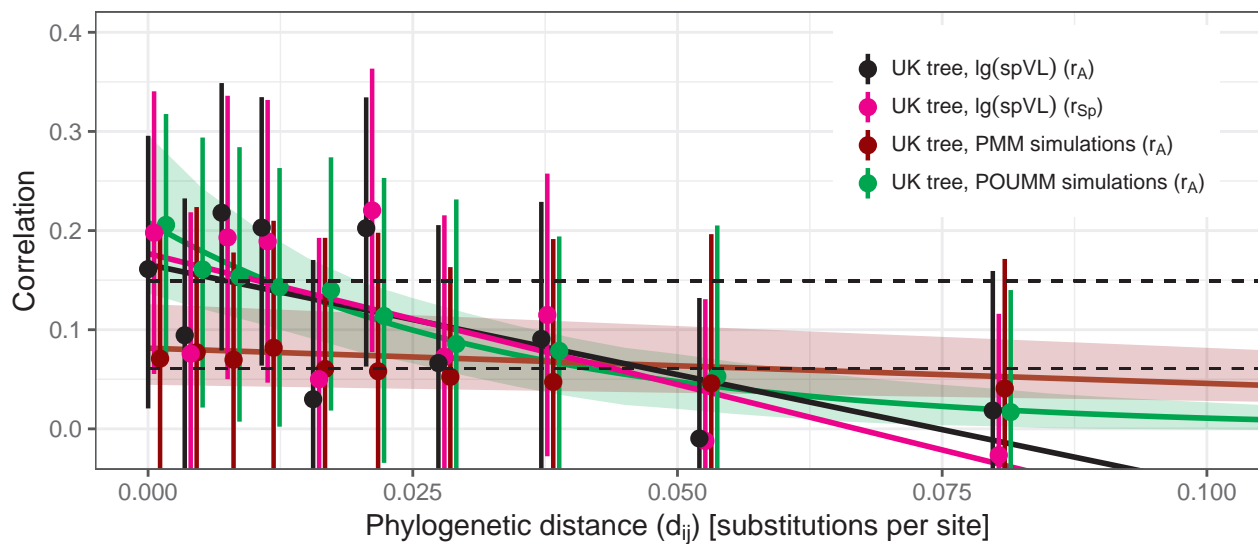


Fig S2: Covariance and variance as functions of  $t$  and  $t_{ij}$  in the UK data

