

Fixing up our range plots

Range Recap

Last week we made a plot of ranges through time for the canid fossil record

```
library(ggplot2)
```

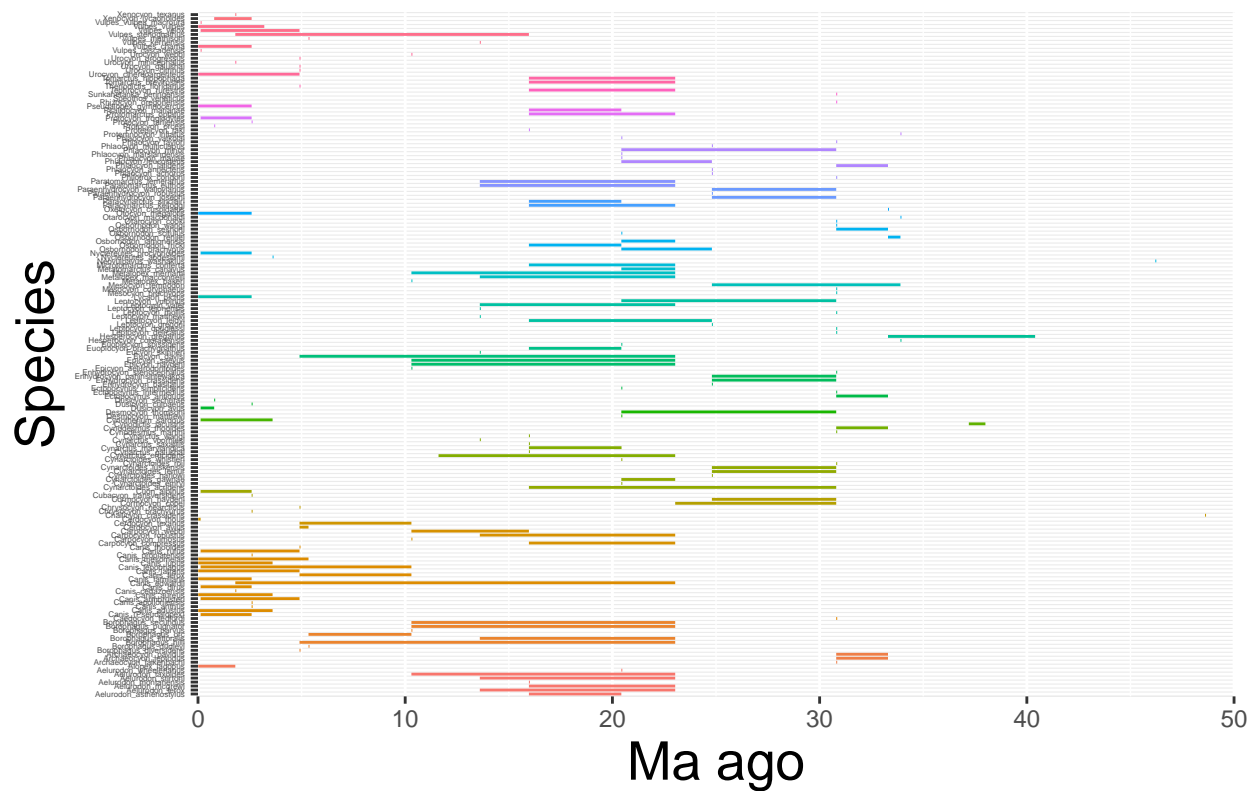
```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```
setwd("/Users/michael_alfaro/Dropbox/teaching/EEB 177:234 2016/range-plotting/") #you will need to chan
canids <- read.csv("/Users/michael_alfaro/Dropbox/teaching/EEB 177:234 2016/range-plotting/canid_fa_la.
names(canids) <- c("genus", "species", "minage", "maxage")
head(canids)
```

```
##      genus      species minage maxage
## 1 Aelurodon Aelurodon_asthenostylus 15.97 20.43
## 2 Aelurodon      Aelurodon_ferox 13.60 23.03
## 3 Aelurodon      Aelurodon_mcgrewi 15.97 23.03
## 4 Aelurodon Aelurodon_montanensis 15.97 15.97
## 5 Aelurodon      Aelurodon_stirtoni 13.60 23.03
## 6 Aelurodon      Aelurodon_taxoides 10.30 23.03
```

```
canid_occ <- ggplot(canids, aes( x = species, maxage, colour = genus))
# everything
canid_occ + geom_linerange(aes(ymin = minage, ymax = maxage)) + theme(legend.position="none") + coord_
```

