Australian Journal of Botany 66(1), 48–58

http://dx.doi.org/10.1071/BT17116_AC

Supplementary material

Genetic and environmental parameters show associations with essential oil composition in West Australian sandalwood (*Santalum spicatum*)

Jessie Moniodis[A,D], *Michael Renton*[A,B], *Christopher G. Jones*[A], *E. Liz Barbour*[A] *and Margaret Byrne*[A,C]

[A]School of Biological Sciences (M084), University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia.

[B]School of Agriculture and Environment, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia.

[C]Science and Conservation Division, Department of Biodiversity, Conservation and Attractions, Locked Bag 104, Bentley Delivery Centre, WA 6983, Australia.

[D]Corresponding author. Email: jess.moniodis@uwa.edu.au

R script

Mantel Tests (Table 1)

```
### read and process data


mydata <- read.csv('complete_FINAL_genetics.csv') ## for all or

mydata <- read.csv('chemistry_data_for_w_genetics.csv') ## for wheatbelt

head(mydata)

names(mydata)

mydata$DISTANCE..M. <- mydata$DISTANCE.M

substr(mydata$DISTANCE..M.,1,1)

unique(substr(mydata$DISTANCE..M.,1,1))

inds <- which(substr(mydata$DISTANCE..M.,1,1) %in% c('W','G','N'))

nsamples <- length(inds)

data1 <- mydata[inds,]

allhosts <- NULL

for (i in 1:nsamples ){
```

```
        ind <- inds[i]
        startind <- ind+1
        if (i<nsamples) stopind <- inds[i+1]-2 else stopind <- length(mydata$DISTANCE..M.)
        print(paste('for',ind,mydata$DISTANCE..M.[ind]))
        for (ii in startind:stopind) {
                dist <- as.numeric(levels(mydata$DISTANCE..M.))[mydata$DISTANCE..M.[ii]]
                print(dist)
                print(as.character(mydata$Host[ii]))
                allhosts <- rbind(allhosts, c(mydata$DISTANCE..M.[ind],
as.character(mydata$Host[ii]), levels(mydata$DISTANCE..M.)[mydata$DISTANCE..M.[ii]]))
        }
}

data1
head(data1)
names(data1)
str(data1)
length(data1[,1])

## read genetic distances
gendat <- read.csv('genetic_data_for_all.csv',skip=2) ## for 186 or
gendat <- read.csv('w_genetics.csv',skip=2) ## wheatbelt
rownames(gendat) <- gendat$site
head(gendat)
gendist <- dist(gendat[1:186,1:186])   ## for 186 or
gendist <- dist(gendat[1:144,1:144]) ## wheatbelt
names(gendat[1:186])  ## 186 or
names(gendat[1:144]) ## wheatbelt
## 186 sites for all, 144 for wheatbelt
```

```
### match the two data sets
data1$name <- data1$DISTANCE..M.
length(data1$name)
## 186 sites for all, 144 for wheatbelt
sort(data1$DISTANCE..M.)
names(gendat)

## 186 all trees
inds <- NULL
indsgen <- NULL
for (i in 1:186){
        ind <- which(data1$name==names(gendat)[i])
        inds <- c(inds, ind)
        if (length(ind)>0) indsgen <- c(indsgen,i)
        print(paste(names(gendat)[i],data1$name[ind],ind ))
}
cbind(inds,indsgen,1:186)
length(inds)
sort(inds)==1:186

## 144 wheatbelt only
inds <- NULL
indsgen <- NULL
for (i in 1:144){
        ind <- which(data1$name==names(gendat)[i])
        inds <- c(inds, ind)
        if (length(ind)>0) indsgen <- c(indsgen,i)
```

```
	print(paste(names(gendat)[i],data1$name[ind],ind ))
}
cbind(inds,indsgen,1:144)
length(inds)
sort(inds)==1:144
## looks good

library(vegan)

gendist <- dist(gendat[indsgen,indsgen])
names(gendat)[indsgen]

geodist <- dist(cbind(data1[["GPS.UTM.E"]][inds],data1[["GPS.UTM.N"]][inds]))
geodist

region <- factor(substr(data1$DISTANCE.M,1,1))

#### mantel test for chem-gen-geo relationship overall

mantel(gendist,geodist)
plot(gendist,geodist)

#### mantel test for chem-gen-geo relationship

niceones <- c(40,41,42,44,45,46,47,48,49,50,51,52,53,54,55)
myimportantoils <- data1[inds,niceones]
head(myimportantoils )
myimportantoils.dist <- dist(myimportantoils)
```

```
mantel(gendist,myimportantoils.dist,permutations=999)

mantel(geodist,myimportantoils.dist,permutations=999)

for (i in niceones){
thisimportantoil <- data1[inds,i]
thisimportantoil.dist <- dist(thisimportantoil )
print(names(data1)[i])
print(mantel(gendist,thisimportantoil.dist))
}

### mantel test for individuals

## gendist

p.alpha.santalene.dist <- dist(data1$p.alpha.santalene[inds])
mantel(gendist, p.alpha.santalene.dist)

p.alpha.santalol.dist <- dist(data1$p.alpha.santalol[inds])
mantel(gendist, p.alpha.santalol.dist)

p.beta.santalol.dist <- dist(data1$p.beta.santalol[inds])
mantel(gendist, p.beta.santalol.dist)

Santalol.total.percent.dist <- dist(data1$ Santalol.total.percent[inds])
mantel(gendist, Santalol.total.percent.dist)

p.trans.trans.farnesol.dist <- dist(data1$ p.trans.trans.farnesol[inds])
```

```
mantel(gendist, p.trans.trans.farnesol.dist)

Santalol.farnesol.ratio.dist <- dist(data1$ Santalol.farnesol.ratio[inds])
mantel(gendist, Santalol.farnesol.ratio.dist)

Total.percent.dist <- dist(data1$ Total.percent[inds])
mantel(gendist, Total.percent.dist)

## geodist
p.alpha.santalene.dist <- dist(data1$p.alpha.santalene[inds])
mantel(geodist, p.alpha.santalene.dist)

p.alpha.santalol.dist <- dist(data1$p.alpha.santalol[inds])
mantel(geodist, p.alpha.santalol.dist)

p.beta.santalol.dist <- dist(data1$p.beta.santalol[inds])
mantel(geodist, p.beta.santalol.dist)

Santalol.total.percent.dist <- dist(data1$ Santalol.total.percent[inds])
mantel(geodist, Santalol.total.percent.dist)

p.trans.trans.farnesol.dist <- dist(data1$ p.trans.trans.farnesol[inds])
mantel(geodist, p.trans.trans.farnesol.dist)

Santalol.farnesol.ratio.dist <- dist(data1$ Santalol.farnesol.ratio[inds])
mantel(geodist, Santalol.farnesol.ratio.dist)

Total.percent.dist <- dist(data1$ Total.percent[inds])
```

```
mantel(geodist, Total.percent.dist)
```