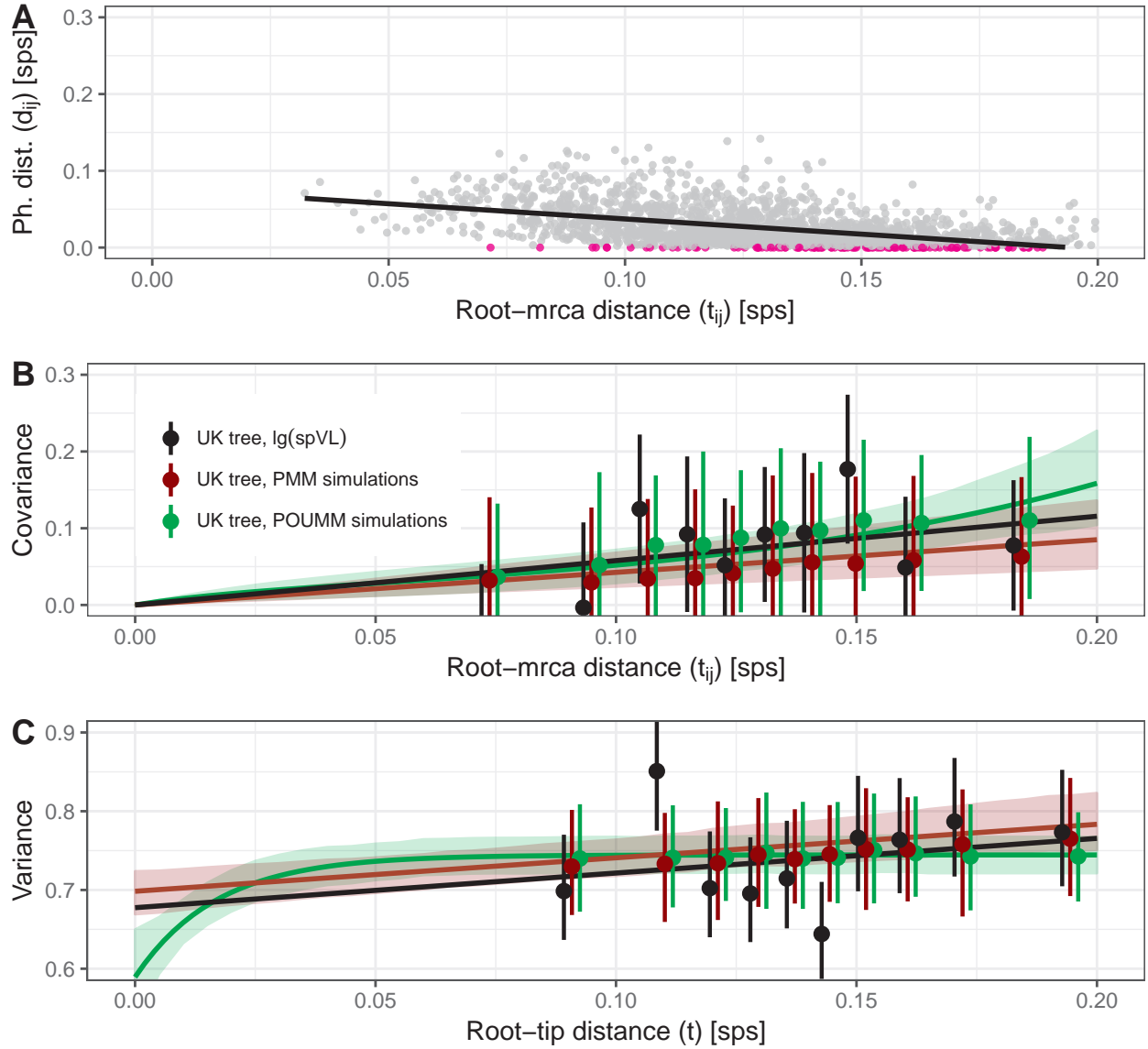


Fig S8: MCMC samples of POUMM parameters:  $\alpha$ ,  $\theta$ ,  $\sigma^2$ ,  $\sigma_e^2$

```
pl <- plot(hivdata$anH2[[1]]$fits$POUMM, doPlot = FALSE, doZoomIn = TRUE, stat=c("alpha", "theta", "sigma^2", "sigma_e^2"))
cowplot::plot_grid(pl$traceplot + scale_x_continuous(name = "iteration", labels = scales::scientific) +
  pl$densplot + theme_bw(), nrow = 2, labels=c("A", "B"))
```

