Exploring Fractions

Alison Hoyt

October 5, 2018

Join Fraction Data to relevant data from other levels of the hierarchy

```
#Create flat data to work with:
frc_data <- ISRaD_data$fraction %>% as_tibble() %>% #Start with fraction data
  left_join(ISRaD_data$layer) %>% #Join to layer data
  left_join(ISRaD_data$profile) %>% #Join to profile data
  left_join(ISRaD_data$site) %>% #Join to site data
  left_join(ISRaD_data$metadata) #Join to metadata

#Take a look at it:
View(frc_data)
```

Fill in missing delta 14C values from fraction modern

```
#Example of how to add a new column to your data frame or tibble:
#note will add the column temporarily, but won't save it unless you assign it to an R object
frc_data %>%
  dplyr::mutate(frc_pmC = frc_fraction_modern*100)

#First fill in observation year across levels
#Add a column using mutate
#Use if_else(is.na, true statement, false statement) to check for missing values
frc_data <- frc_data %>%
  dplyr::mutate(frc_obs_date_y =  if_else(is.na(frc_obs_date_y), lyr_obs_date_y, frc_obs_date_y))


#Then calculate delta 14C from fraction modern
#delta14C=[fraction_modern*exp(lambda*(1950-Yc))-1]*1000
#Where lambda is 1/(true mean-life) of radiocarbon = 1/8267 = 0.00012097
#Yc is year of collection.
lambda <- .00012097

#Fill in lyr_14c using fraction modern
#First calculate the values
lyr_14c_from_fraction_modern <- (frc_data$lyr_fraction_modern * exp(lambda*(1950-frc_data$lyr_obs_date_y
#Then fill in NA values
frc_data <- frc_data %>%
  dplyr::mutate(lyr_14c = if_else(is.na(lyr_14c), lyr_14c_from_fraction_modern, lyr_14c))

#Repeat for layers:
frc_14c_from_fraction_modern <- (frc_data$frc_fraction_modern * exp(lambda*(1950-frc_data$frc_obs_date_y
frc_data <- frc_data %>%
  dplyr::mutate(frc_14c = if_else(is.na(frc_14c), frc_14c_from_fraction_modern, frc_14c))
```

Working with a reduced dataset - filter and summarise

```
#Select more limited data columns to work with:
frc_data_small <- frc_data %>%
  select(ends_with("name"), ends_with("obs_date_y"), ends_with("_14c"), lyr_top, lyr_bot, "frc_input", "
  unique() #note: unique reduces from 1712 to 1665. what type of entries would we be removing here?


#Group and look at summary statistics by land cover
frc_data_small %>%
  group_by(pro_land_cover) %>%
  summarise(num_data_points = n(),
            mean_frc_14c = mean(frc_14c, na.rm=TRUE))
```

```
## # A tibble: 8 x 3
##   pro_land_cover      num_data_points mean_frc_14c
##   <fct>                         <int>        <dbl>
## 1 bare                              1     -890
## 2 cultivated                      145     -213.
## 3 forest                          848      -27.0
## 4 rangeland/grassland             186       45.3
## 5 shrubland                       510     -259.
## 6 tundra                            1     -901
## 7 wetland                          33     -216.
## 8 <NA>                            267        0.369
```

```
#Same, but only soil above 20cm
frc_data_small %>%
  filter(lyr_bot < 20) %>%
  group_by(pro_land_cover) %>%
  summarise(num_data_points = n(),
            mean_frc_14c = mean(frc_14c, na.rm=TRUE))
```

```
## # A tibble: 8 x 3
##   pro_land_cover      num_data_points mean_frc_14c
##   <fct>                         <int>        <dbl>
## 1 bare                              1      -890
## 2 cultivated                       36       -97.5
## 3 forest                          558        62.2
## 4 rangeland/grassland             125        59.5
## 5 shrubland                       133        -9.28
## 6 tundra                            1      -901
## 7 wetland                           6       169.
## 8 <NA>                            133        43.6
```

```
#Check out number and mean values for different fractions
frc_data_small %>%
  group_by(frc_scheme, frc_property) %>%
  summarise(num = n(),
            mean_frc_14c = mean(frc_14c, na.rm=TRUE))
```

```
## # A tibble: 26 x 4
## # Groups:   frc_scheme [?]
##    frc_scheme      frc_property     num mean_frc_14c
##    <fct>           <fct>          <int>        <dbl>
```

```
##  1 Acid            acid insoluble    54       -149.
##  2 Acid            <NA>              34        184.
##  3 Acid_Evolution  carbonate         11       -424.
##  4 Aggregate_Size  clay               8        170
##  5 Aggregate_Size  non-clay          19         66.5
##  6 Aggregate_Size  <NA>             104        -13.4
##  7 Base            base insoluble    25       -112.
##  8 Base            base soluble      27        NaN
##  9 Chem_Extraction ABA residual      62       -345.
## 10 Chem_Extraction acid insoluble    15       -242.
## # ... with 16 more rows
```

```
#Check out number and mean values for different fractions, additionally check out differences based on
frc_data_small %>%
  group_by(frc_scheme, frc_property, pro_parent_material) %>%
  summarise(num = n(),
            mean_frc_14c = mean(frc_14c, na.rm=TRUE))
```