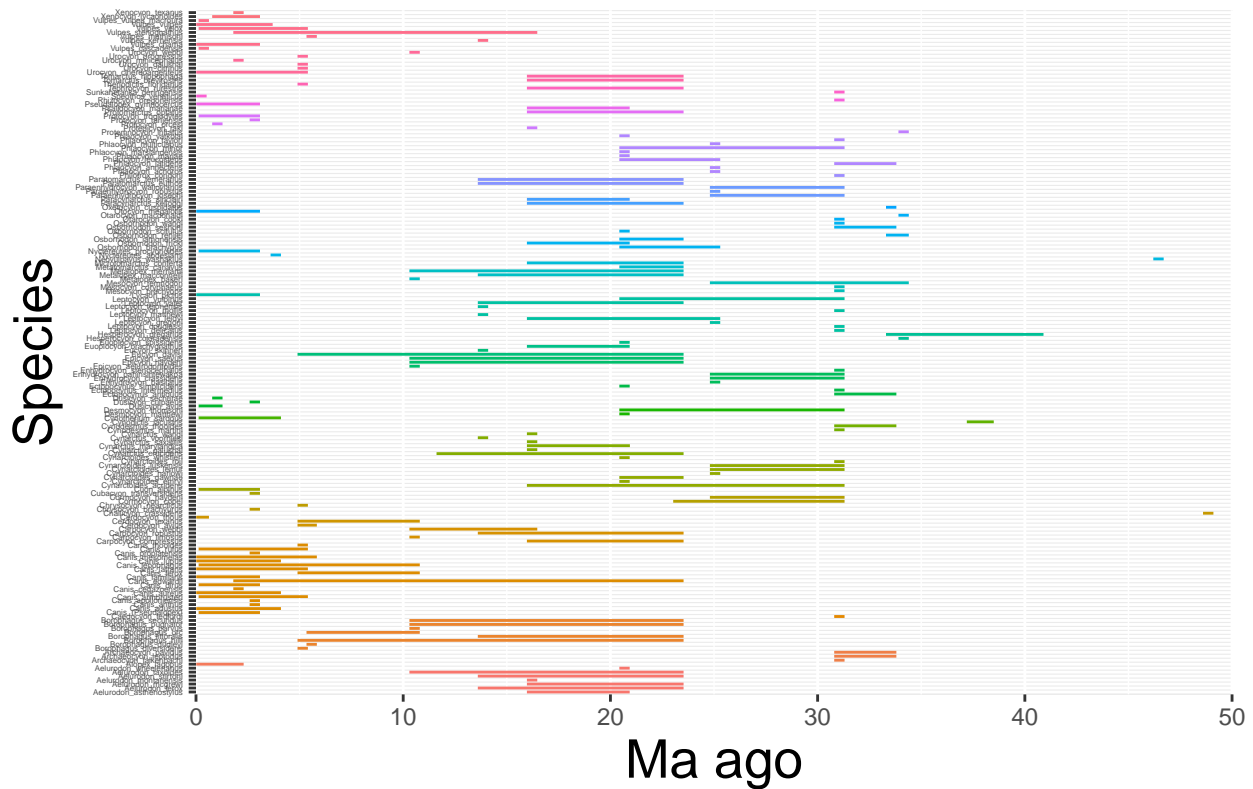
Canid Fossil Occurrences



One problem with this plot is that the point occurrences do not show up well. Lets stretch them just a little bit so we can see them better by setting maxage to be the original max age + 500,00 years

```
canid_occ <- ggplot(canids, aes( x = species, maxage, colour = genus))
# everything
canid_occ + geom_linerange(aes(ymin = minage, ymax = maxage + 0.5)) + theme(legend.position="none") +
```

