There are some things that are great about the tidyverse family of R packages and the style they encourage. There are also a few gotchas. Here’s a reminder to myself about [this phenomenon](https://onunicornsandgenes.blog/2017/02/12/using-r-tibbles-and-the-t-test-function/): tidyverse-style data frames (”tibbles”) do not simplify to vectors upon extracting a single column with hard bracket indexing.

Because some packages rely on specific *data.frame* behaviours that tibbles don’t show, functions that work nicely with data frames, and normally have nice interpretable error messages, may mysteriously collapse in all kinds of ways when fed a tibble.

Here’s an example with *MCMCglmm*. This is not to pick on *MCMCglmm*; it just happened to be one of the handful of packages where I’ve run into this issue. Here, we use *readr*, the tidyverse alternative to the *read.table* family of functions to read some simulated data. The base function is called *read.csv*, and the readr alternative is *read\_csv*.

Reading in tabular data is a surprisingly hard problem: tables can be formatted in any variety of obnoxious ways, and the reading function also needs to be fast enough to deal with large files. Using *readr* certainly isn’t always painless, but it reduces the friction a lot compared to read.table. One of the improvements is that *read\_csv* will return a data.frame with the class *tbl\_df*, affectionately called ”[tibble](https://tibble.tidyverse.org/" \t "_blank)”

After reading the data, we centre and scale the trait, set up some priors and run an animal model. Unfortunately, *MCMCglmm* will choke on the tibble, and deliver a confusing error message.

library(MCMCglmm)

library(readr)

ped <- read\_csv("sim\_ped.csv")

pheno <- read\_csv("sim\_pheno.csv")

pheno$scaled <- scale(pheno$pheno)

prior\_gamma <- list(R = list(V = 1, nu = 1),

G = list(G1 = list(V = 1, nu = 1)))

model <- MCMCglmm(scaled ~ 1,

random = ~ animal,

family = "gaussian",

prior = prior\_gamma,

pedigree = ped,

data = pheno,

nitt = 100000,

burnin = 10000,

thin = 10)

Error in inverseA(pedigree = pedigree, scale = scale, nodes = nodes) :

individuals appearing as dams but not in pedigree

In addition: Warning message:

In if (attr(pedigree, "class") == "phylo") { :

the condition has length > 1 and only the first element will be used

In this pedigree, it is not the case that there are individuals appearing as dams but not listed. If we turn the data and pedigree into vanilla data frames instead, it will work:

ped <- as.data.frame(ped)

pheno <- as.data.frame(pheno)

model <- MCMCglmm(scaled ~ 1,

random = ~ animal,

family = "gaussian",

prior = prior\_gamma,

pedigree = ped,

data = pheno,

nitt = 100000,

burnin = 10000,

thin = 10)

MCMC iteration = 0

MCMC iteration = 1000

MCMC iteration = 2000