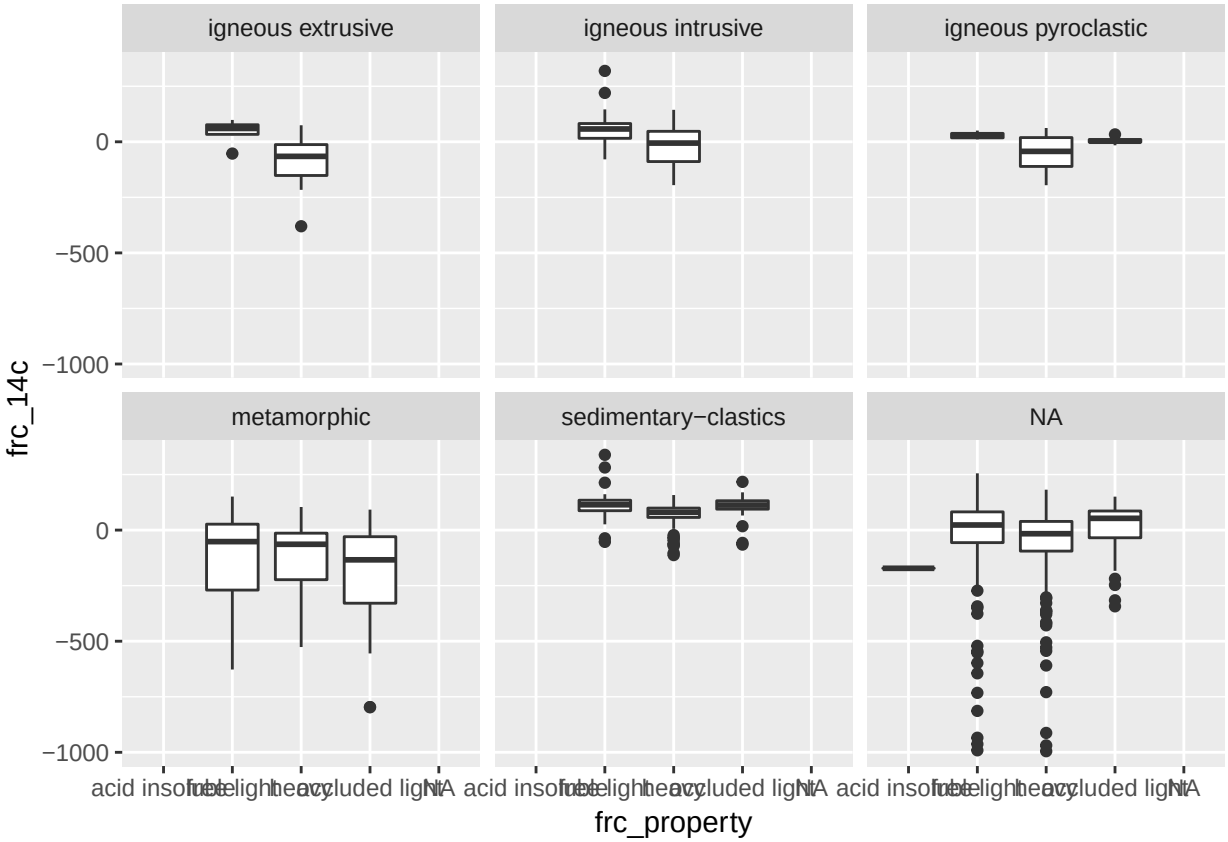
```
## # A tibble: 64 x 5
## # Groups:   frc_scheme, frc_property [?]
##    frc_scheme     frc_property   pro_parent_material    num mean_frc_14c
##    <fct>          <fct>          <fct>                <int>        <dbl>
##  1 Acid           acid insoluble igneous extrusive       10        -1.22
##  2 Acid           acid insoluble <NA>                    44      -182.
##  3 Acid           <NA>           <NA>                    34       184.
##  4 Acid_Evolution carbonate      igneous extrusive        4      -694.
##  5 Acid_Evolution carbonate      igneous intrusive        2       -21.5
##  6 Acid_Evolution carbonate      <NA>                     5      -369.
##  7 Aggregate_Size clay           igneous intrusive        2       170
##  8 Aggregate_Size clay           sedimentary-clastics     6       NaN
##  9 Aggregate_Size non-clay       igneous intrusive        4       112.
## 10 Aggregate_Size non-clay       igneous pyroclastic      3         6.43
## # ... with 54 more rows
```

```
#Check out number of observations for different fractionation schemes & agents
frc_data_small %>%
  group_by(frc_scheme, frc_agent) %>%
  summarise(num = n(),
            mean_frc_14c = mean(frc_14c, na.rm=TRUE))
```