Canid Fossil Occurrences



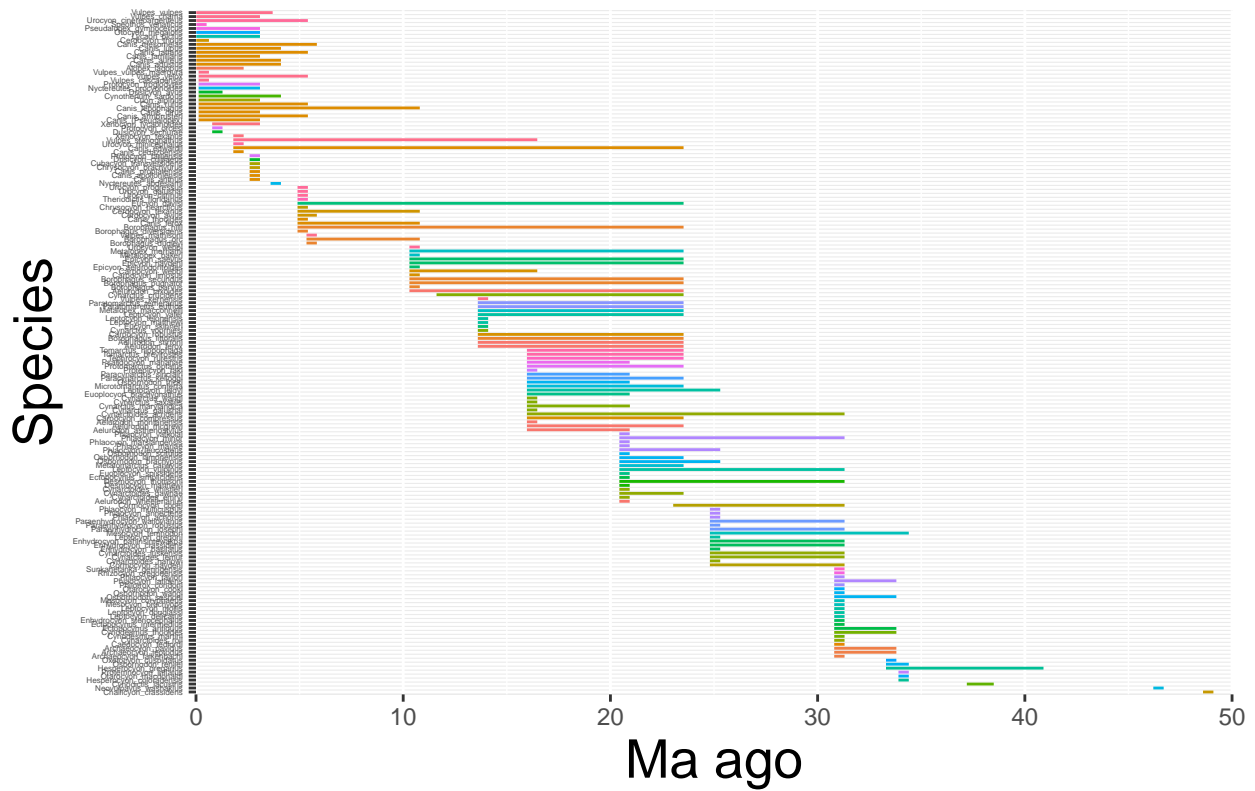
A second problem is that the age ranges are not sorted in any way. The output we saw from paleotree on Tuesday was much nicer. Lets try to fix this. We are going to use an R package called forcats (for categorical variables,) to reorder the species by their minimum age. This will force ggplot to arrange the ranges so that the youngest are on top and the oldest are on the bottom.

```
library(forcats)
```

```
## Warning: package 'forcats' was built under R version 3.3.2
```

```
canid_occ <- ggplot(canids, aes( x = fct_reorder(species, minage, .desc = T), maxage, colour = genus))
canid_occ + geom_linerange(aes(ymin = minage, ymax = maxage + 0.5)) + theme(legend.position="none") +
```

