

# Supplement A for ‘A 2500 Historical Ecology of Agricultural Production under Empire in Dhiban, Jordan’

*Alan Farahani*

## Contents

Introduction . . . . .	1
Instructions for R Users . . . . .	1
Supplementary Information Regarding Sorting Methodology . . . . .	2
General Summary Statistics of Assemblage . . . . .	2
Summary of Analyzed Contexts . . . . .	4
Identification of Depositional Routes . . . . .	5
All Major Crops Encountered by Period . . . . .	6
Summary Table of Domesticated Taxa . . . . .	7
Rarefaction . . . . .	9
Mean Inverse Simpson’s . . . . .	12
Bootstrap ubiquities . . . . .	14
Weighted Proportions of Abundance and General Counts . . . . .	16
Seed Morphometrics . . . . .	18

## Introduction

This document provides additional information regarding and the full ability to reproduce the analyses conducted in Farahani, A. “A 2,500 Year Historical Ecology of Agricultural Production under Empire in Dhiban, Jordan”. The language in which this is written is R. In the event that the R Markdown file is not available, this PDF document contains embedded code which must be copied into an R console (whether the regular R terminal or the R Studio IDE). Once copied, the code below will reproduce nearly all of the analyses presented in the paper (and some additional ones as well).

If there are any questions, please contact Alan Farahani at [alanfarahani@gmail.com](mailto:alanfarahani@gmail.com)

## Instructions for R Users

Reproducing the data requires the following programs and / or packages to be present. First it is highly recommended that R Studio (or IDE equivalent) is used, with the latest version of **knitr**. Moreover, most of the analyses presented below rely on the **tidyverse** ecosystem, or at least **dplyr** and **reshape2**. The packages **vegan** and **bootstrap** are also critical. See the libraries loaded at the beginning of the document for the full range of required packages that must be installed prior to use.

Many of the commands have `%>% kable(.)` added to them, which enhances the readability of charts in compiled markdown files. The latter should be *removed* in order to use any of these analyses effectively outside of the R Markdown environment.

Finally, ensure that the data are in the same folder as the R Markdown in order for it to successfully compile.

## Supplementary Information Regarding Sorting Methodology

All archaeological plant remains were separated, quantified, and identified down to the 0.5mm level using a stereoscopic microscope. Wood charcoal was counted and weighed at the >2mm mesh size, and only weighed, but not counted, at the 1mm mesh size. Identifiable and unidentifiable seed fragments were collected until the 1mm mesh level. When possible, seed fragments identifiable to some taxonomic level were recorded as fragments of that item (e.g. “Fabaceae fragments”). The only taxa whose fragments were included into the count of whole seeds were Poaceae (grasses) – in cases where a Poaceae seed was fragmented but the embryo / scutellum was visible, that item was recorded as “Poaceae apex”, but included in the count of whole Poaceae. Unidentifiable seed fragments were distinguished from “clinkered” remains, which were classified by severe distortion to the point of an inability of recognition. The quantification of rachis (chaff) remains was predicated on the number of identifiable rachis internodes if found on a spike-chain. Non-botanical, organic remains such as shell, dung, and carbonized insect remains were also counted and identified down to the 0.5mm mesh size, although operationalized into larger categories.

## General Summary Statistics of Assemblage

```
library(tidyverse)
library(reshape2)
library(moments)
library(vegan)
library(bootstrap)
library(knitr)
library(kableExtra)

#note that the data frame is called "dhb.peb" which is used throughout the analyses
#if the CSV has another file name, be sure to either change the CSV name
#or change the path below

dhb.peb.mesh <- read.csv("farahani_raw_data_mesh_v3.csv")

#the raw data includes each of the three sieve sizes used to process the material
#this is removed for the subsequent analyses and all of the counts are aggregated
#by sample number ("LF")

dhb.peb <- dhb.peb.mesh %>%
  select(-Mesh) %>%
  group_by_at(1:13) %>%
  summarise_all(funs(sum)) %>%
  ungroup

#the period designations are reordered so as to follow their chronological sequence
dhb.peb <- dhb.peb %>%
  mutate(Period = as.factor(Period)) %>%
  mutate(Period = factor(Period, levels = levels(Period)[c(1,2,7,4,3,5,6)]))

all.summary <- dhb.peb %>%
  mutate(
    'Crop Seeds' = rowSums(select(., DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS), na.rm=TRUE),
    'Weed Seeds' = rowSums(select(., FAMILIES:UNKS_AND_FRAGS, -Poaceae_Frags), na.rm=TRUE),
    'Seed Fragments (Identifiable)' = rowSums(select(., DOM_FRAGS:FAMILIES, Poaceae_Frags,
```

```

    -contains("wt")), na.rm=TRUE),
  Rachises = rowSums(select(., RACHIS_REMAINS:Wild_or_Weed_rachis, contains("rachis")),
    na.rm=TRUE),
  'Meso and Endocarp Fragments' = rowSums(select(., contains("endo"), -contains("wt")),
    na.rm=TRUE)
) %>%
mutate(
  'Seed Fragments (Unidentifiable)' = Clinker + Unidentifiable_seed_fragment,
  'Unidentified Seeds' = Unknown_fruit_seed + Unknown_non_Poaceae,
  'Cereal Stems and Roots' = Poaceae_culm + Poaceae_root,
  'Charcoal (>2mm)' = Charcoal_ct,
  Parenchyma = Parenchyma_fragment
) %>%
select('Crop Seeds':Parenchyma, Dung, Period) %>%
group_by(Period) %>%
summarise_if(is.numeric, funs(sum)) %>% as.data.frame()

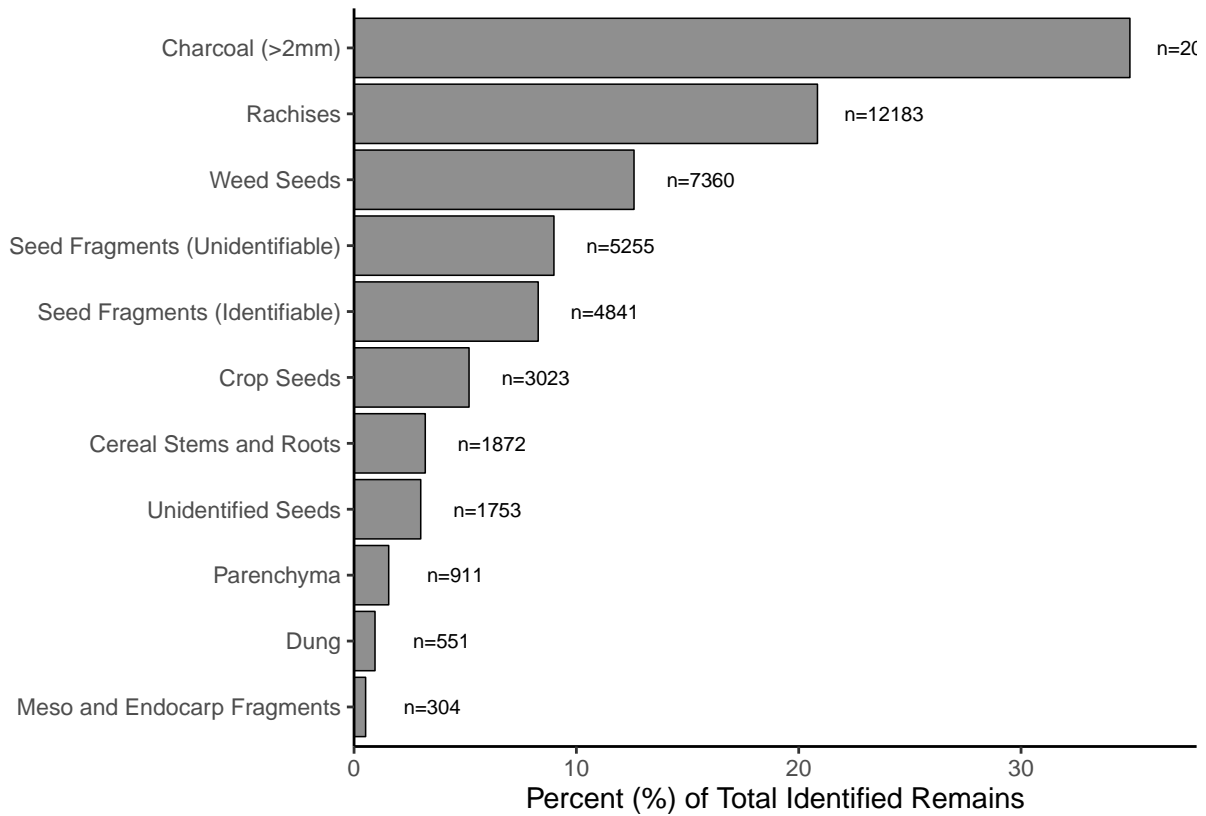
all.summary <- setNames(data.frame(t(all.summary[, -1])), all.summary[, 1])

all.summary %>% kable(.) %>% kable_styling(latex_options = "scale_down")