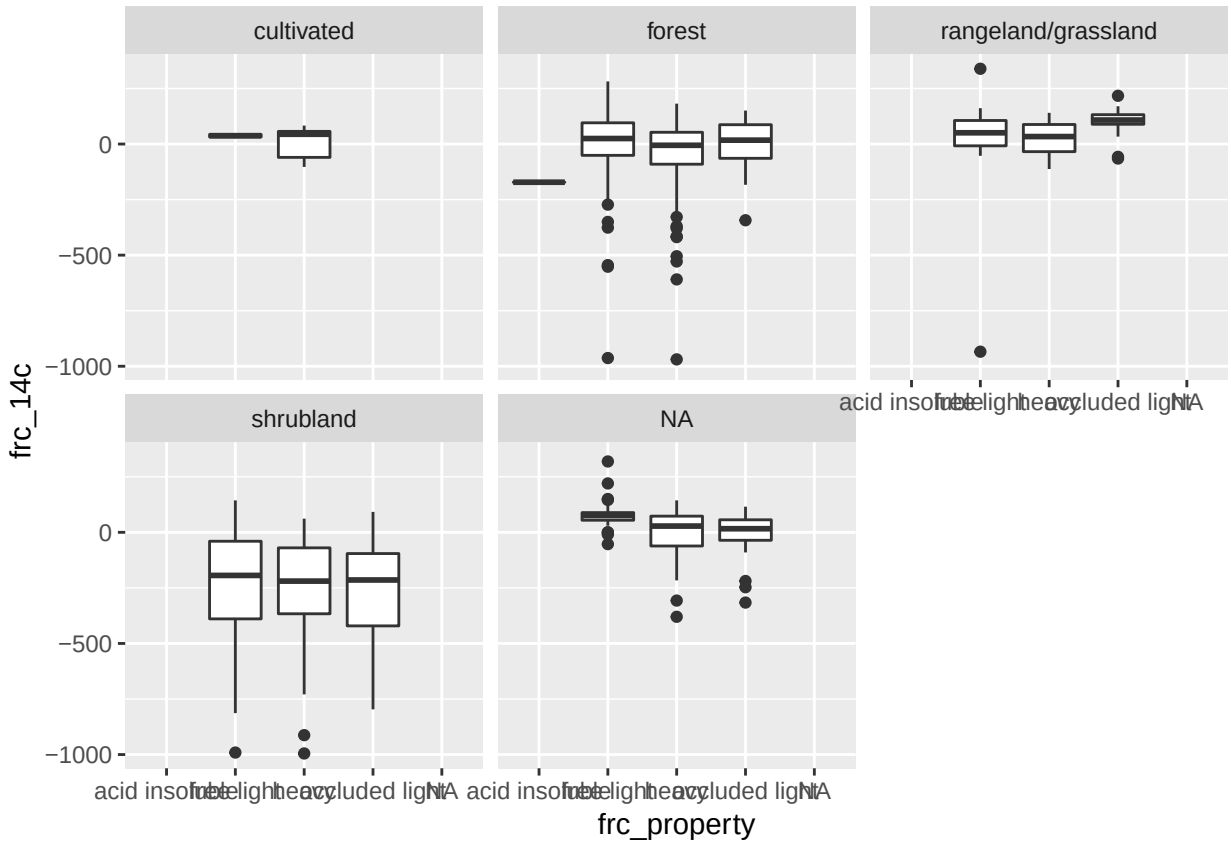
```
## # A tibble: 23 x 4
## # Groups:   frc_scheme [?]
##    frc_scheme     frc_agent             num mean_frc_14c
##    <fct>          <fct>               <int>        <dbl>
##  1 Acid           HCL                    79         10.1
##  2 Acid           NaOH                    9       -343.
##  3 Acid_Evolution HCL                    11       -424.
##  4 Aggregate_Size dry sieve               6        131
##  5 Aggregate_Size wet sieve             125        -10.6
##  6 Base           NaOH                   36        -76.8
##  7 Base           Pyrophosphate          16       -132.
##  8 Chem_Extraction Acid-Base-Acid (ABA)  62       -345.
##  9 Chem_Extraction H2O                    6       -603
## 10 Chem_Extraction HCL                   13       -255.
## # ... with 13 more rows
```