Canid Fossil Occurrences

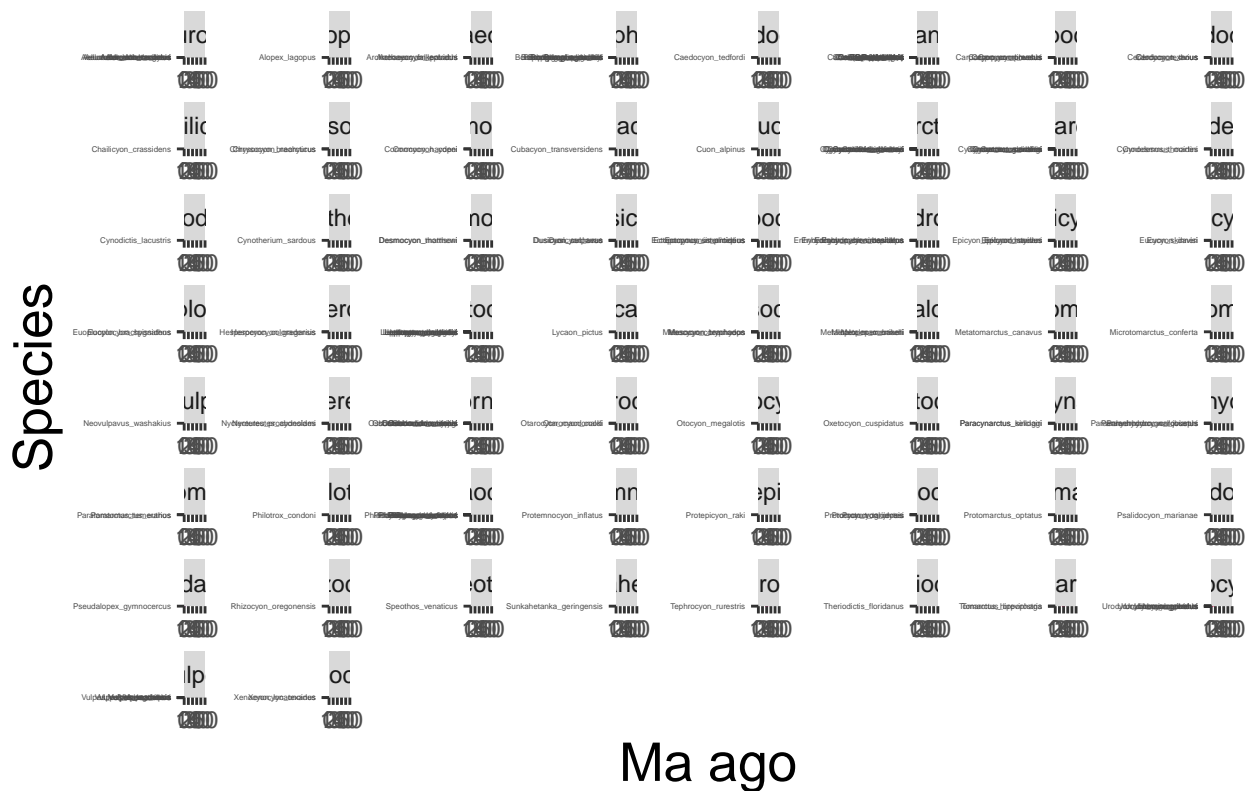


This looks even better than the paleotree output because we can see the species names and we can control the labels and titles.