```

	Iron I	Iron II	Nabataean-Roman	Late Byzantine	Late Antique Transitional	Middle Islamic I	Middle Islamic II
Crop Seeds	2	92	224	787	400	359	1159
Weed Seeds	2	76	280	117	339	2767	3779
Seed Fragments (Identifiable)	0	97	165	1270	491	961	1857
Rachises	0	42	360	73	173	3522	8013
Meso and Endocarp Fragments	0	0	0	289	0	0	15
Seed Fragments (Unidentifiable)	0	109	484	762	1509	682	1709
Unidentified Seeds	0	66	179	60	145	359	944
Cereal Stems and Roots	0	5	34	24	44	557	1208
Charcoal (>2mm)	9	1121	1410	1016	3958	2088	10795
Parenchyma	0	0	0	35	359	106	411
Dung	0	0	0	2	84	44	421

This is a graph of the latter table:



## Summary of Analyzed Contexts

The following presents a table of the contexts that are associated with each of the periods, the amount of deposit sampled in each, and how many there were:

```
#first, by count
dhb.peb %>%
  mutate(N = 1) %>%
  group_by(Period, Context) %>%
  summarise(N = sum(N), Volume = sum(Vol_L)) %>%
  dcast(Period ~ Context, value.var = "N", fill = "-") %>% kable(.)
```

Period	Bin	Drain	Fill	Foundation Trench	Pit	Supra-Surface	Surface	Tabun
Iron I	-	-	-	-	-	-	4	-
Iron II	-	-	14	-	-	-	8	-
Nabataean-Roman	-	-	5	2	-	-	-	-
Late Byzantine	-	6	8	2	-	13	10	-
Late Antique Transitional	-	3	54	-	3	-	-	1
Middle Islamic I	-	-	6	1	-	-	11	-
Middle Islamic II	4	-	12	-	5	-	38	1

```
#now, by volume
dhb.peb %>%
  mutate(N = 1) %>%
  group_by(Period, Context) %>%
  summarise(N = sum(N), Volume = sum(Vol_L)) %>%
  dcast(Period ~ Context, value.var = "Volume", fill = "-") %>% kable(.)
```

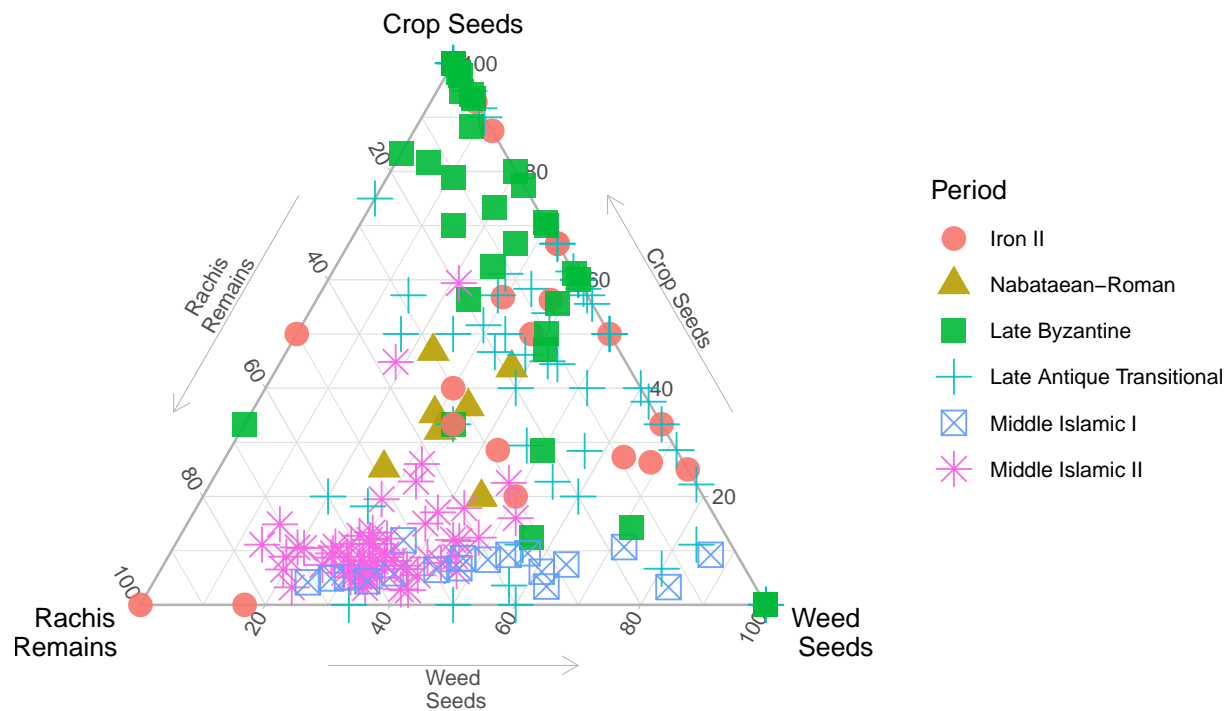
Period	Bin	Drain	Fill	Foundation Trench	Pit	Supra-Surface	Surface	Tabun
Iron I	-	-	-	-	-	-	24	-
Iron II	-	-	142.5	-	-	-	65	-
Nabataean-Roman	-	-	57	49.5	-	-	-	-
Late Byzantine	-	47	49	9	-	91	69.5	-
Late Antique Transitional	-	26.5	439.5	-	29	-	-	3
Middle Islamic I	-	-	182.5	3.5	-	-	167	-
Middle Islamic II	45.5	-	116	-	40.5	-	876	16

## Identification of Depositional Routes

The specific coding for what constitutes a “rachis” in this analysis can be found in the code below.

```
library(ggtern)

dhb.peb %>%
  mutate(Weed_Seeds = rowSums(select(.,FAMILIES:UNKS_AND_FRAGS, -Poaceae_Frags), na.rm = TRUE)) %>%
  mutate(Crop_Seeds = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS), na.rm = TRUE)) %>%
  mutate(Rachises = Wild_or_Weed_rachis + Hordeum_rachis + Triticum_aestivum_durum_rachis +
    Triticum_aestivum_rachis + Triticum_aestivum_durum_rachis_node + Triticum_durum_rachis +
    Triticum_mono_di_rachis + Triticum_node + Triticum_rachis) %>%
  filter(Period!="Ottoman"&Period!="Iron I") %>%
  ggtern(aes(Rachises,Crop_Seeds,Weed_Seeds))+
  geom_mask()+
  geom_point(aes(colour=Period,shape=Period),size=4,alpha=.9)+
  theme_light(base_size = 10)+
  theme_showarrows()+
  ylab("Crop Seeds")+
  xlab("Rachis\n Remains")+
  zlab("Weed\n Seeds")
```



## All Major Crops Encountered by Period

All of the major crops present or absent in each period are displayed below. Several are excluded however, such as “indeterminate cereal” and “large Fabaceae”, both of which are not precise enough in their identification and may “double count” existing taxa. Moreover, identifications of taxa which were not secure, such as *Punica granatum*, *Secale cereale*, and *Triticum spelta*, were excluded.

```
#this codes the Olea fragment as a count of at least one Olea for the Middle Islamic II Period
#this is because there were Olea fragments found in the heavy fraction for this period
#although this count is not reported as such in the overview of the domesticates in the article
#as it was not found in the light fraction
```

```
dhb.peb[dhb.peb$LF == 209, "Olea_sp"] <- 1
```