Plotting - Plot density fractions in different ways

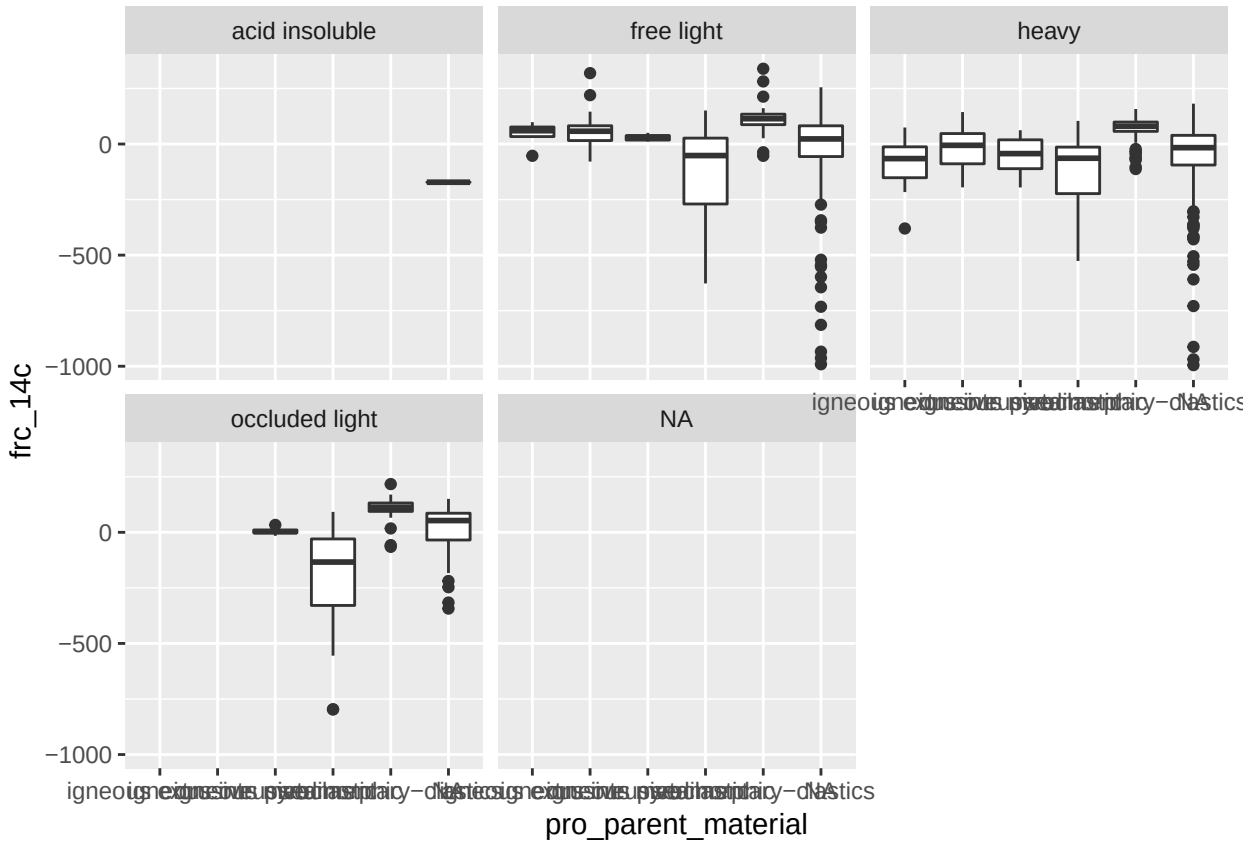
```
frc_data_small %>%
  filter(frc_scheme == "Density") %>%
  ggplot( aes(x=frc_property, y = frc_14c)) +
  geom_boxplot()+
  facet_wrap(~pro_parent_material)
```