What if we wanted to see these range plots broken up by genus? In ggplot this is easily done with `facet_wrap`

```
canid_occ + geom_linerange(aes(ymin = minage, ymax = maxage + 0.5)) + theme(legend.position="none") +
```

Canid Fossil Occurrences



```
ggsave("faceted-ranges.pdf", scale = 6)
```

```
## Saving 39 x 27 in image
```

We don't have many records for some genera so this plot looks bad in the plotting window. However it might still be useful in a data exploration phase. We can save it as a pdf and look at the cells in more details.

Finally, on Tuesday we saw how paleotree could make a disparity through time plot from the occurrence data. Can we do this directly in R with our data frame? The answer is yes using the dplyr and tidyr packages.

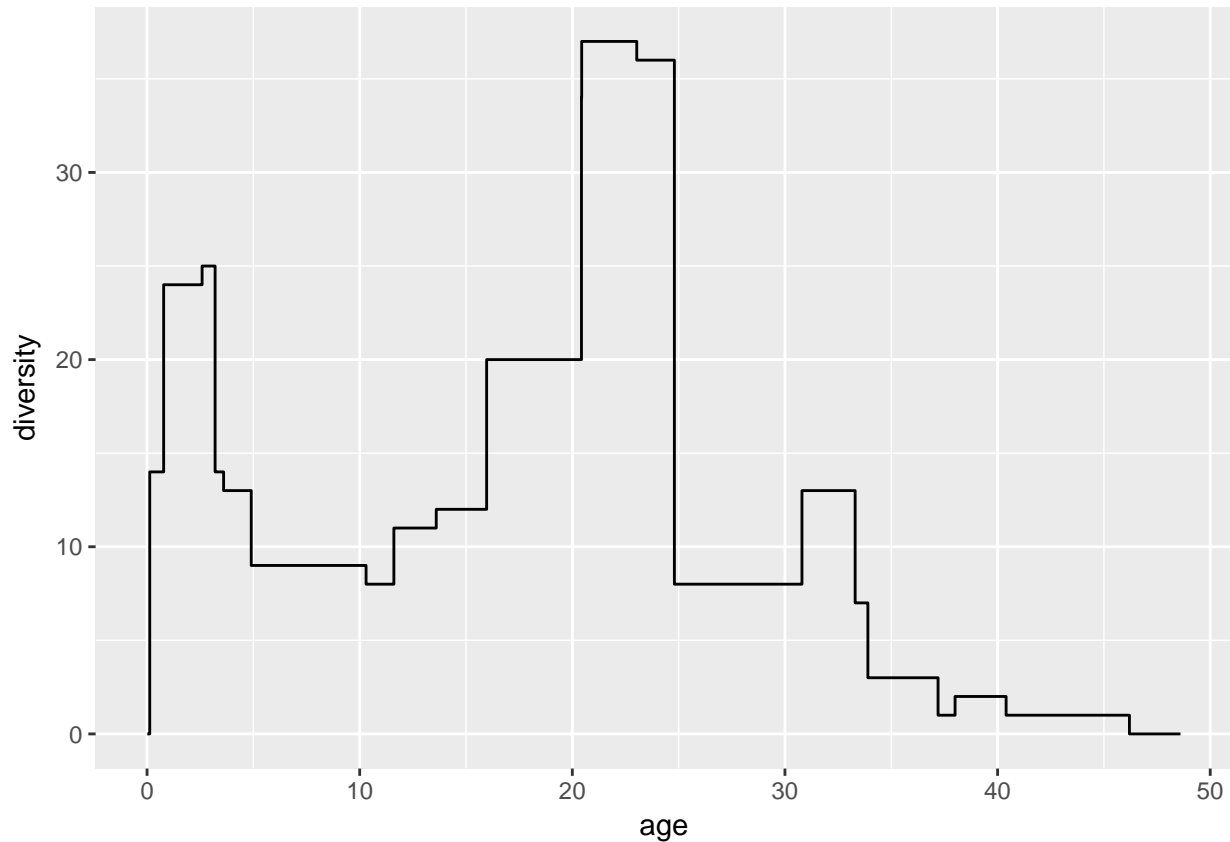
```
library(tidyr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
diversity <- canids %>% gather(key = type, value = age, minage, maxage) %>% mutate(count = ifelse(type == "minage", 1, -1))
ggplot(diversity, aes(x = age, y = diversity)) + geom_step()
```