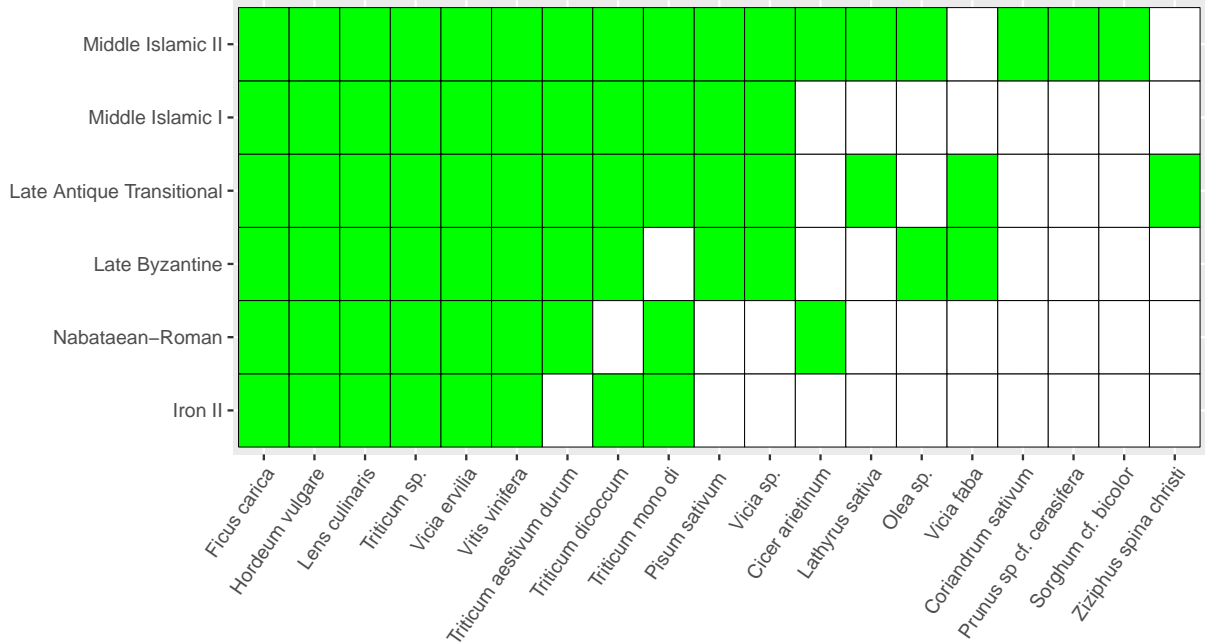
```
excluded_taxa <- c("Indeterminate Cereal", "Punica granatum", "Secale cereale",
                  "Triticum spelta", "Large Fabaceae")
```

```
dhb.peb %>% select(Period, Cicer_arietinum:Ziziphus_spina_christi) %>%
  group_by(Period) %>%
  filter(Period != "Iron I") %>%
  summarise_if(is.numeric, funs(sum)) %>%
  melt %>%
  mutate(value = ifelse(value>0,1,0)) %>%
  mutate( variable = gsub("_"," ",variable),
         variable = gsub("sp$", "sp.", variable),
```

```

    variable = gsub("cf", "cf.", variable),
    variable = gsub("ct", "", variable)
  ) %>%
  filter(!(variable %in% excluded_taxa)) %>%
  ggplot(aes(reorder(variable, desc(value)), Period)) +
  geom_tile(aes(fill = value), colour = "black") +
  scale_fill_continuous(low = "white", high = "green") +
  theme(legend.position = "none", axis.text.x = element_text(angle = 55, hjust = 1)) +
  xlab("") + ylab("")

```



## Summary Table of Domesticated Taxa

The following produces a rather lengthy list of the identified remains aggregated by period. The code relies on a later snippet that is used for the bootstrapping of the domesticate ubiquities.

```

tot.sum <- ddb.peb %>% filter(Period != "Iron I") %>%
  mutate(
    'Crop Seeds' = rowSums(select(., DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS), na.rm=TRUE),
    'Weed Seeds' = rowSums(select(., FAMILIES:UNKS_AND_FRAGS, -Poaceae_Frags), na.rm=TRUE),
    'Seed Fragments (Identifiable)' = rowSums(select(., DOM_FRAGS:FAMILIES, Poaceae_Frags,
      -contains("wt")), na.rm=TRUE),
    Rachises = rowSums(select(., RACHIS_REMAINS:Wild_or_Weed_rachis,
      contains("rachis")), na.rm=TRUE),
    'Meso and Endocarp Fragments' = rowSums(select(., contains("endo"),
      -contains("wt")), na.rm=TRUE)
  ) %>%
  mutate(
    'Seed Fragments (Unidentifiable)' = Clinker + Unidentifiable_seed_fragment,
    'Unidentified Seeds' = Unknown_fruit_seed + Unknown_non_Poaceae,
    'Cereal Stems and Roots' = Poaceae_culm + Poaceae_root,

```

```

  'Large Charcoal' = Charcoal_ct,
  Parenchyma = Parenchyma_fragment,
  N = 1
) %>%
mutate(
  Triticum_aestivum_durum_rachis = Triticum_aestivum_durum_rachis +
    Triticum_aestivum_durum_rachis_node,
  Triticum_indet_rachis = Triticum_rachis + Triticum_node
) %>%
select(Period, Vol_L, N, DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS,
  Triticum_aestivum_durum_rachis,
  Triticum_aestivum_rachis, Triticum_durum_rachis,
  Triticum_mono_di_rachis, Triticum_indet_rachis,
  Hordeum_rachis, 'Crop Seeds':Parenchyma, Dung) %>%
group_by(Period) %>%
summarise_if(is.numeric, funs(sum)) %>%
melt() %>%
filter(!(variable %in% c("Punica_granatum", "Secale_cereale", "Triticum_spelta"))) %>%
dcast(Period ~ ..., fun.aggregate = sum) %>% t() %>% data.frame()

names(tot.sum) <- as.character(unlist(tot.sum[1,]))
tot.sum <- tot.sum[-1,]

tot.sum %>% kable(.) %>% kable_styling(latex_options = "scale_down")

```



	Iron II	Nabataean-Roman	Late Byzantine	Late Antique Transitional	Middle Islamic I	Middle Islamic II
Vol_L	207.5	106.5	265.5	498.0	353.0	1094.0
N	22	7	39	61	18	60
DOMESTICATE_COUNTS	0	0	0	0	0	0
Cicer_arietinum	0	2	0	0	0	1
Coriandrum_sativum	0	0	0	0	0	1
Ficus_carica	43	124	104	118	123	418
Hordeum_vulgare	22	27	13	58	121	286
Indeterminate_Cereal	0	1	12	14	9	31
Large_Fabaceae	0	2	2	13	7	28
Lathyrus_sativa	0	0	0	1	0	1
Lens_culinaris	4	8	10	15	5	24
Olea_sp	0	0	1	0	0	1
Pisum_sativum_ct	0	0	132	7	1	11
Prunus_sp_cf_cerasifera	0	0	0	0	0	1
Sorghum_cf_bicolor	0	0	0	0	0	1
Triticum_aestivum_durum	0	15	31	27	37	92
Triticum_dicoccum	1	0	18	8	1	12
Triticum_mono_di	1	1	0	4	1	5
Triticum_sp	12	26	294	40	39	152
Vicia_ervilia	2	4	4	12	13	51
Vicia_faba	0	0	4	1	0	0
Vicia_sp	0	0	3	1	1	1
Vitis_vinifera	7	14	159	80	1	32
Ziziphus_spina_christi	0	0	0	1	0	0
DOMESTICATE_WEIGHTS	0	0	0	0	0	0
Triticum_aestivum_durum_rachis	5	37	31	74	1632	3951
Triticum_aestivum_rachis	3	4	1	1	43	178
Triticum_durum_rachis	0	0	2	1	12	33
Triticum_mono_di_rachis	4	3	0	0	4	4
Triticum_indet_rachis	8	203	1	0	72	49
Hordeum_rachis	9	35	8	18	706	1099
Crop Seeds	92	224	787	400	359	1160
Weed Seeds	76	280	117	339	2767	3779
Seed Fragments (Identifiable)	97	165	1270	491	961	1857
Rachises	42	360	73	173	3522	8013
Meso and Endocarp Fragments	0	0	289	0	0	15
Seed Fragments (Unidentifiable)	109	484	762	1509	682	1709
Unidentified Seeds	66	179	60	145	359	944
Cereal Stems and Roots	5	34	24	44	557	1208
Large Charcoal	1121	1410	1016	3958	2088	10795
Parenchyma	0	0	35	359	106	411
Dung	0	0	2	84	44	421