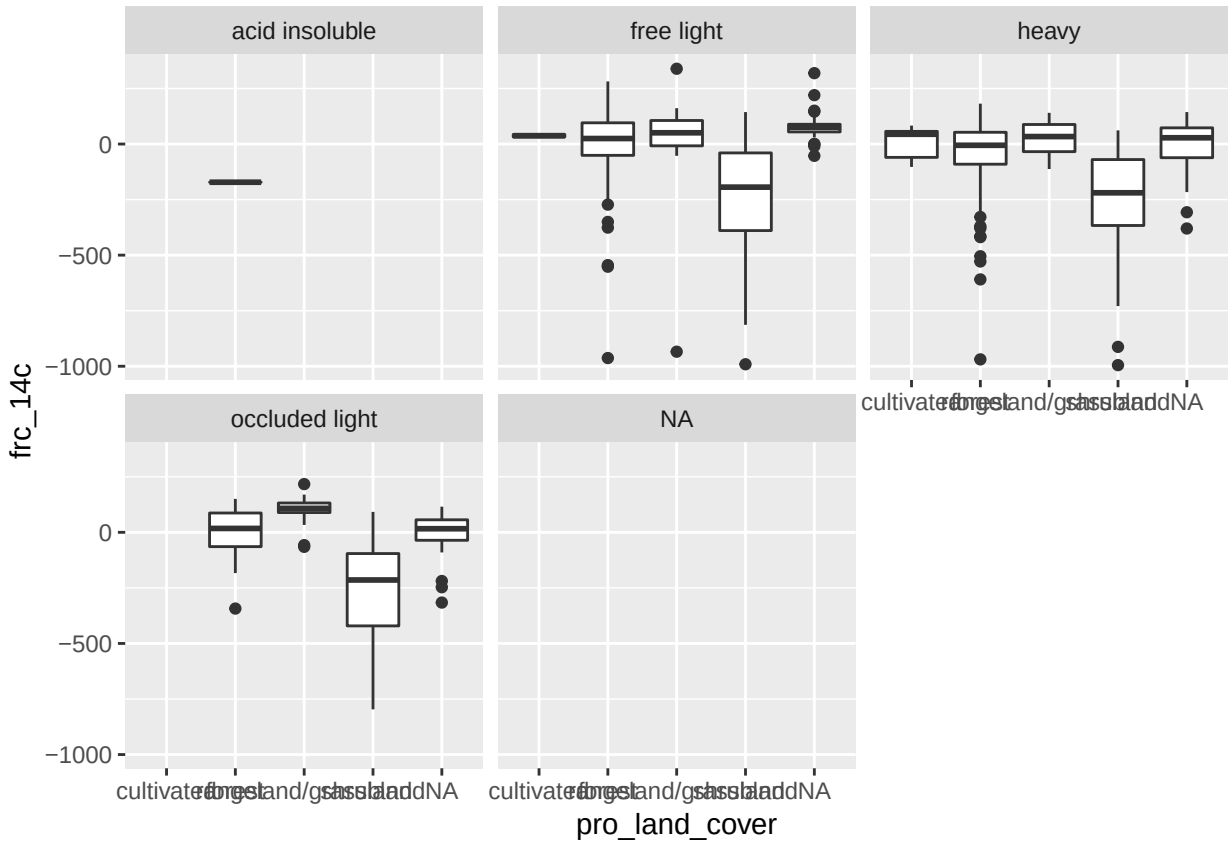

```
frc_data_small %>%
  filter(frc_scheme == "Density") %>%
  ggplot( aes(x=frc_property, y = frc_14c)) +
  geom_boxplot()+
  facet_wrap(~pro_land_cover)
```