Multidimensional scaling (Fig. 4)

```
mds1 <- metaMDS(gendist) ## fits an MDS

plot(mds1$points,col=as.numeric(region[inds]))

siz <- valstouse/max(valstouse)*2+0.1

plot(mds1,cex=siz )


mds1 <- metaMDS(gendist) ## fits an MDS
types = as.numeric(region[inds])
plot(mds1$points,col=c('grey40','grey20','black')[types],pch=c(16,2,3)[types])
```

Linear models (Table 2)

DO ENVIRONMENT AND MORPHOLOGICAL TRAITS INFLUENCE OIL CHEMISTRY?

1. Total santalol
2. Farnesol
3. Santalol:farnesol ratio
4. Total percent
5. Yield percent

####

1. Santalol.total.percent

Environment

a. Does ..... depend on Region.2? ## (wheatbelt, goldfields, north)

```
#help(abline)
```

```
data1$Region.2 <- factor(data1$Region.2)
plot(data1$Santalol.total.percent~data1$Region.2  )   #### plot
mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$Region.2 )  ## linear model
anova(mylinearmodel)
TukeyHSD(aov(mylinearmodel))
#help(plot)
```

b. Does ..... depend on PROVENANCE? ##

```
#help(abline)
```

```
plot(data1$PROVENANCE,data1$Santalol.total.percent  )    #### plot
mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$PROVENANCE )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
TukeyHSD(aov(mylinearmodel))
```

c. Does ..... depend on the number of potential hosts? ##

```
#help(abline)

plot(data1$Host.Number,data1$Santalol.total.percent  )    #### plot
mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$Host.Number )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

d. Does ..... depend on the main host species? ##