## Rarefaction

The logic of rarefaction is presented in the paper and discussed extensively in several different sources, but here the **vegan** package is used for the calculation of individual-based and sample-based rarefaction.

```
periods <- c("Middle Islamic II", "Iron II", "Late Byzantine", "Late Antique Transitional")

#this codes the Olea fragment as a count of at least one Olea for the Middle Islamic Period
dhb.peb[dhb.peb$LF == 209, "Olea_sp"] <- 1

#create colors for our different strata
redtrans <- rgb(255, 0, 0, 127, maxColorValue=255)
chartrans <- rgb(102, 205, 0, 200, maxColorValue=255)
lightbluetrans <- rgb(173, 216, 230, 200, maxColorValue=255)
purpletrans <- rgb(200, 100, 200, 200, maxColorValue = 255)

#The following code loops through each of the time periods
#and creates a different species accumulation curve based on the domesticates only
#(with some exluded, as can be seen in the select below) through random permutations
#of the data weighted by flotation volume

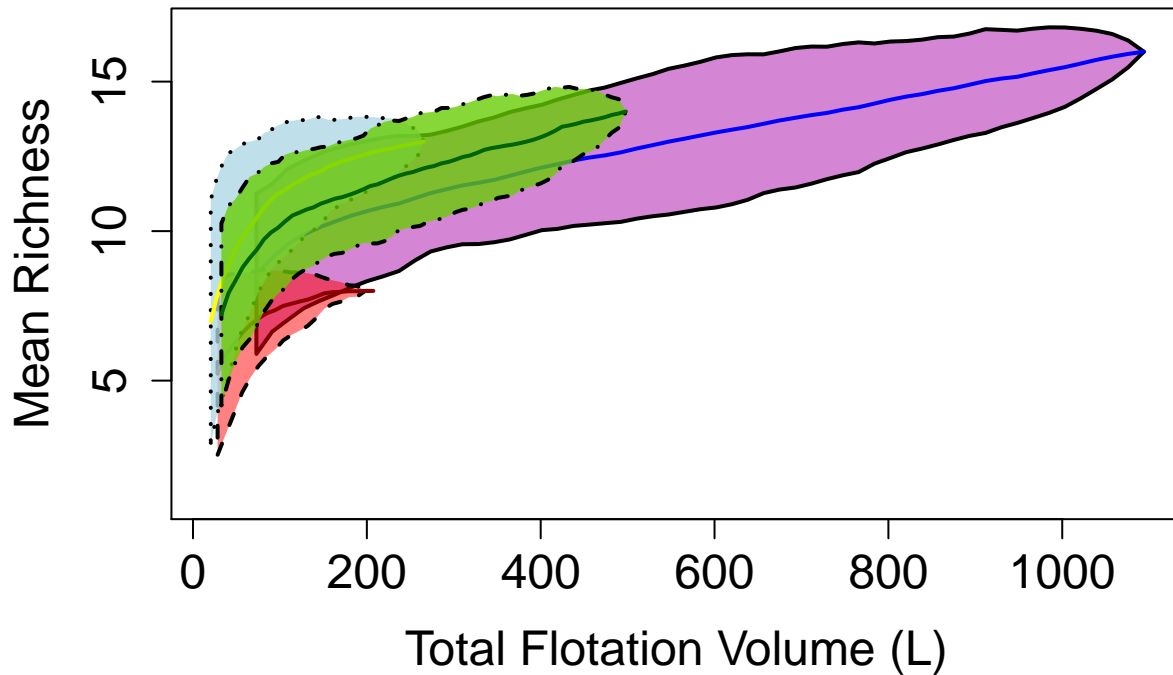
for (i in 1:length(periods)){
```

```

x <- dhb.peb %>%
  select(Period, Cicer_arietinum:Ziziphus_spina_christi,Vol_L,
    -c(Secale_cereale, Punica_granatum, Indeterminate_Cereal,
      Coriandrum_sativum, Large_Fabaceae, Triticum_spelta)) %>%
  mutate(Triticum_mono_di = Triticum_mono_di + Triticum_dicoccum) %>%
  select(-Triticum_dicoccum) %>%
  filter(Period == periods[i]) %>%
  select(-Period)

if(i == 1){
  plot(specaccum(x, method = "random",w=x$Vol_L),xvar="effort",ci.type="poly",col="blue",
    lwd=2, ci.lty=i, ci.col=purpletrans,
    xlab = "Total Flotation Volume (L)", ylab = "Mean Richness",
    cex.axis = 1.5, cex.lab = 1.5)
}
else if(i == 2){
  plot(specaccum(x, method = "random",w=x$Vol_L),xvar="effort",ci.type="poly", col="darkred",
    lwd=2, ci.lty=i, ci.col=redtrans, add=TRUE)
}
else if(i == 3){
  plot(specaccum(x, method = "random",w=x$Vol_L),xvar="effort",ci.type="poly", col="yellow",
    lwd=2, ci.lty=i, ci.col=lightbluetrans, add=TRUE)
}
else{
  plot(specaccum(x, method = "random",w=x$Vol_L),xvar="effort",ci.type="poly", col="darkgreen",
    lwd=2, ci.lty=i, ci.col=chartrans, add=TRUE)
}
}

```



```
#classic rarefaction of the data, by individuals
```

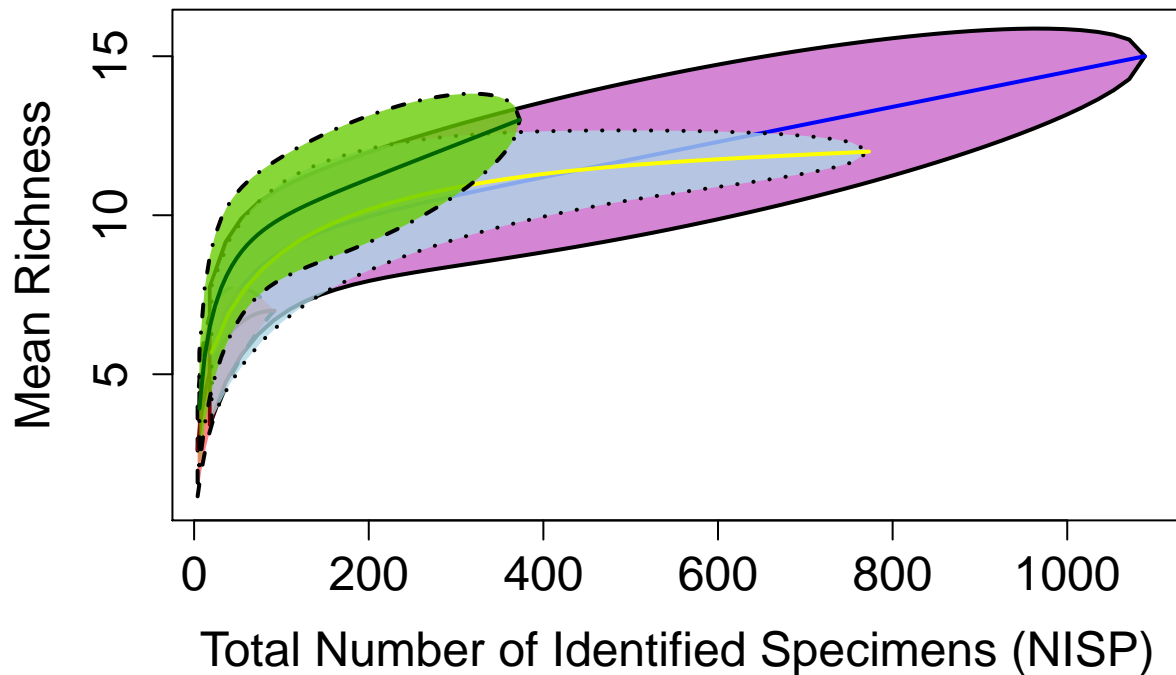
```
for (i in 1:length(periods)){
  x <- dnb.peb %>%
    select(Period, Cicer_arietinum:Ziziphus_spina_christi,
           -c(Secale_cereale, Punica_granatum, Indeterminate_Cereal,
              Coriandrum_sativum, Large_Fabaceae, Triticum_spelta)) %>%
    mutate(Triticum_mono_di = Triticum_mono_di + Triticum_dicoccum) %>%
    select(-Triticum_dicoccum) %>%
    filter(Period == periods[i]) %>%
    select(-Period)

  if(i == 1){
    plot(specaccum(x, method = "rarefaction"), ci.type="poly", col="blue", lwd=2,
         ci.lty=i, ci.col=purpletrans, xvar = "individuals",
         xlab = "Total Number of Identified Specimens (NISF)", ylab = "Mean Richness",
         cex.axis = 1.5, cex.lab = 1.5)
  }
  else if(i == 2){
    plot(specaccum(x, method = "rarefaction"), xvar = "individuals", ci.type="poly",
         col="darkred", lwd=2, ci.lty=i, ci.col=redtrans, add=TRUE)
  }
  else if(i == 3){
    plot(specaccum(x, method = "rarefaction"), xvar = "individuals", ci.type="poly",
         col="yellow", lwd=2, ci.lty=i, ci.col=lightbluetrans, add=TRUE)
  }
}
```

```

else{
  plot(specaccum(x, method = "rarefaction"), xvar = "individuals", ci.type="poly",
       col="darkgreen", lwd=2, ci.lty=i, ci.col=chartrans, add=TRUE)
}
}