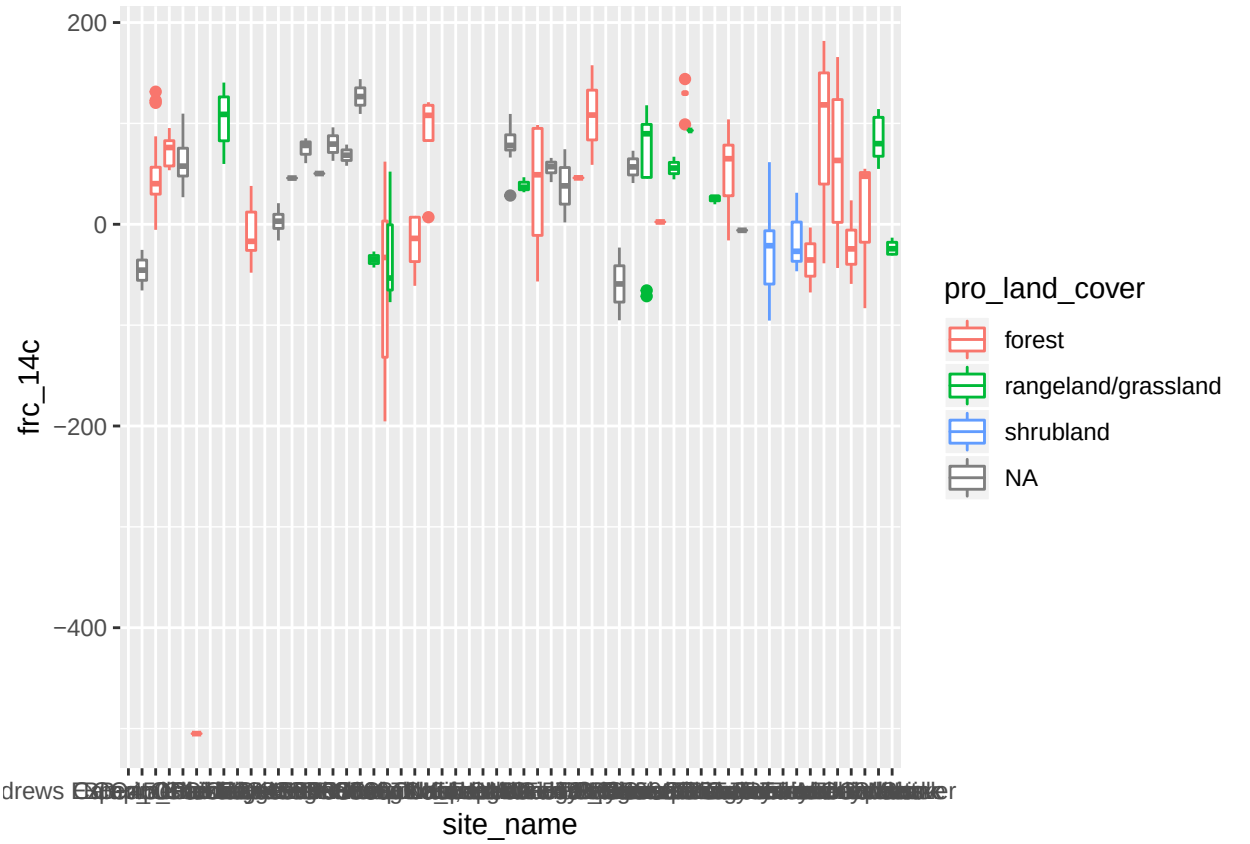

```
frc_data_small %>%
  filter(frc_scheme == "Density") %>%
  ggplot( aes(x=pro_parent_material, y = frc_14c)) +
  geom_boxplot()+
  facet_wrap(~frc_property)
```



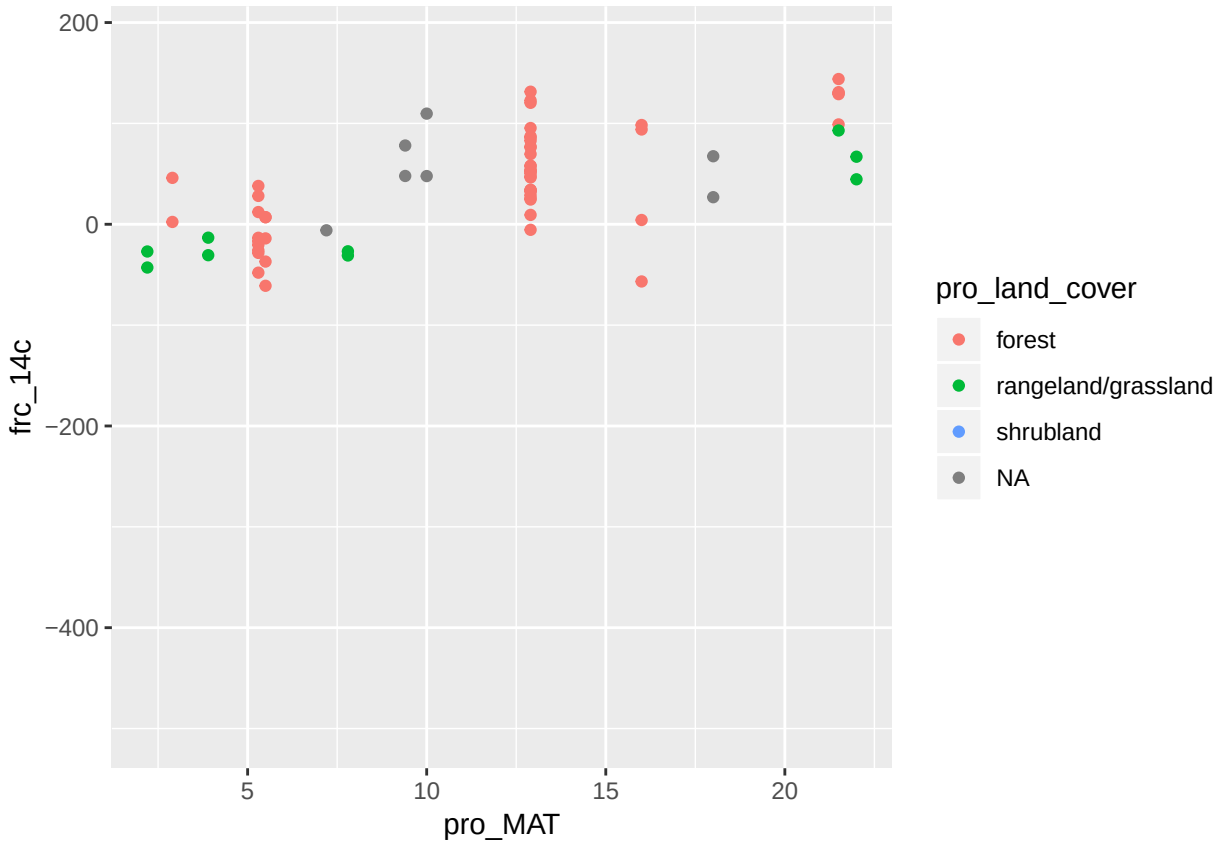

```
frc_data_small %>%
  filter(frc_scheme == "Density") %>%
  ggplot( aes(x=pro_land_cover, y = frc_14c)) +
  geom_boxplot()+
  facet_wrap(~frc_property)
```




```
#How much of a role does site level variability play?
#Let's look only at heavy fraction above 20cm
frc_data_small %>%
  filter(frc_property == "heavy" & lyr_bot < 20) %>%
  ggplot(aes(x= site_name, y = frc_14c, color = pro_land_cover)) +
  geom_boxplot()
```