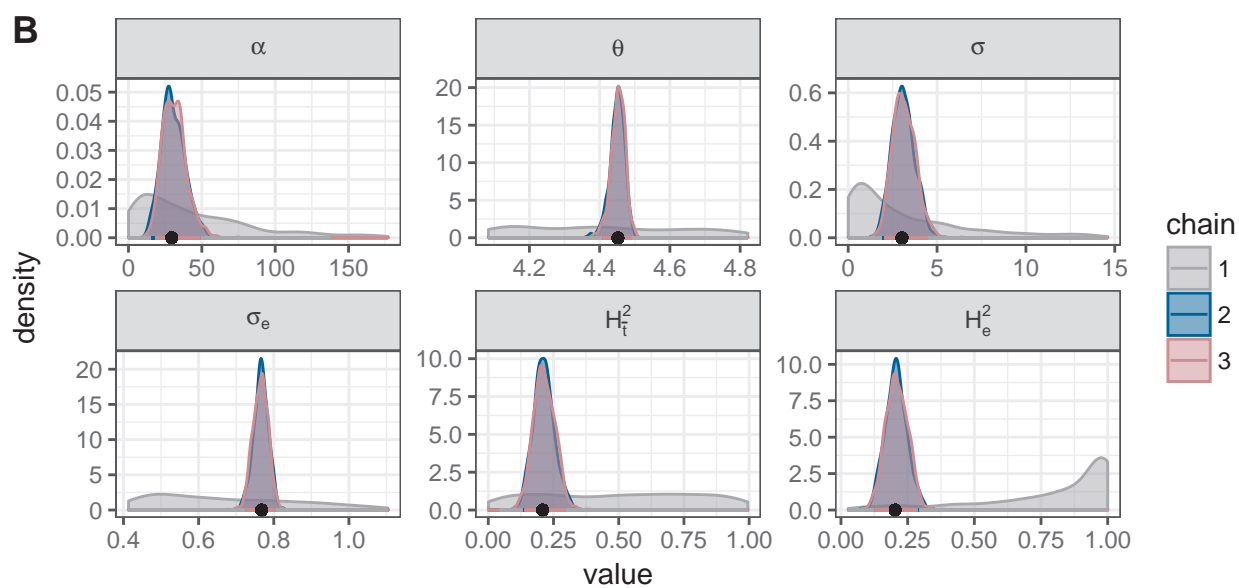
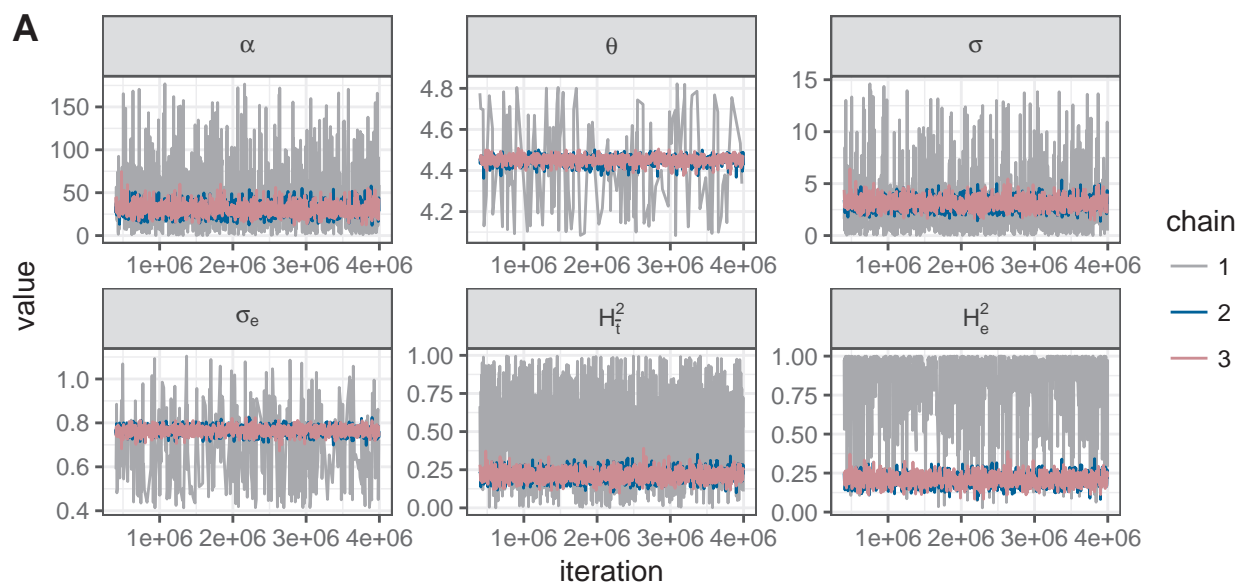