```



## Mean Inverse Simpson's

In the event that the raw Estimate S data are not available, the following code can be run in order to procure the necessary data that must be modified for the Estimate S input system. Unfortunately this can not be exhaustively discussed in this space, however the following code should generate the data needed for the software.

```

for (i in 1:length(levels(dhb.peb$Period))){
  x <- dhb.peb %>%
    select(Period, Cicer_arietinum:Ziziphus_spina_christi,
          -c(Secale_cereale, Punica_granatum, Indeterminate_Cereal,
             Coriandrum_sativum, Large_Fabaceae, Triticum_spelta)) %>%
    mutate(Triticum_mono_di = Triticum_mono_di + Triticum_monococcum + Triticum_dicoccum) %>%
    select(-c(Triticum_monococcum, Triticum_dicoccum)) %>%
    filter(Period == periods[i]) %>%
    select(-Period)
  name <- paste(periods[i])
  q <- t(x)
}

```

```

q <- q[-1,]
write.csv(q, file = paste0("C:/Users/Username/Desktop/EstimateSdata/",name,".csv"))

rm(q, x, name)
}

```

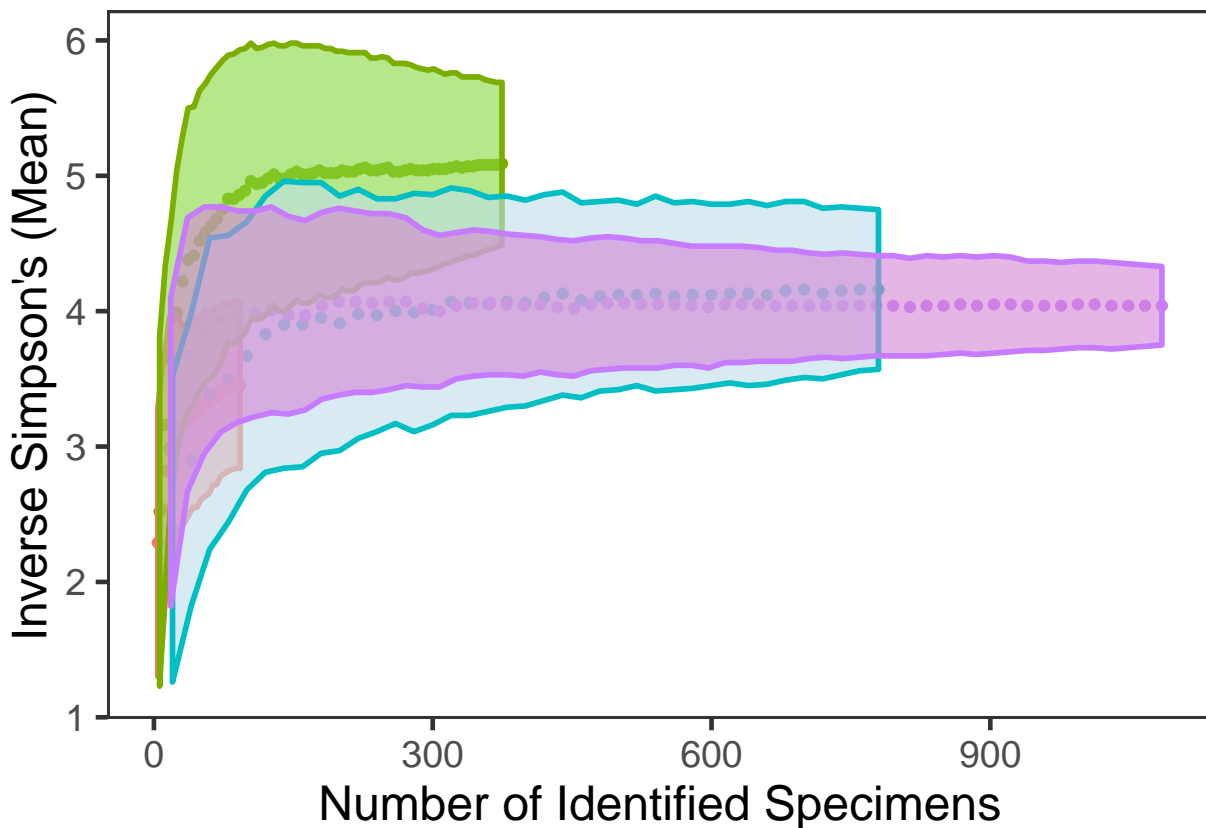
Once the EstimateS output has been calculated, a modified data frame can then be passed and plotted via regular `ggplot` methods. Note here that the confidence interval of the inverse Simpson's index. is in fact predicated upon the standard deviation of the estimate.

```

estimate <- read.csv("estimate_s_output.csv")

estimate %>% ggplot(aes(Individuals..computed.,Simpson.Inv.Mean)) +
  geom_point(aes(colour = Period)) +
  geom_ribbon(aes(ymin = Simpson.Inv.Mean - Simpson.Inv.SD..runs.,
                ymax = Simpson.Inv.Mean + Simpson.Inv.SD..runs.,
                colour = Period, fill = Period), size = 1,
            alpha = .6) +
  theme_bw(base_size = 18) +
  theme(panel.grid = element_blank(), legend.position = "none") +
  xlab("Number of Identified Specimens") +
  ylab("Inverse Simpson's (Mean)") +
  scale_fill_manual(values=c("#ff8081", "#86d73c", "#bfdfea", "#d485d4"))

```



## Bootstrap ubiquities

The use of a bootstrap for the ubiquity (proportional presence) required several steps. The steps are outlined below in the code itself.

```
#first, isolate and aggregate the taxa of interest by their genus name
dom <- dhb.peb %>%
  mutate(Triticum = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS) %>%
    select(.,contains("Triticum"))),
    Hordeum = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS) %>%
    select(.,contains("Hordeum"))),
    Ficus = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS) %>%
    select(.,contains("Ficus"))),
    Lens = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS) %>%
    select(.,contains("Lens"))),
    Pisum = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS) %>%
    select(.,contains("Pisum"))),
    Vicia = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS) %>%
    select(.,contains("Vicia"))),
    Vitis = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS) %>%
    select(.,contains("Vitis"))),
    All_Doms = rowSums(select(.,DOMESTICATE_COUNTS:DOMESTICATE_WEIGHTS))) %>%
  select(LF, Period, Vol_L, Triticum, Hordeum, Ficus, Lens, Pisum, Vicia, Vitis, All_Doms)

#using these taxa, recode so that they are present or absent
#and save this as a new data frame minus two periods

temp <- dom %>%
  #the following line can be included as necessary
  #mutate(Rem_Doms = All_Doms - rowSums(select(.,Triticum:Vitis))) %>%
  select(-c(LF, Vol_L, All_Doms)) %>%
  filter(Period != "Iron I" & Period != "Ottoman") %>%
  mutate_if(is.numeric, funs(ifelse((.)>0,1,0))) %>%
  droplevels(.) %>% as.data.frame(.)

#now loop through the data frame, extracting each period
#bootstrapping the taxa in each period, saving those in a list
#and then combining the whole list

test <- vector("list", 6)

for (i in 1:length(levels(temp$Period))){
  per <- levels(temp$Period)[i]
  n <- filter(temp, Period == per)
  for (z in 1:length(temp[,-1])){
    p <- n
    p <- p[,-1]
    boo <- bootstrap(p[,z], 1000, function(x) mean(x))$thetastar
    test[[i]][[z]] <- list(mean = mean(p[,z]),
      mean_boot = mean(boo),
      high = quantile(boo, .95),
      low = quantile(boo, .05),
      Period = per,
      Taxon = names(temp[z+1]),
```

```

      TimeOp = i,
      N = nrow(temp[temp$Period == per,]))

    }
  }

#combine the list with bind_rows and map_df from purrr
test <- test %>%
  map_df(bind_rows)
  #pipe this in if necessary with %>%
  #mutate_at(vars(mean:low), function(x) round(x*100, 1))

test <- test %>% mutate(Period_N = as.factor(paste(Period, " (N = ", N, ")", sep = "")),
  Period = as.factor(Period)) %>% select(-N)
test$Period_N <- factor(test$Period_N, levels = levels(test$Period_N)[c(1,6,3,2,4,5)])
test$Period <- factor(test$Period, levels = levels(test$Period)[c(1,6,3,2,4,5)])

rm(n,p,boo,i,per,z, temp)

test %>% kable(.) %>% kable_styling(latex_options = "scale_down")