```
#help(abline)

plot(data1$Main.Host,data1$Santalol.total.percent  )    #### plot
mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$Main.Host )  ## linear model
anova(mylinearmodel)
tapply(data1$Santalol.total.percent , data1$Main.Host, mean)
```

e. Does ..... depend on altitude? ##

```
#help(abline)

plot(data1$Altitude,data1$Santalol.total.percent  )    #### plot
mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$Altitude )  ## linear model
summary(mylinearmodel)
```

```
anova(mylinearmodel)
```

f. Does ..... depend on rainfall? ##

```
#help(abline)

plot(data1$rainfall,data1$Santalol.total.percent  )    #### plot
mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$rainfall )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

g. Does ..... depend on soil type? ##

```
#help(abline)

plot(data1$SOIL,data1$Santalol.total.percent )    #### plot
mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$SOIL )  ## linear model
anova(mylinearmodel)
tapply(data1$Santalol.total.percent , data1$SOIL, mean)
```

h. Does ..... depend on pH? ##

```
#help(abline)

plot(data1$PH,data1$Santalol.total.percent  )    #### plot
mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$PH )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

Morphology

i. Does ..... depend on tree height? ##

```
#help(abline)


plot(data1$HEIGHT.M,data1$Santalol.total.percent  )    #### plot

mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$HEIGHT.M )  ## linear model

summary(mylinearmodel)

anova(mylinearmodel)
```

ii. Does ..... depend on Stem diameter ? ##

```
#help(abline)


plot(data1$STEM.DIAMETER.15.CM,data1$Santalol.total.percent  )    #### plot

mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$STEM.DIAMETER.15.CM )  ## linear model

summary(mylinearmodel)

anova(mylinearmodel)


plot(data1$STEM.DIAMETER.30.CM,data1$Santalol.total.percent  )    #### plot

mylinearmodel <- lm( data1$Santalol.total.percent  ~ data1$STEM.DIAMETER.30.CM )  ## linear model

summary(mylinearmodel)

anova(mylinearmodel)


####
```

2. p.trans.trans.farnesol

Environment

a. Does ..... depend on Region.2? ## (wheatbelt, goldfields, north)

```
#help(abline)

data1$Region.2 <- factor(data1$Region.2)
plot(data1$p.trans.trans.farnesol~data1$Region.2  )   #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$Region.2 )  ## linear model
anova(mylinearmodel)
TukeyHSD(aov(mylinearmodel))
#help(plot)
```

b. Does ..... depend on PROVENANCE? ##

```
#help(abline)

plot(data1$PROVENANCE,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$PROVENANCE )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

c. Does ..... depend on the number of potential hosts? ##

```
#help(abline)

plot(data1$Host.Number,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$Host.Number )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

d. Does ..... depend on the main host species? ##

```
#help(abline)
```

```
plot(data1$Main.Host,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$Main.Host )  ## linear model
anova(mylinearmodel)
tapply(data1$p.trans.trans.farnesol , data1$Main.Host, mean)
```

e. Does ..... depend on altitude? ##

```
#help(abline)

plot(data1$Altitude,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$Altitude )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

f. Does ..... depend on rainfall? ##

```
#help(abline)

plot(data1$rainfall,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$rainfall )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

g. Does ..... depend on soil type? ##

```
#help(abline)

plot(data1$SOIL,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$SOIL )  ## linear model
anova(mylinearmodel)
tapply(data1$p.trans.trans.farnesol , data1$SOIL, mean)
```

h. Does ..... depend on pH? ##

```
#help(abline)

plot(data1$PH,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$PH )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

Morphology

i. Does ..... depend on tree height? ##

```
#help(abline)

plot(data1$HEIGHT.M,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$HEIGHT.M )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

ii. Does ..... depend on Stem diameter ? ##

```
#help(abline)

plot(data1$STEM.DIAMETER.15.CM,data1$p.trans.trans.farnesol  )    #### plot
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$STEM.DIAMETER.15.CM )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)

plot(data1$STEM.DIAMETER.30.CM,data1$p.trans.trans.farnesol  )    #### plot
```

```
mylinearmodel <- lm( data1$p.trans.trans.farnesol  ~ data1$STEM.DIAMETER.30.CM )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)