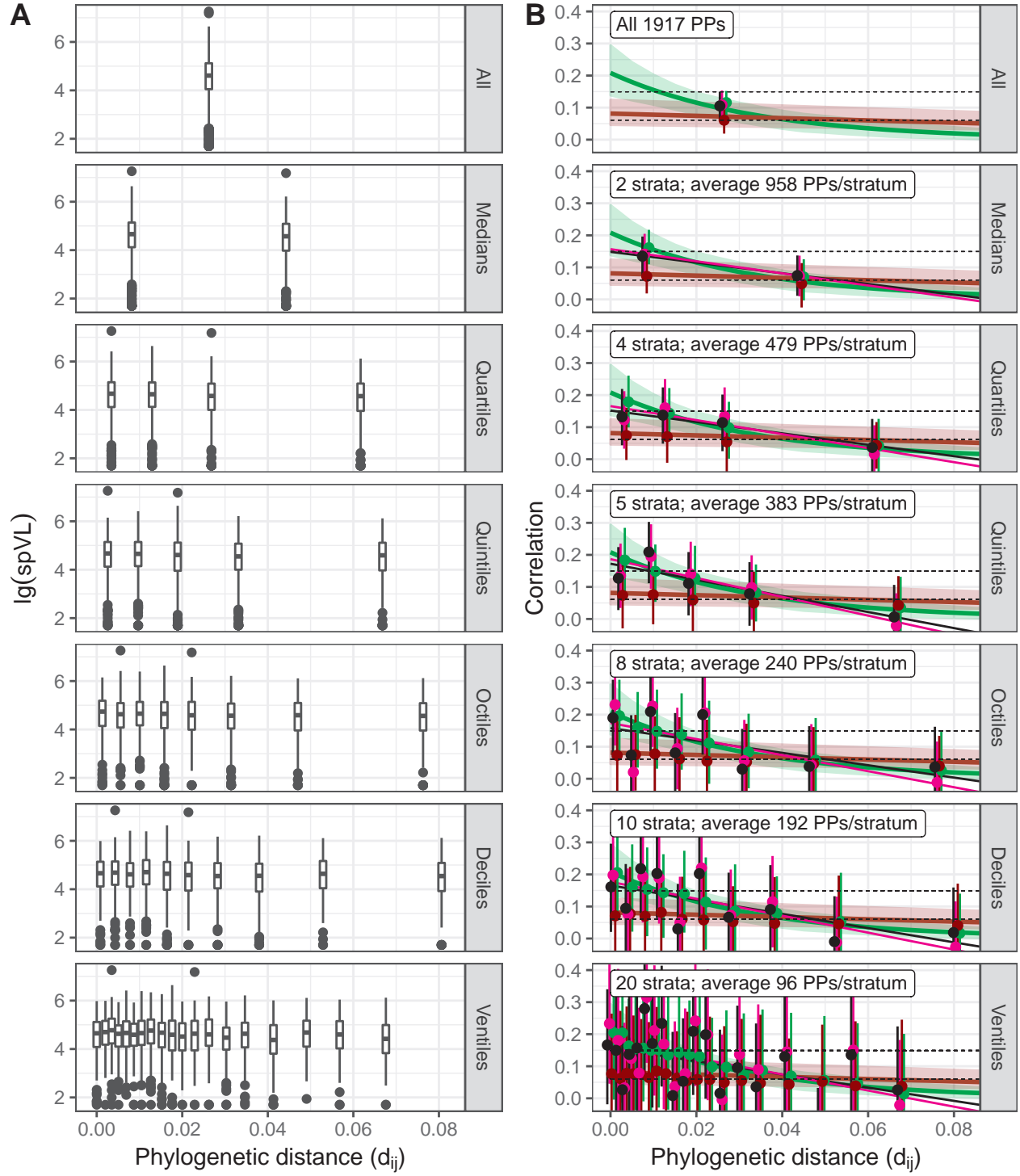
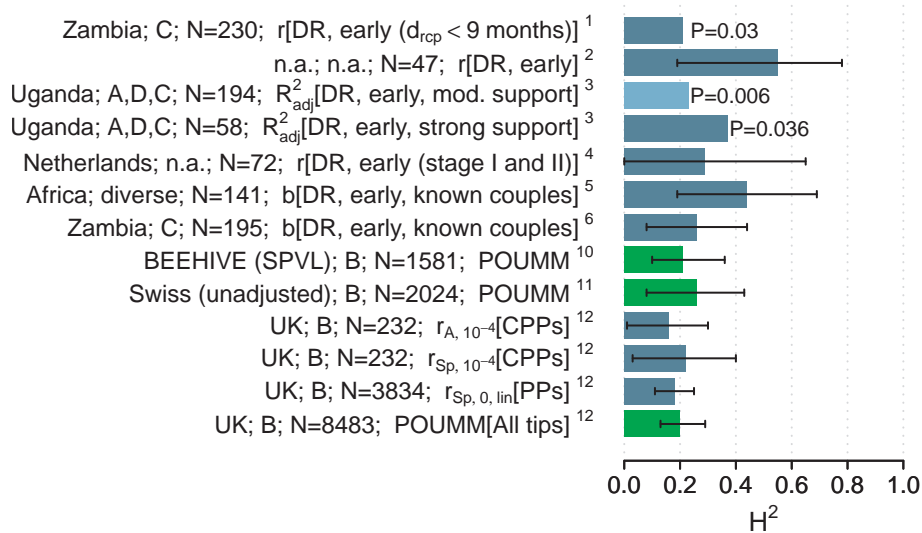