```

mean	mean_boot	high	low	Period	Taxon	TimeOp	Period_N
0.4090909	0.4125000	0.5909091	0.2272727	Iron II	Triticum	1	Iron II (N = 22)
0.4090909	0.4098182	0.5909091	0.2272727	Iron II	Hordeum	1	Iron II (N = 22)
0.4545455	0.4499545	0.5909091	0.2727273	Iron II	Ficus	1	Iron II (N = 22)
0.1818182	0.1793636	0.3181818	0.0454545	Iron II	Lens	1	Iron II (N = 22)
0.0000000	0.0000000	0.0000000	0.0000000	Iron II	Pisum	1	Iron II (N = 22)
0.0909091	0.0916364	0.1818182	0.0000000	Iron II	Vicia	1	Iron II (N = 22)
0.2727273	0.2760000	0.4545455	0.1363636	Iron II	Vitis	1	Iron II (N = 22)
0.7142857	0.7051429	1.0000000	0.4285714	Nabataean-Roman	Triticum	2	Nabataean-Roman (N = 7)
0.8571429	0.8567143	1.0000000	0.5714286	Nabataean-Roman	Hordeum	2	Nabataean-Roman (N = 7)
1.0000000	1.0000000	1.0000000	1.0000000	Nabataean-Roman	Ficus	2	Nabataean-Roman (N = 7)
0.4285714	0.4247143	0.7142857	0.1428571	Nabataean-Roman	Lens	2	Nabataean-Roman (N = 7)
0.0000000	0.0000000	0.0000000	0.0000000	Nabataean-Roman	Pisum	2	Nabataean-Roman (N = 7)
0.5714286	0.5745714	0.8571429	0.2857143	Nabataean-Roman	Vicia	2	Nabataean-Roman (N = 7)
0.7142857	0.7128571	1.0000000	0.4285714	Nabataean-Roman	Vitis	2	Nabataean-Roman (N = 7)
0.6666667	0.6665385	0.7948718	0.5384615	Late Byzantine	Triticum	3	Late Byzantine (N = 39)
0.2564103	0.2570769	0.3589744	0.1538462	Late Byzantine	Hordeum	3	Late Byzantine (N = 39)
0.7179487	0.7180513	0.8461538	0.5897436	Late Byzantine	Ficus	3	Late Byzantine (N = 39)
0.1538462	0.1508205	0.2564103	0.0512821	Late Byzantine	Lens	3	Late Byzantine (N = 39)
0.2307692	0.2265128	0.3333333	0.1282051	Late Byzantine	Pisum	3	Late Byzantine (N = 39)
0.1794872	0.1778718	0.2820513	0.0769231	Late Byzantine	Vicia	3	Late Byzantine (N = 39)
0.5897436	0.5872308	0.7179487	0.4615385	Late Byzantine	Vitis	3	Late Byzantine (N = 39)
0.5081967	0.5076230	0.6065574	0.4098361	Late Antique Transitional	Triticum	4	Late Antique Transitional (N = 61)
0.4754098	0.4772787	0.5737705	0.3770492	Late Antique Transitional	Hordeum	4	Late Antique Transitional (N = 61)
0.5409836	0.5397377	0.6393443	0.4426230	Late Antique Transitional	Ficus	4	Late Antique Transitional (N = 61)
0.1967213	0.1959344	0.2786885	0.1147541	Late Antique Transitional	Lens	4	Late Antique Transitional (N = 61)
0.0491803	0.0492459	0.0983607	0.0163934	Late Antique Transitional	Pisum	4	Late Antique Transitional (N = 61)
0.1147541	0.1132787	0.1803279	0.0491803	Late Antique Transitional	Vicia	4	Late Antique Transitional (N = 61)
0.4098361	0.4072951	0.5081967	0.3114754	Late Antique Transitional	Vitis	4	Late Antique Transitional (N = 61)
0.9444444	0.9457778	1.0000000	0.8333333	Middle Islamic I	Triticum	5	Middle Islamic I (N = 18)
1.0000000	1.0000000	1.0000000	1.0000000	Middle Islamic I	Hordeum	5	Middle Islamic I (N = 18)
0.8888889	0.8903333	1.0000000	0.7777778	Middle Islamic I	Ficus	5	Middle Islamic I (N = 18)
0.1666667	0.1653889	0.3333333	0.0555556	Middle Islamic I	Lens	5	Middle Islamic I (N = 18)
0.0555556	0.0543889	0.1666667	0.0000000	Middle Islamic I	Pisum	5	Middle Islamic I (N = 18)
0.5555556	0.5526111	0.7222222	0.3888889	Middle Islamic I	Vicia	5	Middle Islamic I (N = 18)
0.0555556	0.0547778	0.1666667	0.0000000	Middle Islamic I	Vitis	5	Middle Islamic I (N = 18)
0.8166667	0.8154833	0.9000000	0.7333333	Middle Islamic II	Triticum	6	Middle Islamic II (N = 60)
0.9500000	0.9521167	1.0000000	0.9000000	Middle Islamic II	Hordeum	6	Middle Islamic II (N = 60)
0.7833333	0.7824833	0.8666667	0.7000000	Middle Islamic II	Ficus	6	Middle Islamic II (N = 60)
0.2666667	0.2681333	0.3666667	0.1833333	Middle Islamic II	Lens	6	Middle Islamic II (N = 60)
0.1166667	0.1149667	0.1833333	0.0500000	Middle Islamic II	Pisum	6	Middle Islamic II (N = 60)
0.4333333	0.4342333	0.5333333	0.3333333	Middle Islamic II	Vicia	6	Middle Islamic II (N = 60)
0.2500000	0.2530167	0.3500000	0.1666667	Middle Islamic II	Vitis	6	Middle Islamic II (N = 60)

## Weighted Proportions of Abundance and General Counts

The calculations for the weighted proportions can be seen in the code below, as well as that used to generate the graph.

```
dom.summary <- dom %>%
  #mutate(Rem_Doms = All_Doms - rowSums(select(.,Triticum:Vitis))) %>%
  filter(Period != "Iron I" & Period != "Ottoman") %>%
  droplevels(.) %>%
  add_count(Period) %>%
  melt(id.vars = c("Period","LF", "All_Doms", "n", "Vol_L")) %>%
  mutate(Ubiq = ifelse(value > 0, 1, 0)) %>%
  group_by(Period, LF) %>%
  mutate(Prop = ifelse(value == 0, 0, value/sum(value)), Prop_Ubiq = Ubiq / n) %>%
  ungroup() %>%
  rename(Taxon =variable) %>%
  group_by(Period, Taxon) %>%
  summarise(N = unique(n),
            Vol = sum(Vol_L),
```



```

CT = sum(value),
Ubiq = sum(Ubiq),
Ubiq_prop = Ubiq / N,
WT_Prop = weighted.mean(Prop, All_Doms),
#WT_Prop = weighted.mean(Prop, Taxon), #weighted by count of seeds per taxon
Mean_Prop = mean(Prop, na.rm = TRUE),
Mean_Dens = mean(value / Vol),
WT_Dens = weighted.mean((value/Vol), All_Doms)) %>%
mutate(CT_Prop = CT / sum(CT)) %>%
mutate_at(vars(Ubiq_prop:Mean_Prop,CT_Prop), funs(round(., 2))) %>% as.data.frame()

dom.summary %>% kable(.) %>% kable_styling(latex_options = "scale_down")