What just happened? Let's break this down. dplyr and tidyr support the use of chaining (you can think of this like the pipe command we learned in the shell) with the %>% character. This allows us to send the output of one operation as the input for another.

```
#create a new column in the data frame that contains all of the min and max age values
diversity <- canids %>% gather(key = type, value = age, minage, maxage)
```

```
head(diversity)
```

```
##      genus      species  type  age
## 1 Aelurodon Aelurodon_asthenostylus minage 15.97
## 2 Aelurodon      Aelurodon_ferox minage 13.60
## 3 Aelurodon      Aelurodon_mcgrewi minage 15.97
## 4 Aelurodon Aelurodon_montanensis minage 15.97
## 5 Aelurodon      Aelurodon_stirtoni minage 13.60
## 6 Aelurodon      Aelurodon_taxoides minage 10.30
```

```
#make minimum ages count as 1 (species richness increases at that time) and max ages count as -1 (species richness decreases at that time)
diversity <- diversity %>% mutate(count = ifelse(type == "maxage", 1, -1))
```

```
head(diversity)
```

```
##      genus      species  type  age count
## 1 Aelurodon Aelurodon_asthenostylus minage 15.97 -1
## 2 Aelurodon      Aelurodon_ferox minage 13.60 -1
## 3 Aelurodon      Aelurodon_mcgrewi minage 15.97 -1
## 4 Aelurodon Aelurodon_montanensis minage 15.97 -1
## 5 Aelurodon      Aelurodon_stirtoni minage 13.60 -1
## 6 Aelurodon      Aelurodon_taxoides minage 10.30 -1
```

```
#put all taxa with the same min age or max age in the same group
diversity <- diversity %>% group_by(age)
head(diversity)
```

```
## Source: local data frame [6 x 5]
## Groups: age [3]
##
##      genus      species  type  age count
##      <chr>      <chr> <chr> <dbl> <dbl>
## 1 Aelurodon Aelurodon_asthenostylus minage 15.97 -1
## 2 Aelurodon      Aelurodon_ferox minage 13.60 -1
## 3 Aelurodon      Aelurodon_mcgrewi minage 15.97 -1
## 4 Aelurodon Aelurodon_montanensis minage 15.97 -1
## 5 Aelurodon      Aelurodon_stirtoni minage 13.60 -1
## 6 Aelurodon      Aelurodon_taxoides minage 10.30 -1
```

```
# for each age count the number of taxa
diversity <- diversity %>% summarise(count = sum(count))
head(diversity)
```

```
## # A tibble: 6 × 2
##   age count
##   <dbl> <dbl>
## 1 0.0117 -14
## 2 0.1260 -10
## 3 0.7810 0
## 4 1.8000 -1
## 5 2.5880 11
## 6 3.2000 1
```

```
#sort this
diversity <- diversity %>% arrange(-age, -count)
head(diversity)
```

```
## # A tibble: 6 × 2
##   age count
##   <dbl> <dbl>
## 1 48.6 0
## 2 46.2 0
## 3 40.4 1
## 4 38.0 1
## 5 37.2 -1
## 6 33.9 2
```

```
#get the cumulative sum
diversity <- diversity %>% mutate(diversity = cumsum(count))
head(diversity)
```

```
## # A tibble: 6 × 3
##   age count diversity
##   <dbl> <dbl>     <dbl>
## 1  48.6     0         0
## 2  46.2     0         0
## 3  40.4     1         1
## 4  38.0     1         2
## 5  37.2    -1         1
## 6  33.9     2         3
```

```
#plot it!
ggplot(diversity, aes(x = age, y = diversity)) + geom_step()
```