Fig S1: Different stratifications on tau

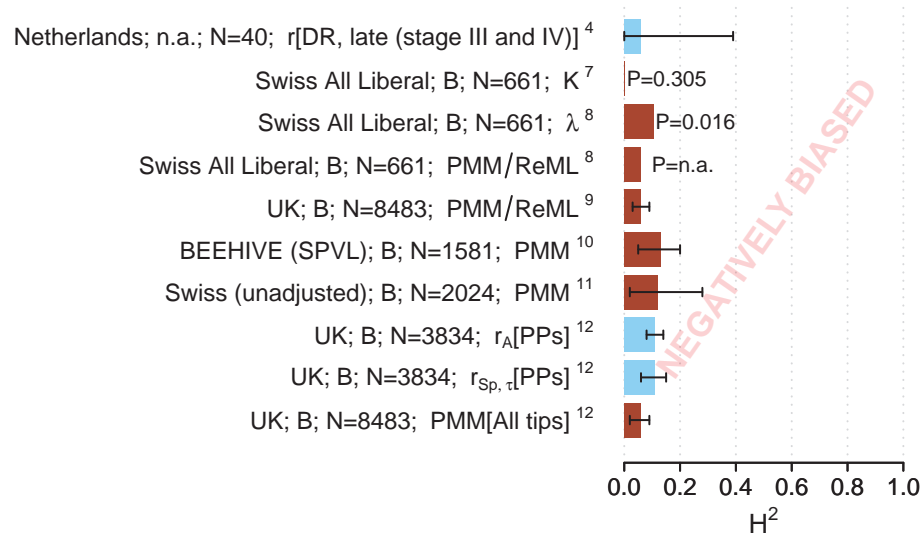


**Fig 6: Compare spVL-studies**

**A Cohort; Subtype; N; Method[data, filter]**



**B Cohort; Subtype; N; Method[data, filter]**



## Embedding fonts in pdf-figures

The embedding of the fonts didn't work smoothly. A solution that seems to work better is to drag-drop the figure in an MS powerpoint and then to right-click the image object and save as image (choosing .pdf as a format).

```
library(extrafont)
```

```
#font_import()
```

```
# To embed greek letter-fonts you also need the fontcm package
```

```
#font_install("fontcm")
```

```
loadfonts()
```

```

dir <- "GenerateFigures_files/figure-latex/"
figs <- c("Fig2C-F-1", "Fig3-1", "Fig4-1", "Fig5-1", "Fig6-1", "plot-random-distribution-1", "plot-random-distribution-2")

figsInLatex <- c(NA, "Fig3", "Fig4", "Fig5", "Fig6", NA, NA, "FigS1", "FigS2", "FigS3", "FigS4", "FigS5")

ext <- ".pdf"

for(i in 1:length(figs)) {
  infile <- paste0(dir, figs[i], ext)
  outfile <- paste0(dir, figs[i], "-embed", ext)
  print(infile)
  embed_fonts(infile, outfile = outfile)

  # store the figures with embedded fonts in the directory containing the manuscript latex file

  if(!is.na(figsInLatex[i])) {
    file.copy(outfile, paste0("../MS/", figsInLatex[i], ext), overwrite = TRUE)
  }
}

```