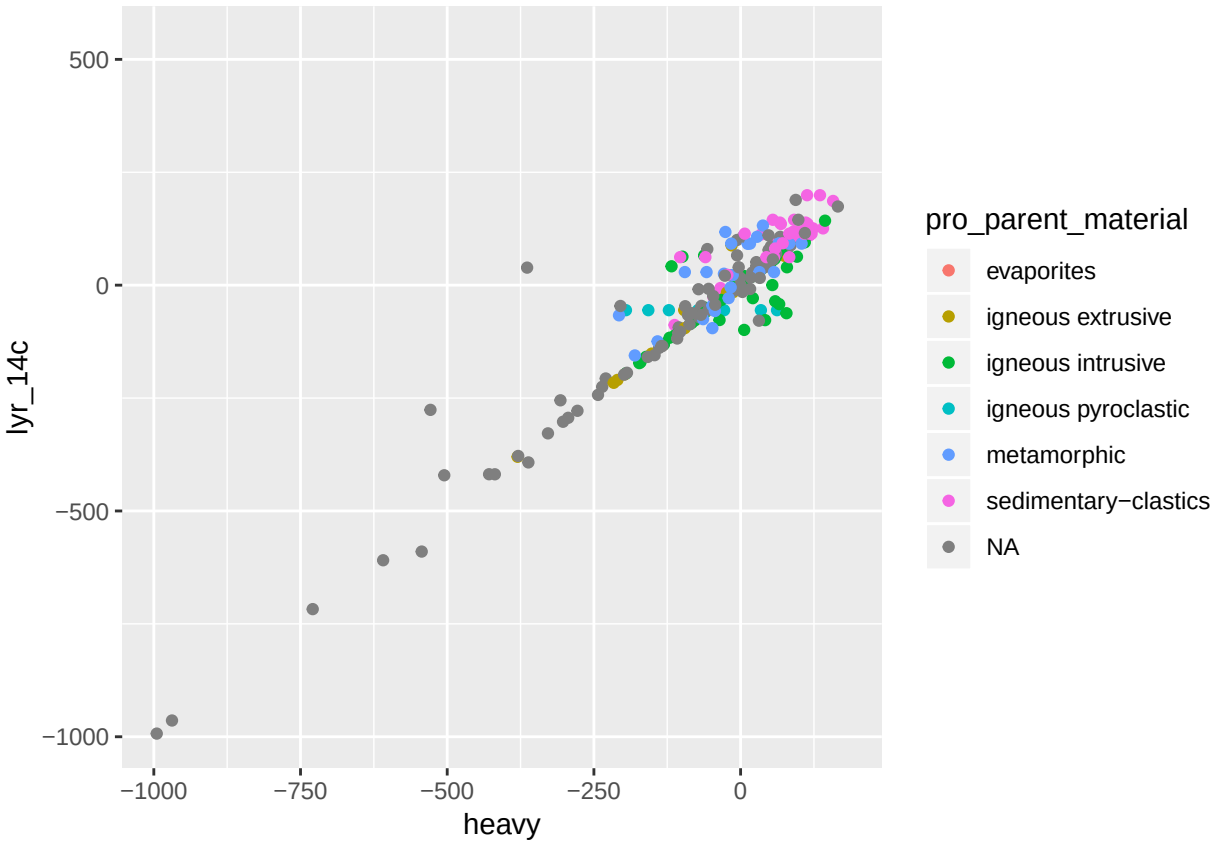



```
#Heavy fraction above 20cm vs temperature
frc_data_small %>%
  filter(frc_property == "heavy" & lyr_bot < 20) %>%
  ggplot(aes(x= pro_MAT, y = frc_14c, color = pro_land_cover)) +
  geom_point()
```