```

Period	Taxon	N	Vol	CT	Ubiq	Ubiq_prop	WT_Prop	Mean_Prop	Mean_Dens	WT_Dens	CT_Prop
Iron II	Triticum	22	207.5	14	9	0.41	0.15	0.16	0.0030668	0.0071765	0.15
Iron II	Hordeum	22	207.5	22	9	0.41	0.24	0.20	0.0048193	0.0148769	0.24
Iron II	Ficus	22	207.5	43	10	0.45	0.47	0.30	0.0094195	0.0287061	0.47
Iron II	Lens	22	207.5	4	4	0.18	0.04	0.07	0.0008762	0.0011001	0.04
Iron II	Pisum	22	207.5	0	0	0.00	0.00	0.00	0.0000000	0.0000000	0.00
Iron II	Vicia	22	207.5	2	2	0.09	0.02	0.01	0.0004381	0.0015191	0.02
Iron II	Vitis	22	207.5	7	6	0.27	0.08	0.08	0.0015334	0.0018334	0.08
Nabataean-Roman	Triticum	7	106.5	42	5	0.71	0.19	0.14	0.0563380	0.0995976	0.19
Nabataean-Roman	Hordeum	7	106.5	27	6	0.86	0.13	0.16	0.0362173	0.0551224	0.12
Nabataean-Roman	Ficus	7	106.5	124	7	1.00	0.57	0.56	0.1663313	0.2671445	0.57
Nabataean-Roman	Lens	7	106.5	8	3	0.43	0.04	0.05	0.0107311	0.0209172	0.04
Nabataean-Roman	Pisum	7	106.5	0	0	0.00	0.00	0.00	0.0000000	0.0000000	0.00
Nabataean-Roman	Vicia	7	106.5	4	4	0.57	0.02	0.04	0.0053655	0.0071680	0.02
Nabataean-Roman	Vitis	7	106.5	14	5	0.71	0.06	0.06	0.0187793	0.0249832	0.06
Late Byzantine	Triticum	39	265.5	343	26	0.67	0.44	0.27	0.0331257	0.1841650	0.44
Late Byzantine	Hordeum	39	265.5	13	10	0.26	0.02	0.05	0.0012555	0.0016463	0.02
Late Byzantine	Ficus	39	265.5	104	28	0.72	0.14	0.28	0.0100439	0.0159992	0.13
Late Byzantine	Lens	39	265.5	10	6	0.15	0.01	0.02	0.0009658	0.0014836	0.01
Late Byzantine	Pisum	39	265.5	132	9	0.23	0.17	0.07	0.0127481	0.0832167	0.17
Late Byzantine	Vicia	39	265.5	11	7	0.18	0.01	0.01	0.0010623	0.0033453	0.01
Late Byzantine	Vitis	39	265.5	159	23	0.59	0.21	0.22	0.0153556	0.0367220	0.21
Late Antique Transitional	Triticum	61	498.0	79	31	0.51	0.21	0.23	0.0026006	0.0072992	0.21
Late Antique Transitional	Hordeum	61	498.0	58	29	0.48	0.16	0.16	0.0019093	0.0056476	0.16
Late Antique Transitional	Ficus	61	498.0	118	33	0.54	0.31	0.24	0.0038844	0.0146837	0.32
Late Antique Transitional	Lens	61	498.0	15	12	0.20	0.04	0.06	0.0004938	0.0013956	0.04
Late Antique Transitional	Pisum	61	498.0	7	3	0.05	0.02	0.01	0.0002304	0.0014408	0.02
Late Antique Transitional	Vicia	61	498.0	14	7	0.11	0.04	0.02	0.0004609	0.0016315	0.04
Late Antique Transitional	Vitis	61	498.0	80	25	0.41	0.22	0.16	0.0026335	0.0070884	0.22
Middle Islamic I	Triticum	18	353.0	78	17	0.94	0.23	0.25	0.0122757	0.0177468	0.23
Middle Islamic I	Hordeum	18	353.0	121	18	1.00	0.35	0.40	0.0190431	0.0297016	0.35
Middle Islamic I	Ficus	18	353.0	123	16	0.89	0.36	0.29	0.0193579	0.0445919	0.36
Middle Islamic I	Lens	18	353.0	5	3	0.07	0.01	0.01	0.0007869	0.0019727	0.01
Middle Islamic I	Pisum	18	353.0	1	1	0.06	0.00	0.01	0.0001574	0.0000710	0.00
Middle Islamic I	Vicia	18	353.0	14	10	0.56	0.04	0.04	0.0022033	0.0027145	0.04
Middle Islamic I	Vitis	18	353.0	1	1	0.06	0.00	0.00	0.0001574	0.0002131	0.00
Middle Islamic II	Triticum	60	1094.0	266	49	0.82	0.25	0.27	0.0040524	0.0065538	0.24
Middle Islamic II	Hordeum	60	1094.0	286	57	0.95	0.26	0.32	0.0043571	0.0090951	0.26
Middle Islamic II	Ficus	60	1094.0	418	47	0.78	0.38	0.29	0.0063681	0.0188757	0.38
Middle Islamic II	Lens	60	1094.0	24	16	0.27	0.02	0.02	0.0003656	0.0008983	0.02
Middle Islamic II	Pisum	60	1094.0	11	7	0.12	0.01	0.02	0.0001676	0.0001962	0.01
Middle Islamic II	Vicia	60	1094.0	52	26	0.43	0.05	0.05	0.0007922	0.0014531	0.05
Middle Islamic II	Vitis	60	1094.0	32	15	0.25	0.03	0.02	0.0004875	0.0008069	0.03

```

#this generates a graph that combines the bootstrapped ubiquity data
#with aspects of the weighted proportion summary above

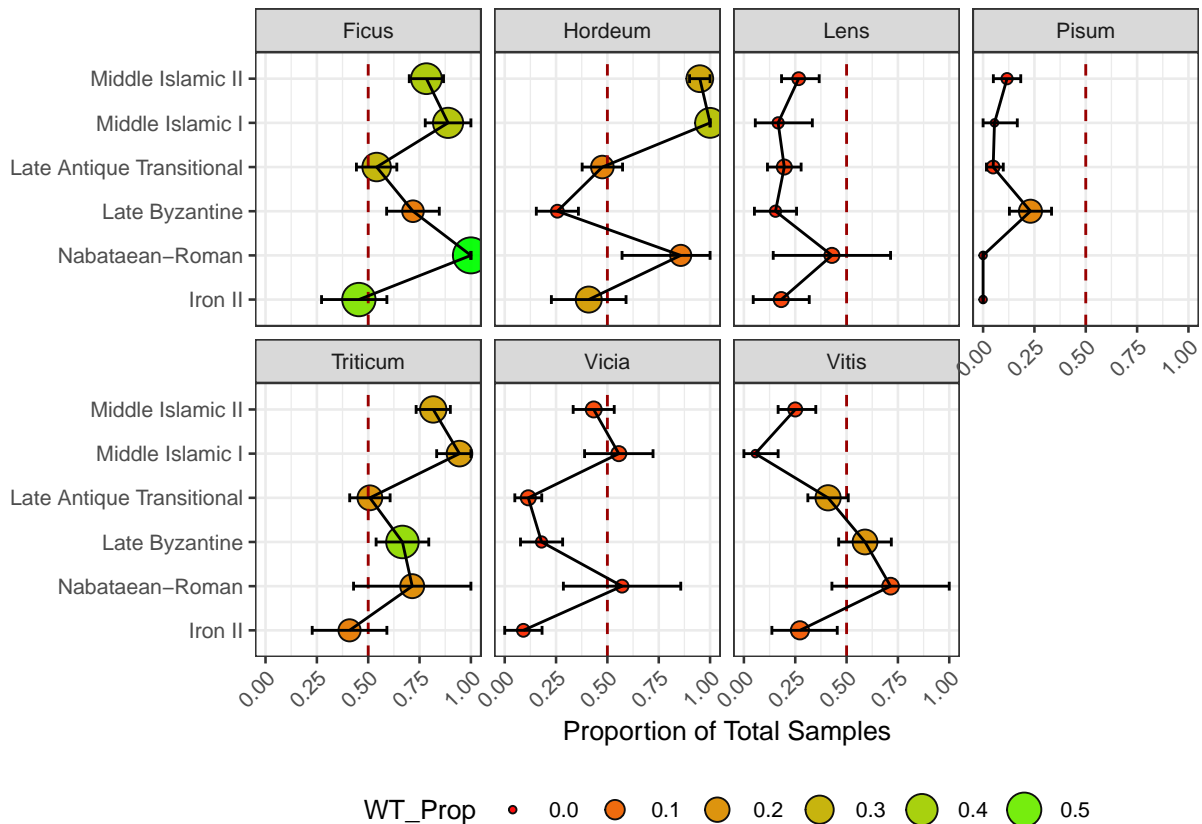
test %>%
  left_join(dom.summary %>% select(Period, Taxon, WT_Prop), by = c("Period", "Taxon")) %>%
  ggplot(aes(Period,mean))+
  geom_point(alpha=.95,shape = 21, aes(size = WT_Prop, fill = WT_Prop))+
  scale_fill_continuous(low="red", high="green") +

```

```

theme_bw(base_size = 10)+
ylab("Proportion of Total Samples")+
xlab("")+
facet_wrap(~Taxon, nrow = 2) +
geom_hline(aes(yintercept=.5), colour="#990000", linetype="dashed")+
geom_errorbar(aes(ymin=low,ymax=high),width=.2) +
geom_line(aes(group=1)) +
theme(axis.text.x = element_text(angle = 45, hjust = 1), legend.position = "bottom") +
coord_flip() +
guides(fill=guide_legend(nrow = 1), size = guide_legend())

```



## Seed Morphometrics

The seed morphometrics relies on a subset of several different kinds of measurements taken primarily on cereal grains, but also including grape and vetch seeds. Several of the analyzed seeds are not included in the analysis, and their Period designation appears as NA due to the fact that those contexts were removed. Their removal is predicated on clear bioturbation in the deposits from which they were collected, again visible in the AMS C14 chronology.

```

library(broom)
library(DTK)

#import the csv, either change the path or filename
morph <- read.csv("farahani_seed_morph_v2.csv")

```

```
#attach the relevant time periods from the analysed material
morph <- morph %>% left_join(dhb.peb %>% select(LF, Period), by = "LF")
```

```
#basic pairwise comparison using a welch's t-test
morph %>% filter(Genus == "Hordeum", !is.na(Thickness)) %>%
  do(tidy(pairwise.t.test(.$Thickness, .$Period))) %>%
  arrange(p.value) %>% kable(.)
```

group1	group2	p.value
Late Antique Transitional	Iron II	0.0013466
Middle Islamic II	Late Antique Transitional	0.0013466
Late Byzantine	Iron II	0.0480620
Middle Islamic II	Late Byzantine	0.1216450
Middle Islamic I	Iron II	0.4154407
Nabataean-Roman	Iron II	0.6272095
Late Antique Transitional	Nabataean-Roman	0.6621703
Middle Islamic I	Late Antique Transitional	0.6716318
Middle Islamic II	Middle Islamic I	0.7970466
Middle Islamic II	Iron II	1.0000000
Late Byzantine	Nabataean-Roman	1.0000000
Middle Islamic I	Nabataean-Roman	1.0000000
Middle Islamic II	Nabataean-Roman	1.0000000
Late Antique Transitional	Late Byzantine	1.0000000
Middle Islamic I	Late Byzantine	1.0000000

```
morph %>% filter(Genus == "Triticum", !is.na(Thickness)) %>%
  do(tidy(pairwise.t.test(.$Thickness, .$Period))) %>%
  arrange(p.value) %>% kable(.)
```

group1	group2	p.value
Late Antique Transitional	Iron II	0.9675700
Middle Islamic II	Iron II	0.9866075
Nabataean-Roman	Iron II	1.0000000
Late Byzantine	Iron II	1.0000000
Middle Islamic I	Iron II	1.0000000
Late Byzantine	Nabataean-Roman	1.0000000
Late Antique Transitional	Nabataean-Roman	1.0000000
Middle Islamic I	Nabataean-Roman	1.0000000
Middle Islamic II	Nabataean-Roman	1.0000000
Late Antique Transitional	Late Byzantine	1.0000000
Middle Islamic I	Late Byzantine	1.0000000
Middle Islamic II	Late Byzantine	1.0000000
Middle Islamic I	Late Antique Transitional	1.0000000
Middle Islamic II	Late Antique Transitional	1.0000000
Middle Islamic II	Middle Islamic I	1.0000000

```
#using the Type III SS for anova and thickness
#the Type III SS for ANOVA does not use order of appearance
```

```
morph %>% filter(Genus == "Hordeum", !is.na(Thickness)) %>%
  do(drop1(aov(Thickness ~ Period, data = .), ~., test = "F"))
```

```
## Single term deletions
##
## Model:
```

```
## Thickness ~ Period
##           Df Sum of Sq    RSS      AIC F value    Pr(>F)
## <none>                10.823 -191.19
## Period  5      3.3739 14.197 -175.69  5.4864 0.0001923 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*#a basic pairwise comparison is not correct because  
#of vastly unequal sample sizes, a Tukey HSD is a post-hoc correction for this*

```
morph %>% filter(Genus == "Hordeum", !is.na(Thickness)) %>%
  do(tidy(TukeyHSD(aov(Thickness ~ Period, data = .)))) %>%
  mutate(comparison = gsub("Nabataean-Roman", "NabataeanRoman", comparison)) %>%
  separate(comparison, into = c("Per1", "Per2"), sep = "-") %>%
  arrange(conf.low) %>%
  filter(adj.p.value < .05) %>% kable(.)
```

term	Per1	Per2	estimate	conf.low	conf.high	adj.p.value
Period	Middle Islamic II	Late Antique Transitional	-0.3986749	-0.6822667	-0.1150830	0.0012752
Period	Late Byzantine	Iron II	0.5230000	0.0121258	1.0338742	0.0416377
Period	Late Antique Transitional	Iron II	0.5057304	0.1469042	0.8645566	0.0012288

*#this uses the Dunning method, which is more appropriate  
#for unbalanced sample designs with unequal sample variances*

```
test <- morph %>% filter(Genus == "Hordeum", !is.na(Thickness), !is.na(Period)) %>%
  droplevels %>%
  do(tidy(DTK.test(.$Thickness, .$Period)[[2]])) %>%
  mutate(.rownames = gsub("Nabataean-Roman", "NabataeanRoman", .rownames)) %>%
  separate(.rownames, into = c("Per1", "Per2"), sep = "-")
```

*#this now loops through the Dunning output looking for those periods  
#which are significantly different in terms of thickness for barley*

```
vec <- vector("list", nrow(test))
```

```
for(i in 1:nrow(test)){
  vec[i] <- between(0, test$Lower.CI[i], test$Upper.CI[i])
}
```

```
test[unlist(vec) == FALSE,] %>% kable(.)
```

Per1	Per2	Diff	Lower.CI	Upper.CI
Late Byzantine	Iron II	0.5230000	0.0087063	1.0372937
Late Antique Transitional	Iron II	0.5057304	0.0754911	0.9359697
Middle Islamic II	Late Antique Transitional	-0.3986749	-0.7397446	-0.0576052

*#the variances across periods are similar for barley and wheat  
#but not between individual periods*

```
library(car)
morph %>% filter(Genus == "Hordeum", !is.na(Thickness), !is.na(Period)) %>%
  do(leveneTest(.$Thickness, .$Period, center = median))
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##           Df F value Pr(>F)
## group  5  1.8623 0.1091
##           88
```

```
morph %>% filter(Genus == "Hordeum", !is.na(Thickness), !is.na(Period)) %>%
  do(leveneTest(.$Thickness, .$Period, center = median))
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 5  1.8623 0.1091
##      88
```

*#and now, the f-tests, which are reported in the paper*  
*#this code could be improved with the map function from purrr,*  
*#especially splitting by genus -- I encourage someone to try!*

```
var.test(
  morph %>%
    filter(Genus == "Hordeum", !is.na(Thickness), !is.na(Period)) %>%
    filter(Period == "Late Antique Transitional") %>% pull(Thickness),
  morph %>%
    filter(Genus == "Hordeum", !is.na(Thickness), !is.na(Period)) %>%
    filter(Period == "Middle Islamic II") %>%
    pull(Thickness)
)
```

```
##
## F test to compare two variances
##
## data:  morph %>% filter(Genus == "Hordeum", !is.na(Thickness), !is.na(Period)) %>% and morph %>% fi
## F = 0.38168, num df = 24, denom df = 26, p-value = 0.02038
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1721281 0.8573966
## sample estimates:
## ratio of variances
##           0.38168
```

```
var.test(
  morph %>%
    filter(Genus == "Triticum", !is.na(Thickness), !is.na(Period)) %>%
    filter(Period == "Late Antique Transitional") %>% pull(Thickness),
  morph %>%
    filter(Genus == "Triticum", !is.na(Thickness), !is.na(Period)) %>%
    filter(Period == "Middle Islamic II") %>% pull(Thickness)
)
```

```
##
## F test to compare two variances
##
## data:  morph %>% filter(Genus == "Triticum", !is.na(Thickness), !is.na(Period)) %>% and morph %>% f
## F = 0.70355, num df = 30, denom df = 32, p-value = 0.3359
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.3447345 1.4478937
## sample estimates:
## ratio of variances
##           0.7035468
```

*#this plots the data using a normal\_cl\_boot*  
*#hacky fix for the levels*

```

morph %>%
  add_count(Genus, Period) %>%
  filter(!is.na(Period), !is.na(Thickness)) %>%
  mutate(Genus = ifelse(Genus == "Hordeum", "Barley", "Wheat")) %>%
  mutate(Period2 = as.factor(paste(Period, "(n = ", n, ")")) %>%
  mutate(Period2 = factor(Period2, levels = levels(Period2)[c(1,2,11,12,3,4,5,6,7,8,9,10)])) %>%
  ggplot(aes(Period2, Thickness)) +
  geom_point(aes(colour = Period)) +
  theme_bw(base_size = 10) +
  stat_summary(fun.data = mean_cl_boot, colour = "red", size = .5, aes(shape = Period)) +
  stat_summary(fun.y=mean, colour="red", geom="line", aes(group = 1)) +
  theme(panel.grid = element_blank(), axis.text.x = element_text(angle = 45, hjust = 1),
        legend.position = "none") +
  facet_wrap(~Genus, scales="free_x") + ylab("Thickness (mm)") + xlab("")

```