####
```

3. Santalol.farnesol.ratio

Environment

a. Does ..... depend on Region.2? ## (wheatbelt, goldfields, north)

```
#help(abline)


data1$Region.2 <- factor(data1$Region.2)
plot(data1$Santalol.farnesol.ratio~data1$Region.2  )   #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$Region.2 )  ## linear model
anova(mylinearmodel)
TukeyHSD(aov(mylinearmodel))
#help(plot)
```

b. Does ..... depend on PROVENANCE? ##

```
#help(abline)


plot(data1$PROVENANCE,data1$Santalol.farnesol.ratio  )    #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$PROVENANCE )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

c. Does ..... depend on the number of potential hosts? ##

```
#help(abline)


plot(data1$Host.Number,data1$Santalol.farnesol.ratio  )    #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$Host.Number )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

d. Does ..... depend on the main host species? ##

```
#help(abline)


plot(data1$Main.Host,data1$Santalol.farnesol.ratio  )    #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$Main.Host )  ## linear model
anova(mylinearmodel)
tapply(data1$Santalol.farnesol.ratio , data1$Main.Host, mean)
```

e. Does ..... depend on altitude? ##

```
#help(abline)


plot(data1$Altitude,data1$Santalol.farnesol.ratio  )    #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$Altitude )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

f. Does ..... depend on rainfall? ##

```
#help(abline)


plot(data1$rainfall,data1$Santalol.farnesol.ratio  )    #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$rainfall )  ## linear model
```

```
summary(mylinearmodel)
anova(mylinearmodel)
```

g. Does ..... depend on soil type? ##

```
#help(abline)

plot(data1$SOIL,data1$Santalol.farnesol.ratio  )    #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$SOIL )  ## linear model
anova(mylinearmodel)
tapply(data1$Santalol.farnesol.ratio , data1$SOIL, mean)
```

h. Does ..... depend on pH? ##

```
#help(abline)

plot(data1$PH,data1$Santalol.farnesol.ratio  )    #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$PH )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

Morphology

i. Does ..... depend on tree height? ##

```
#help(abline)

plot(data1$HEIGHT.M,data1$Santalol.farnesol.ratio  )    #### plot
mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$HEIGHT.M )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

ii. Does ..... depend on Stem diameter ? ##

```
#help(abline)


plot(data1$STEM.DIAMETER.15.CM,data1$Santalol.farnesol.ratio  )    #### plot

mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$STEM.DIAMETER.15.CM )  ## linear model

summary(mylinearmodel)

anova(mylinearmodel)


plot(data1$STEM.DIAMETER.30.CM,data1$Santalol.farnesol.ratio  )    #### plot

mylinearmodel <- lm( data1$Santalol.farnesol.ratio  ~ data1$STEM.DIAMETER.30.CM )  ## linear model

summary(mylinearmodel)

anova(mylinearmodel)


####
```

4. Sum of sesquiterpenes

Environment

a. Does ..... depend on Region.2? ## (wheatbelt, goldfields, north)

```
#help(abline)


data1$Region.2 <- factor(data1$Region.2)

plot(data1$Total.percent~data1$Region.2  )   #### plot

mylinearmodel <- lm( data1$Total.percent  ~ data1$Region.2 )  ## linear model

anova(mylinearmodel)

TukeyHSD(aov(mylinearmodel))
```

```
#help(plot)
```

b. Does ..... depend on PROVENANCE? ##

```
#help(abline)

plot(data1$PROVENANCE,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$PROVENANCE )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

c. Does ..... depend on the number of potential hosts? ##

```
#help(abline)

plot(data1$Host.Number,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$Host.Number )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

d. Does ..... depend on the main host species? ##

```
#help(abline)

plot(data1$Main.Host,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$Main.Host )  ## linear model
anova(mylinearmodel)
tapply(data1$Total.percent , data1$Main.Host, mean)
```

e. Does ..... depend on altitude? ##

```
#help(abline)
```

```
plot(data1$Altitude,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$Altitude )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

f. Does ..... depend on rainfall? ##

```
#help(abline)

plot(data1$rainfall,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$rainfall )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

g. Does ..... depend on soil type? ##

```
#help(abline)

plot(data1$SOIL,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$SOIL )  ## linear model
anova(mylinearmodel)
tapply(data1$Total.percent , data1$SOIL, mean)
```

h. Does ..... depend on pH? ##

```
#help(abline)

plot(data1$PH,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$PH )  ## linear model
summary(mylinearmodel)
```

```
anova(mylinearmodel)
```

Morphology

i. Does ..... depend on tree height? ##

```
#help(abline)


plot(data1$HEIGHT.M,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$HEIGHT.M )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

ii. Does ..... depend on Stem diameter ? ##

```
#help(abline)


plot(data1$STEM.DIAMETER.15.CM,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$STEM.DIAMETER.15.CM )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)

plot(data1$STEM.DIAMETER.30.CM,data1$Total.percent  )    #### plot
mylinearmodel <- lm( data1$Total.percent  ~ data1$STEM.DIAMETER.30.CM )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)

####
```

5. YIELD PERCENT

Environment

a. Does ..... depend on Region.2? ## (wheatbelt, goldfields, north)

```
#help(abline)

data1$Region.2 <- factor(data1$Region.2)
plot(data1$Yield.percent~data1$Region.2  )   #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$Region.2 )  ## linear model
anova(mylinearmodel)
TukeyHSD(aov(mylinearmodel))
#help(plot)
```

b. Does ..... depend on PROVENANCE? ##

```
#help(abline)

plot(data1$PROVENANCE,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$PROVENANCE )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

c. Does ..... depend on the number of potential hosts? ##

```
#help(abline)

plot(data1$Host.Number,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$Host.Number )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

d. Does ..... depend on the main host species? ##

```
#help(abline)

plot(data1$Main.Host,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$Main.Host )  ## linear model
anova(mylinearmodel)
tapply(data1$Yield.percent , data1$Main.Host, mean)
```

e. Does ..... depend on altitude? ##

```
#help(abline)

plot(data1$Altitude,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$Altitude )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

f. Does ..... depend on rainfall? ##

```
#help(abline)

plot(data1$rainfall,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$rainfall )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

g. Does ..... depend on soil type? ##

```
#help(abline)

plot(data1$SOIL,data1$Yield.percent  )    #### plot
```

```
mylinearmodel <- lm( data1$Yield.percent  ~ data1$SOIL )  ## linear model
anova(mylinearmodel)
tapply(data1$Yield.percent , data1$SOIL, mean)
```

h. Does ..... depend on pH? ##

```
#help(abline)

plot(data1$PH,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$PH )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

Morphology

i. Does ..... depend on tree height? ##

```
#help(abline)

plot(data1$HEIGHT.M,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$HEIGHT.M )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)
```

ii. Does ..... depend on Stem diameter ? ##

```
#help(abline)

plot(data1$STEM.DIAMETER.15.CM,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$STEM.DIAMETER.15.CM )  ## linear model
summary(mylinearmodel)
```

```
anova(mylinearmodel)

plot(data1$STEM.DIAMETER.30.CM,data1$Yield.percent  )    #### plot
mylinearmodel <- lm( data1$Yield.percent  ~ data1$STEM.DIAMETER.30.CM )  ## linear model
summary(mylinearmodel)
anova(mylinearmodel)

## Adonis test Table 2

#### does overall oil composition vary consistently with space?
adonis(myimportantoils.dist ~ data1[["GPS.UTM.E"]]+data1[["GPS.UTM.N"]])

#### does overall oil composition vary consistently with main host?
adonis(myimportantoils.dist ~ data1[["Main.Host"]])

#### does overall oil composition vary consistently with provenance?
adonis(myimportantoils.dist ~ data1[["PROVENANCE"]])

#### does overall oil composition vary consistently with height?
expvar = data1[["HEIGHT.M"]]
nonnas = !is.na(expvar)
myimportantoils <- data1[nonnas ,niceones]
myimportantoils.dist <- dist(myimportantoils)
adonis(myimportantoils.dist ~ expvar[nonnas])

#### does overall oil composition vary consistently with stem diameter?
expvar = data1[["STEM.DIAMETER.15.CM"]]
nonnas = !is.na(expvar)
```

```
myimportantoils <- data1[nonnas ,niceones]
myimportantoils.dist <- dist(myimportantoils)
adonis(myimportantoils.dist ~ expvar[nonnas])

#### does overall oil composition vary consistently with PH?
expvar = data1[["PH"]]
nonnas = !is.na(expvar)
myimportantoils <- data1[nonnas ,niceones]
myimportantoils.dist <- dist(myimportantoils)
adonis(myimportantoils.dist ~ expvar[nonnas])
```