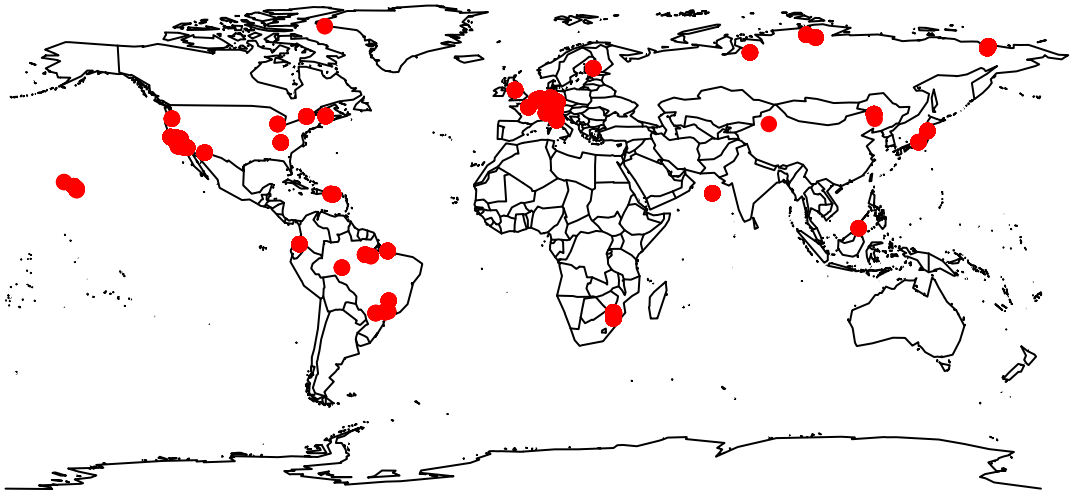
Spread data into multipe columns - Compare fractions from same layer

```
#Create new columns for each of the fractions
frc_data_small %>%
  spread(frc_property, frc_14c, fill = NA) %>% #new columns created on values in frc_property, filled w
  ggplot(aes(x= heavy, y = lyr_14c, color = pro_parent_material)) +
  geom_point()
```



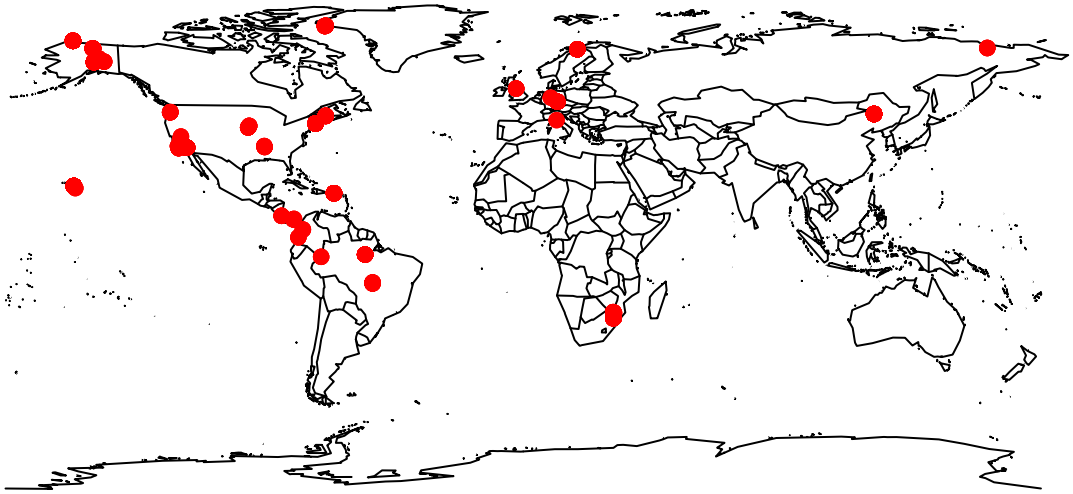
Make some maps of where the data is located

```
# Map sites that have fraction data:
map('world')
points(frc_data$site_long, frc_data$site_lat, col=2, pch=19)
```



```
#Map sites that have incubation data:
inc.site <- ISRaD_data$incubation %>%
  left_join(ISRaD_data$site) %>%
  filter(is.na(inc_14c) != 1) #filters for data with 14C only
```

```r
map('world')
points(inc.site$site_long, inc.site$site_lat, col=2, pch=19)
```



R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this: