

Function to Presentation

Brett

January 04, 2018

Contents

Introduction:	1
Read data from SQL database and tabulated files	1
Defining Functions	2
Plots (final images made in plotting window)	3
BrS1 meta analysis	3
LQT3 meta analysis	5
Functional Perturbation vs. Calculated Penetrance	7
Sensitivity Analysis	29
Empirical Bayes (used in manuscript)	29
Uninformative Prior	31
Optimistic Prior	33
Pessimistic Prior	35

Introduction:

This file is used to judge empirical relationships between functional parameters and calculated penetrance derived from the SCN5A dataset. There is also a sensitivity analysis done at the end using peak current as an example.

```
require(Hmisc)
library(tableone)
library(ggplot2)
library("DBI")
library("RSQLite")
library(wCorr)
library(meta)
knitrSet(lang = 'markdown', echo = TRUE)
```

Read data from SQL database and tabulated files

```
con = dbConnect(SQLite(),
dbname="/Users/B/Dropbox/SCN5A/BrettsSandbox/paper/data/VariantSCN5A-new.db")
alltables = dbListTables(con)
my.data <- dbReadTable(con, 'VariantSCN5A')
my.data[my.data=='NA'] <- NA
d<-my.data
dbDisconnect(con)

d$resnum<-as.integer(d$resnum)
```

```

d$gnomAD[is.na(d$gnomAD)] <- 0
d$gnomAD<-as.numeric(d$gnomAD)
d$ipeak<-100*as.numeric(d$ipeak)

d$vhalfact<-as.numeric(d$vhalfact)

d$vhalfinact<-as.numeric(d$vhalfinact)
d$recovfrominact<-log10(100*as.numeric(d$recovfrominact))
d$recovfrominact_dist<-log10(100*as.numeric(d$recovfrominact_dist))
d$ilate[as.numeric(d$ilate)==0]<-NA
d$ilate_norm<-log10(d$ipeak*as.numeric(d$ilate)+0.00001)
d$ilate<-log10(100*as.numeric(d$ilate)+0.00001)
d$ilate_dist<-log10(100*as.numeric(d$ilate_dist)+0.00001)
d$total_carriers<-d$lqt3+d$brs1+d$unaff+d$gnomAD
d$weight = 1-1/(0.1+d$total_carriers) #weights
d$weightsMilder = 1-1/(1+d$total_carriers) #weights

# Adding in penetrance variables
abrs0=0.32
alqt0=0.11
beta0=1
d$LQT_penetranceBayesian<-(d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0)
d$BrS_penetranceBayesian<-(d$brs1+abrs0)/(d$total_carriers+1+alqt0)
d$all_penetranceBayesian<-(d$brs1+abrs0+d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0+abrs0)

servers<-read.csv("/Users/B/Dropbox/SCN5A/BrettsSandbox/paper/data/annotated_variants-trim.txt", sep =
provean <-read.csv("/Users/B/Dropbox/SCN5A/BrettsSandbox/paper/data/provean.txt", sep = "\t")
pph2 <-read.csv("/Users/B/Dropbox/SCN5A/BrettsSandbox/paper/data/pph2-short.txt", sep = "\t")
sift <-read.csv("/Users/B/Dropbox/SCN5A/BrettsSandbox/paper/data/SIFT.txt", sep = "\t")
d <- merge(d, servers, all = TRUE)
d <- merge(d, provean, all = TRUE)
d <- merge(d, sift, all = TRUE)
d <- merge(d, pph2, all = TRUE)
d<-d[!is.na(d$var), ]

d$seaRate<-as.numeric(d$seaRate)
d$blastpssm<-as.numeric(d$blastpssm)
d$pamscore<-as.numeric(d$pamscore)
d<-d[d$total_carriers>0, ]

```

Defining Functions

```

plt.loess <- function(fit, xlabel="X", ylabel="Y", xlims=NA, ylims=NA){
  if(sum(is.na(xlims) & is.na(ylims))>0){
    print(xlims)
    lines(min(fit$x)-1000, 0, xlim=c(min(fit$x), max(fit$x)), ylim=c(min(fit$y), max(fit$y)), ylab=ylabel)
  } else{ if(is.na(xlims)){
    lines(min(fit$x)-1000, 0, xlim=c(min(fit$x), max(fit$x)), ylim=ylims, ylab=ylabel, xlab=xlabel)
  } else{
    lines(min(fit$x)-1000, 0, xlim=xlims, ylim=ylims, ylab=ylabel, xlab=xlabel)
  }
}
xrange <- seq(min(fit$x), max(fit$x), length.out = 100)

```

```

ps <- predict(fit, xrange, se=T)
lines(xrange, ps$fit*100, lwd=5)
lines(xrange, (ps$fit+1.96*ps$se.fit)*100, lty=2, lwd=5)
lines(xrange, (ps$fit-1.96*ps$se.fit)*100, lty=2, lwd=5)
}

plt.disease <- function(func, funcName, disName, xl=c(0,200), sp=0.7){
  par(cex=1, bty='l', lwd=2)
  if (disName=="BrS_penetranceBayesian"){dis="BrS1"; pcolor="gray"}
  else if (disName=="LQT_penetranceBayesian"){dis="LQT3"; pcolor="gray"}
  else if (disName=="all_penetranceBayesian"){dis="BrS1 and LQT3"; pcolor="gray"}
  plot(d[,func], 100*(d[,disName]), pch=21,
       bg=pcolor, cex=log(d[, "total_carriers"])+1, lwd=1.5, axes=FALSE,
       cex.lab=1.5, ylab=paste("Penetrance (%)",dis,""), sep="", xlab = funcName,
       ylim = c(-1,101), xlim = xl)

  axis(side=1,lwd=3, cex.axis=1.5)
  axis(side=2,lwd=3, cex.axis=1.5)

  plt.loess(loess(d[,disName]~as.numeric(d[,func]),span = sp,
    weights = 1-1/(0.1+d[, "total_carriers"])), xlabel=funcName,
    ylims=c(0, 1))
}

```

Plots (final images made in plotting window)

BrS1 meta analysis

```

#plots:

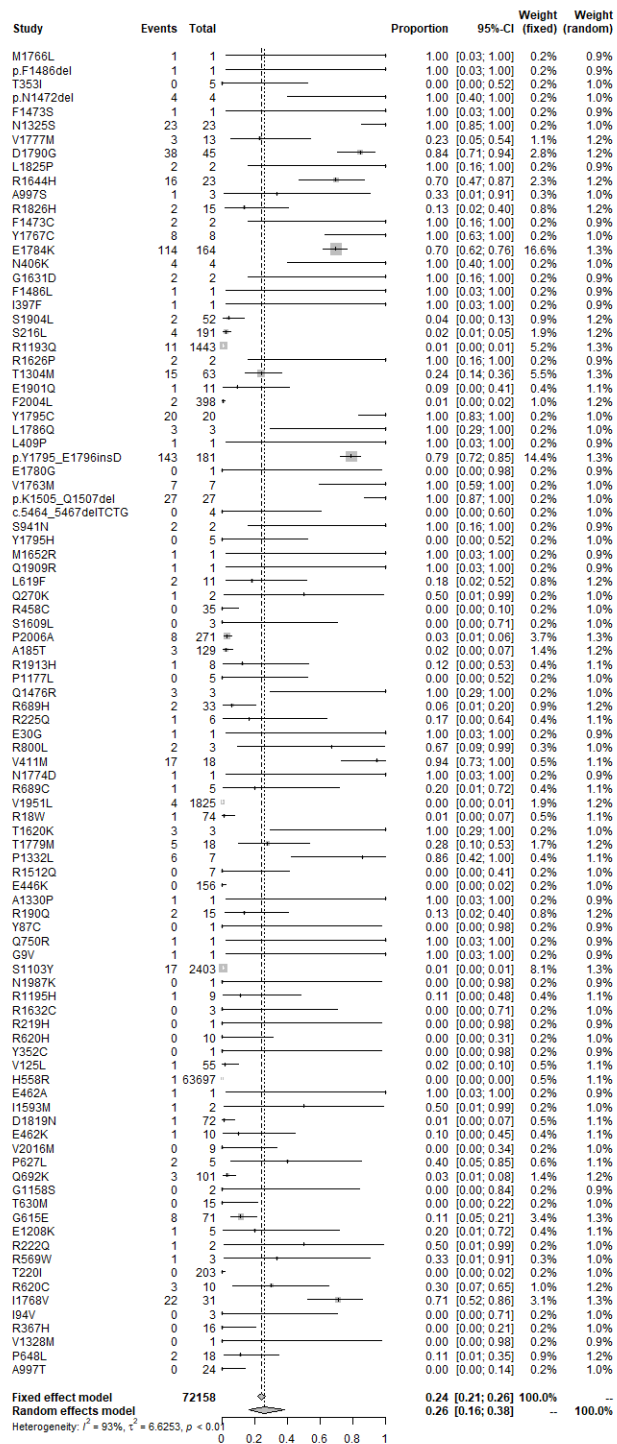
#forest plot of by variant
#BrS1 for variants that have peak current measured
f<-d[!is.na(d$ipeak),]
f<-f[order(f$ipeak, f$var),]
tmp<-metaprop(brs1,total_carriers,data = f,studlab = var)
forest(tmp)

```



LQT3 meta analysis

```
#LQT3 for variants that have late current measured
f<-d[!is.na(d$ilate),]
f<-f[order(-f$ilate, f$var),]
tmp<-metaprop(lqt3,total_carriers,data = f,studlab = var)
forest(tmp)
```

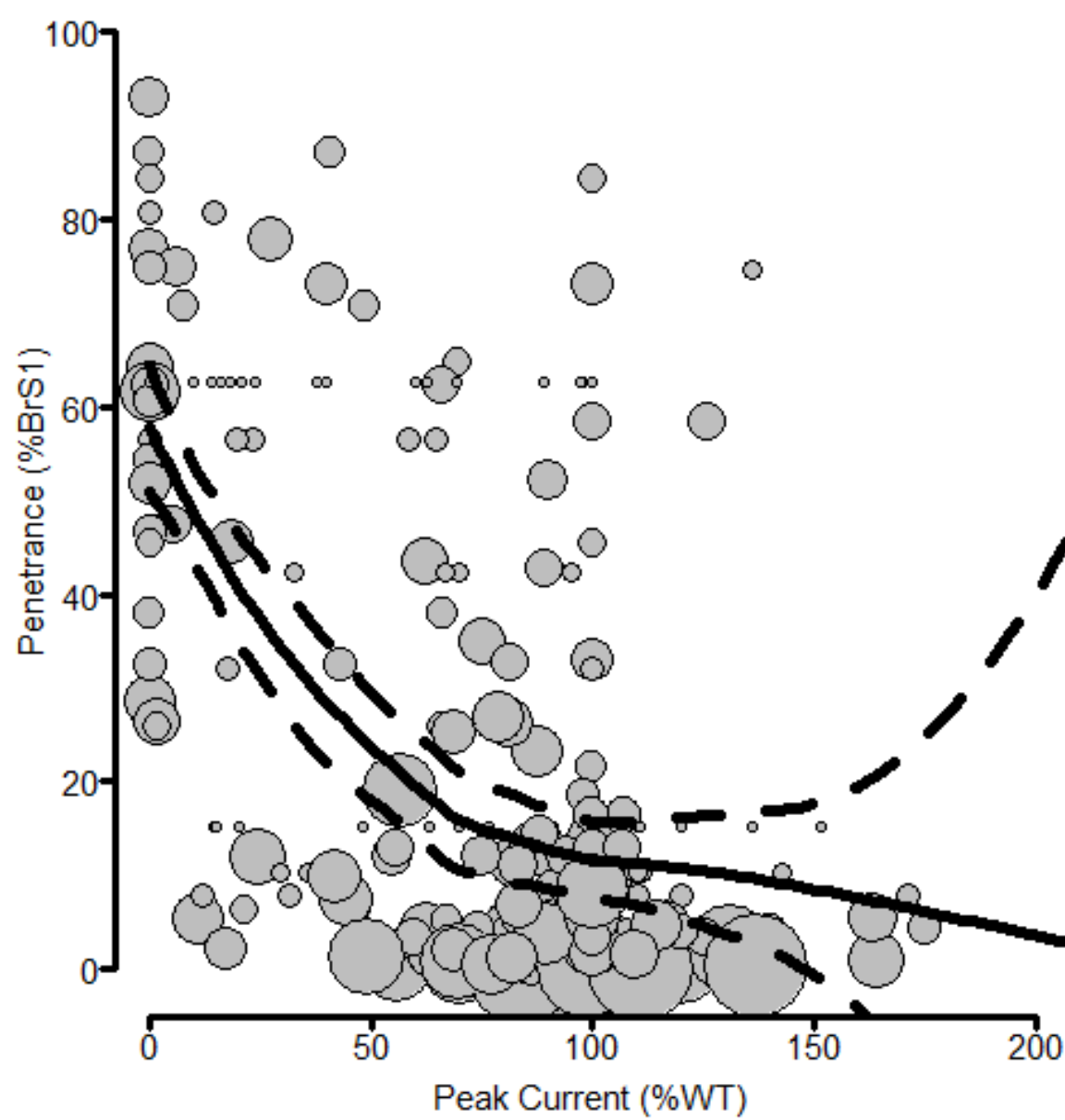


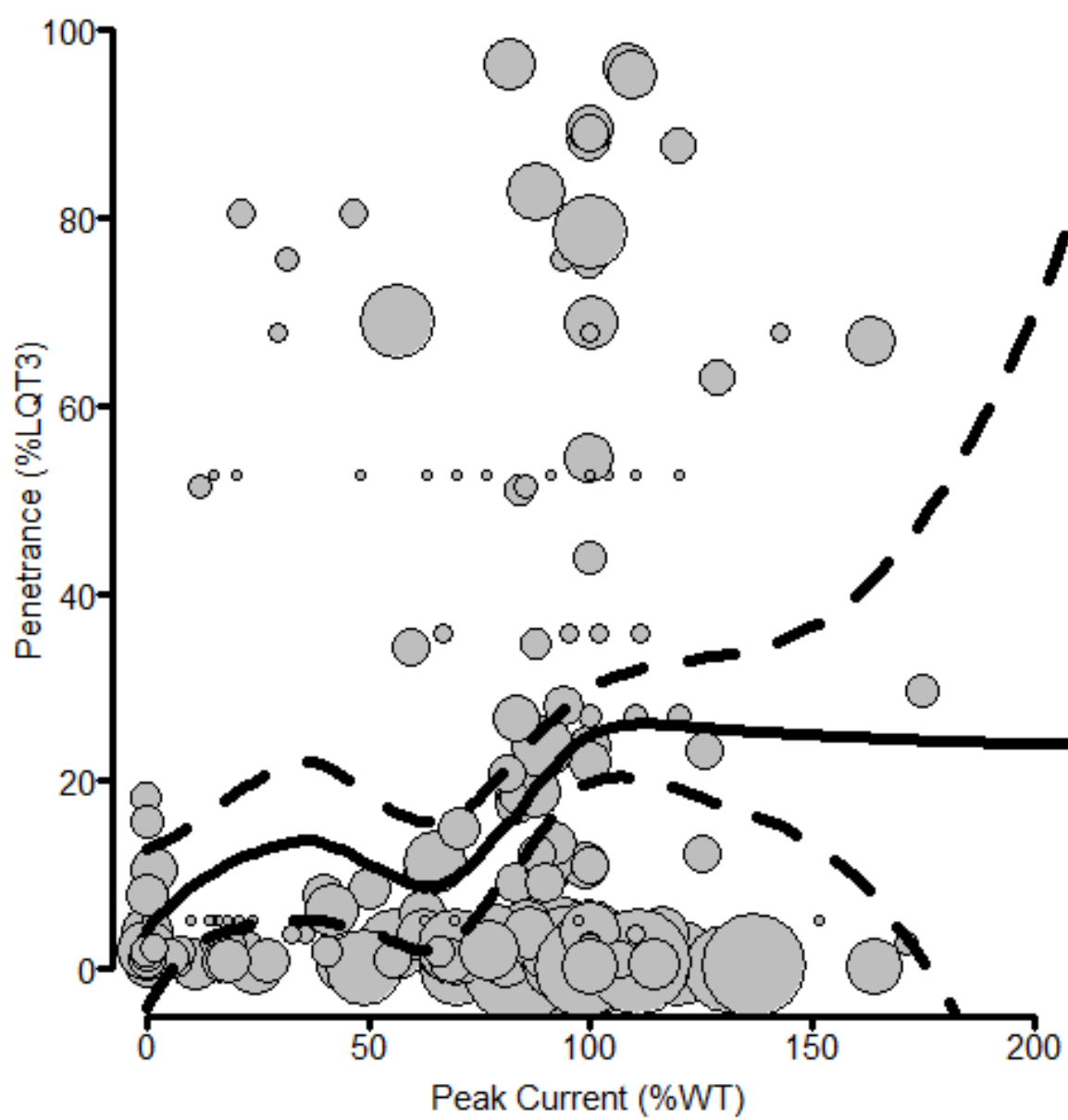
Functional Perturbation vs. Calculated Penetrance

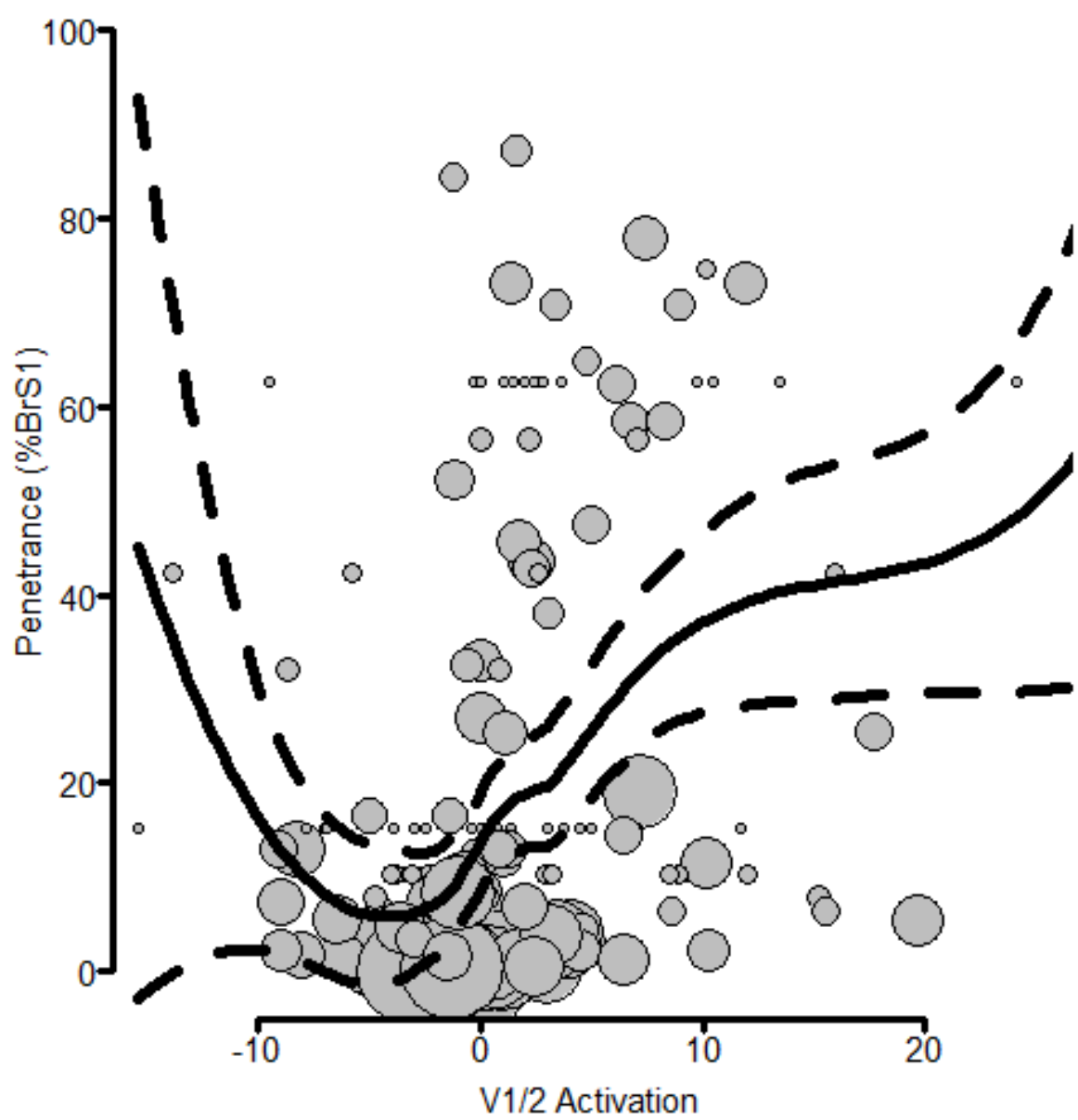
```
plt.disease("ipeak","Peak Current (%WT)", "BrS_penetranceBayesian")
plt.disease("ipeak","Peak Current (%WT)", "LQT_penetranceBayesian")
plt.disease("vhalfact","V1/2 Activation", "BrS_penetranceBayesian", x1 = c(-15,25))
plt.disease("vhalfact","V1/2 Activation", "LQT_penetranceBayesian", x1 = c(-15,25))
plt.disease("vhalfinact","V1/2 Inactivation", "BrS_penetranceBayesian", x1 = c(-25,20))
plt.disease("vhalfinact","V1/2 Inactivation", "LQT_penetranceBayesian", x1 = c(-25,20))
plt.disease("recovfrominact","Recovery from Inactivation (log[%WT])", "BrS_penetranceBayesian", x1 = c(
plt.disease("recovfrominact","Recovery from Inactivation (log[%WT])", "LQT_penetranceBayesian", x1 = c(
plt.disease("ilate","Late Current (log[%WT])", "BrS_penetranceBayesian", x1 = c(1.5,4))
plt.disease("ilate","Late Current (log[%WT])", "LQT_penetranceBayesian", x1 = c(1.5,4))
plt.disease("ilate_norm","Late Current (log[%WT]) Normalized", "BrS_penetranceBayesian", x1 = c(1.5,4))
plt.disease("ilate_norm","Late Current (log[%WT]) Normalized", "LQT_penetranceBayesian", x1 = c(1.5,4))

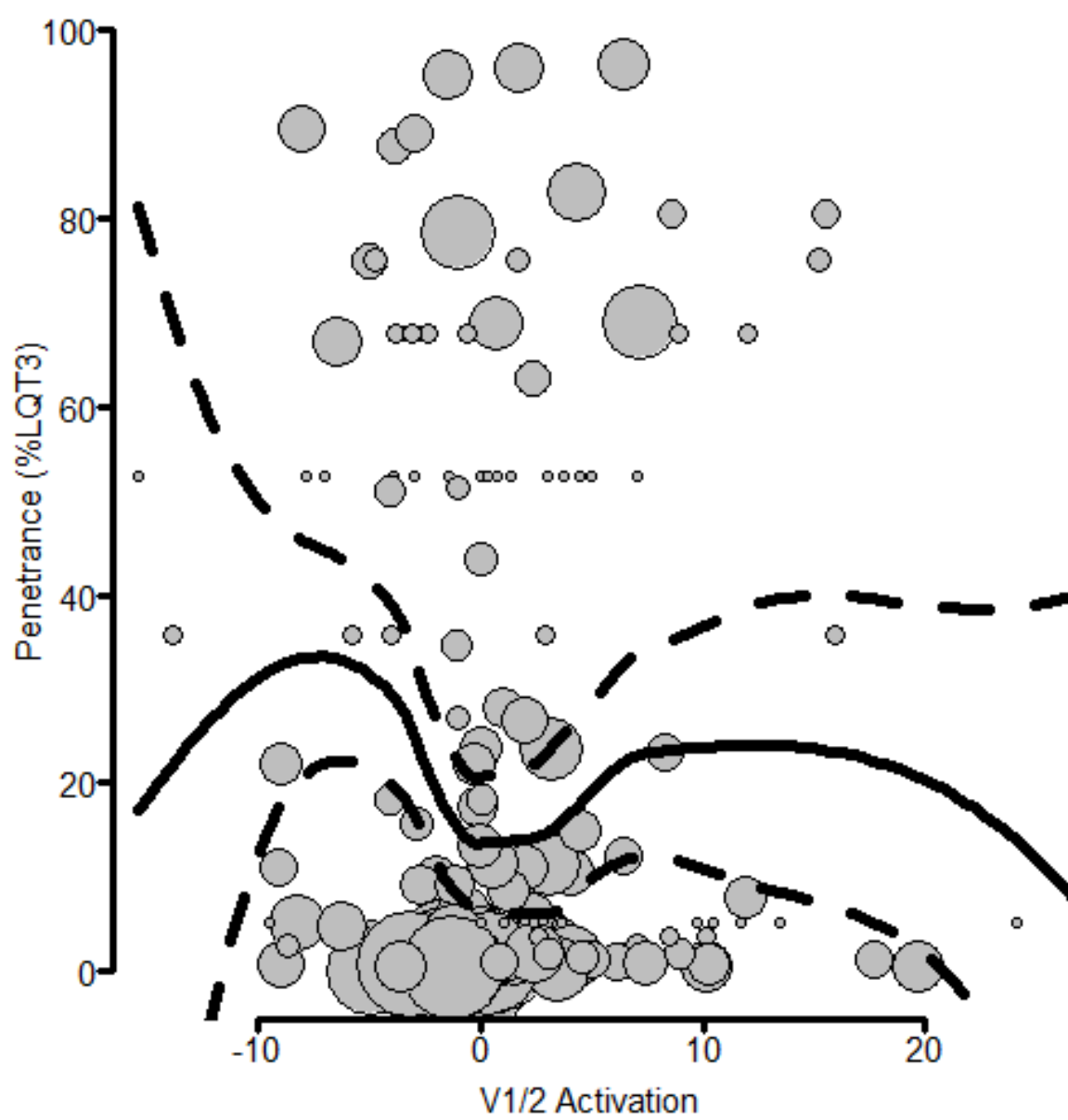
plt.disease("ipeak","Peak Current (%WT)", "all_penetranceBayesian")

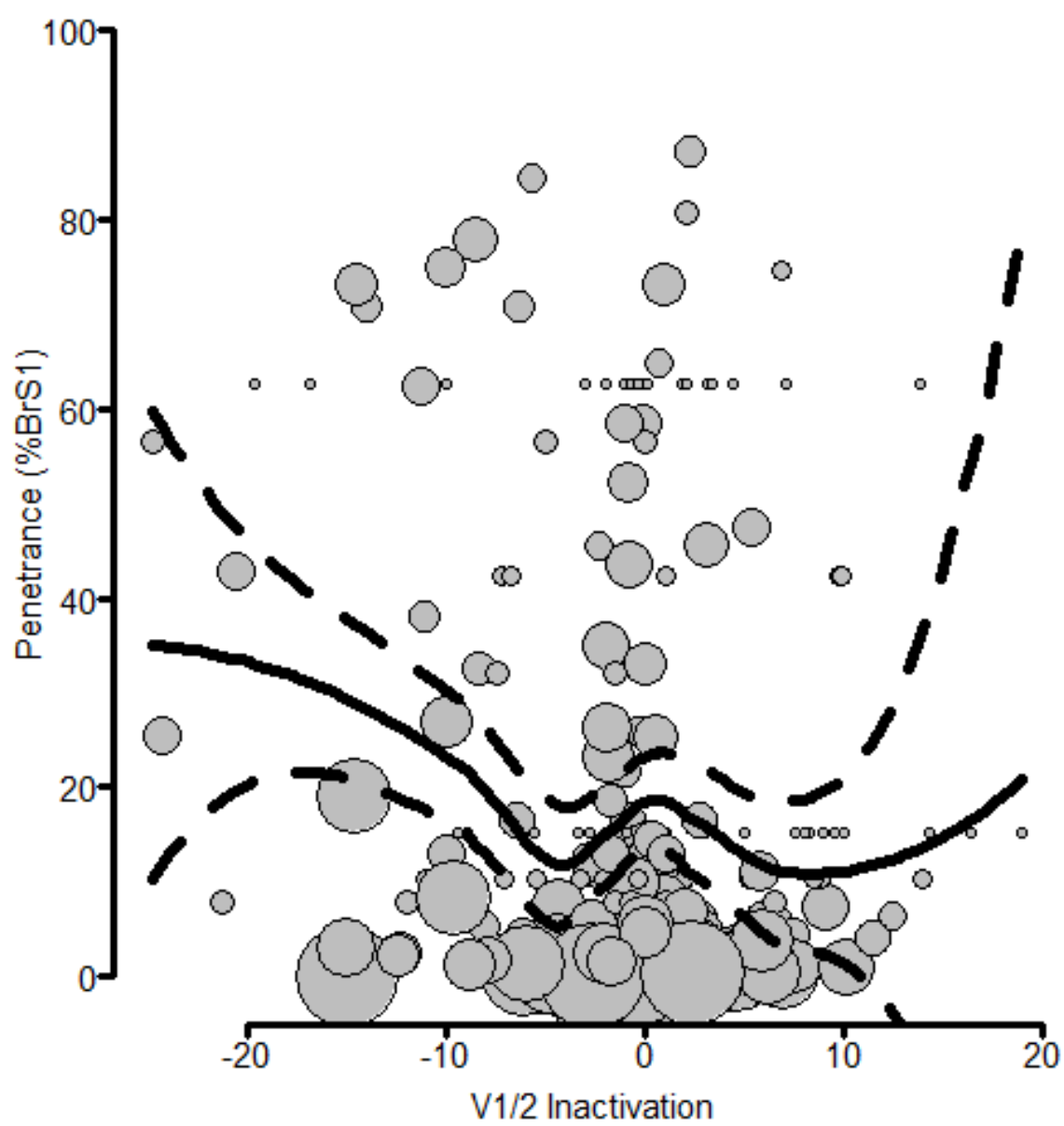
plt.disease('eaRate','Evolution Rate', 'all_penetranceBayesian', x1=c(0,4))
plt.disease('pph2_prob','PolyPhen-2', 'all_penetranceBayesian', x1=c(0,1))
plt.disease('SIFT.Score','SIFT', 'all_penetranceBayesian', x1=c(0,1))
plt.disease('provean_score','PROVEAN', 'all_penetranceBayesian', x1=c(-13,5))
plt.disease('CADD_RAW','CADD', 'all_penetranceBayesian', x1=c(0,8))
plt.disease('Condel_n','ConDel', 'all_penetranceBayesian', x1=c(0,1))
plt.disease('blastpssm','BLAST-PSSM', 'all_penetranceBayesian', x1=c(-7,7))
plt.disease('pamscore','PAM Score', 'all_penetranceBayesian', x1=c(-15,2))
plt.disease('aasimilaritymat','AA Similarity Matrix', 'all_penetranceBayesian', x1=c(-0.5,0.6))
```

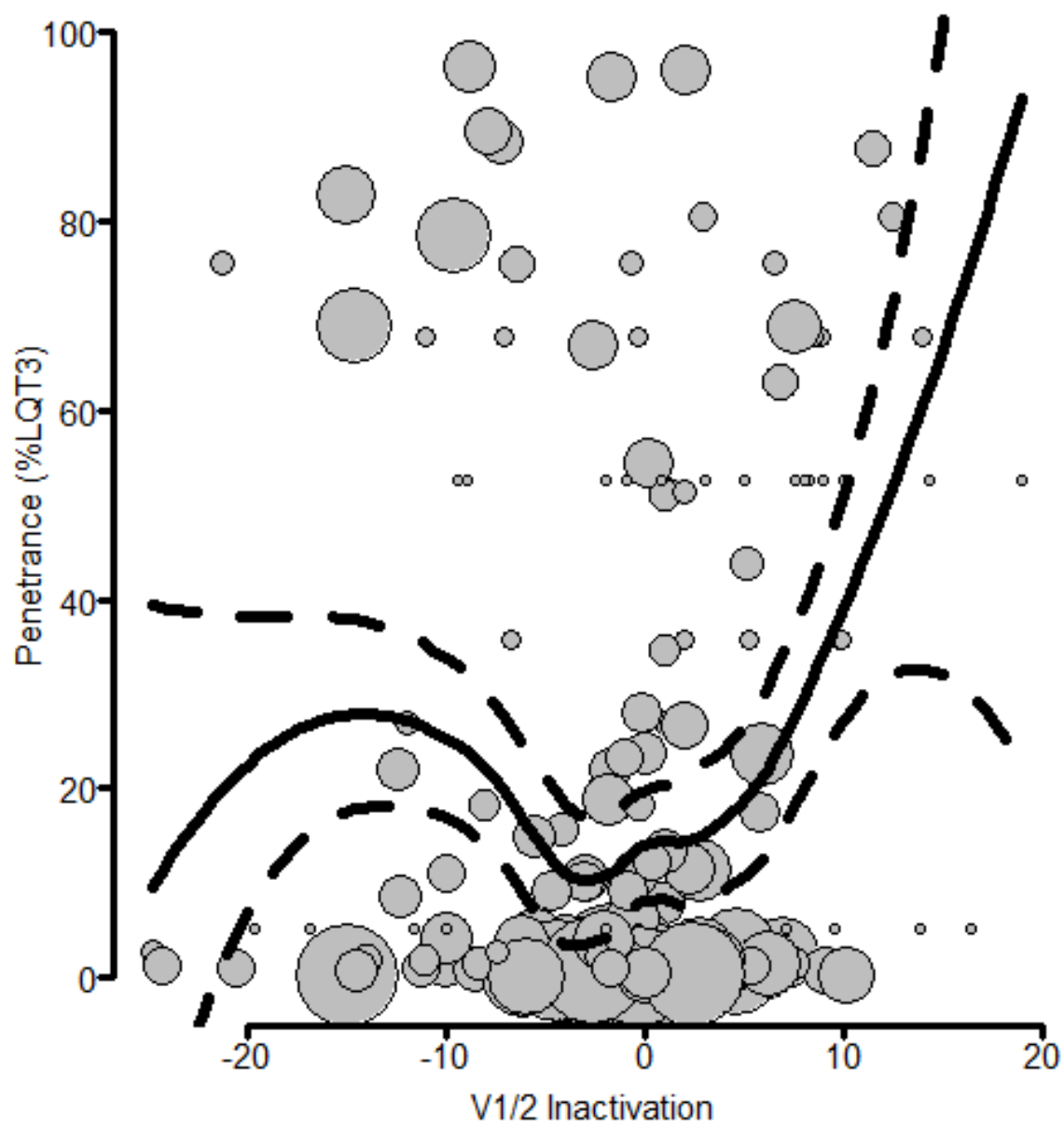


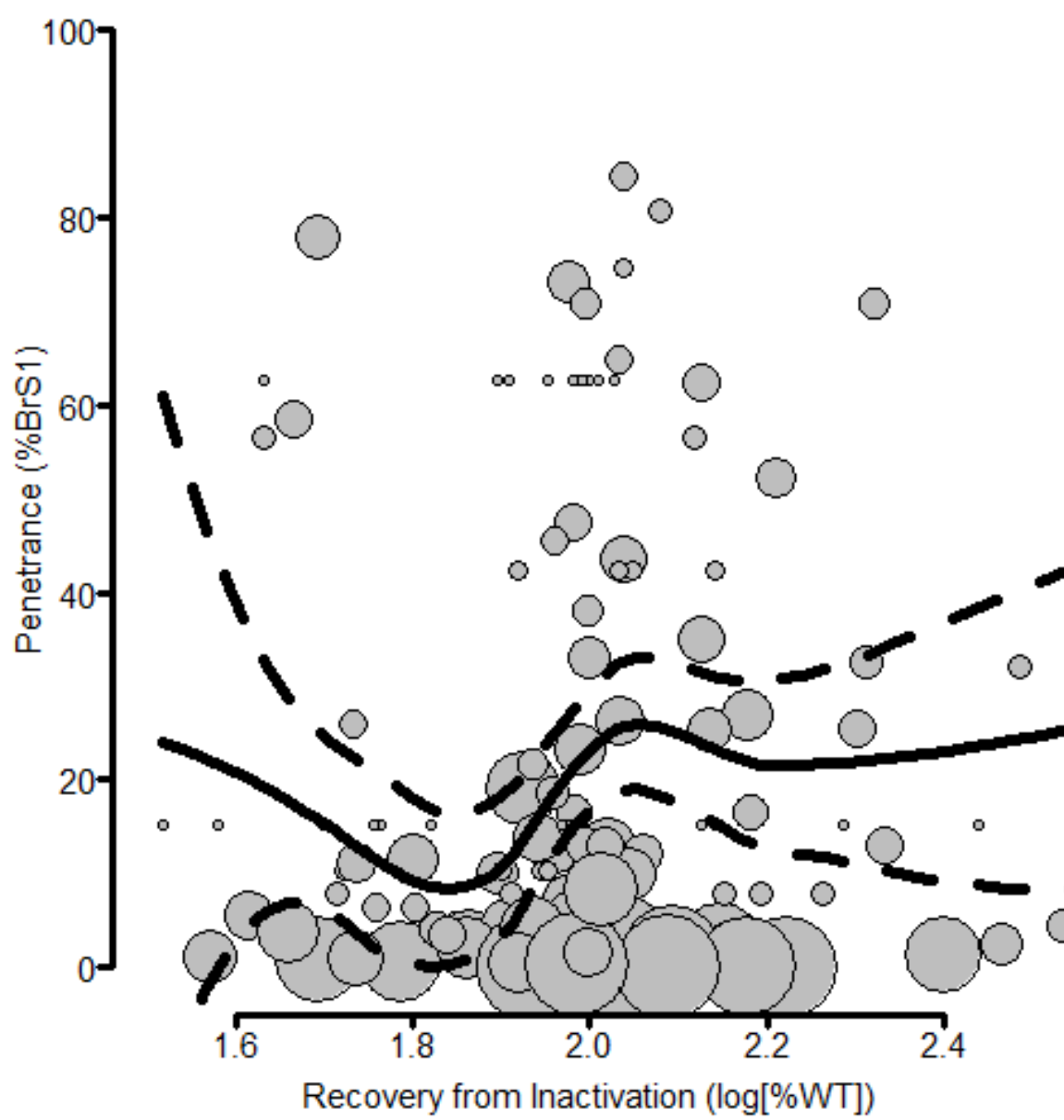


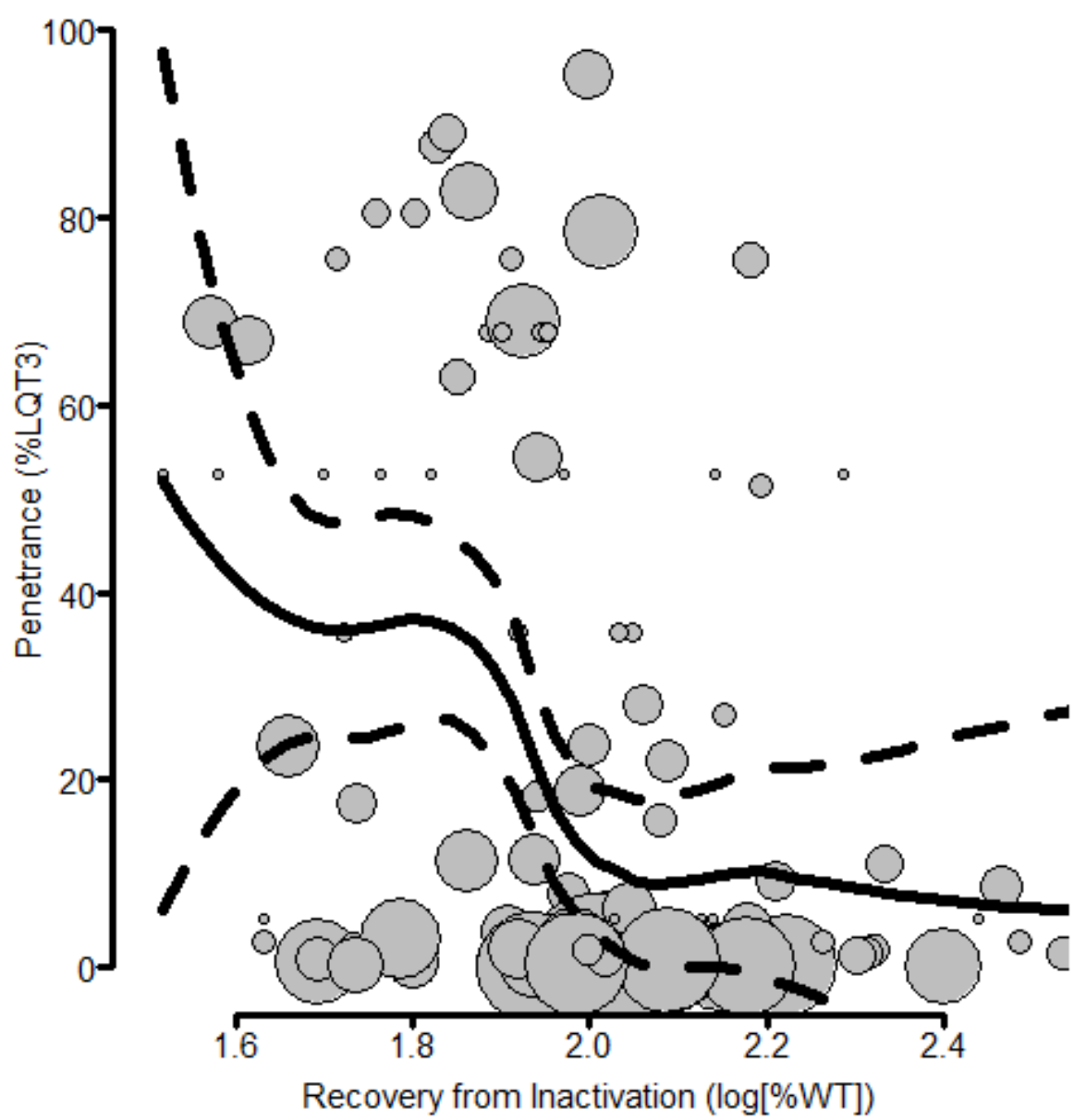


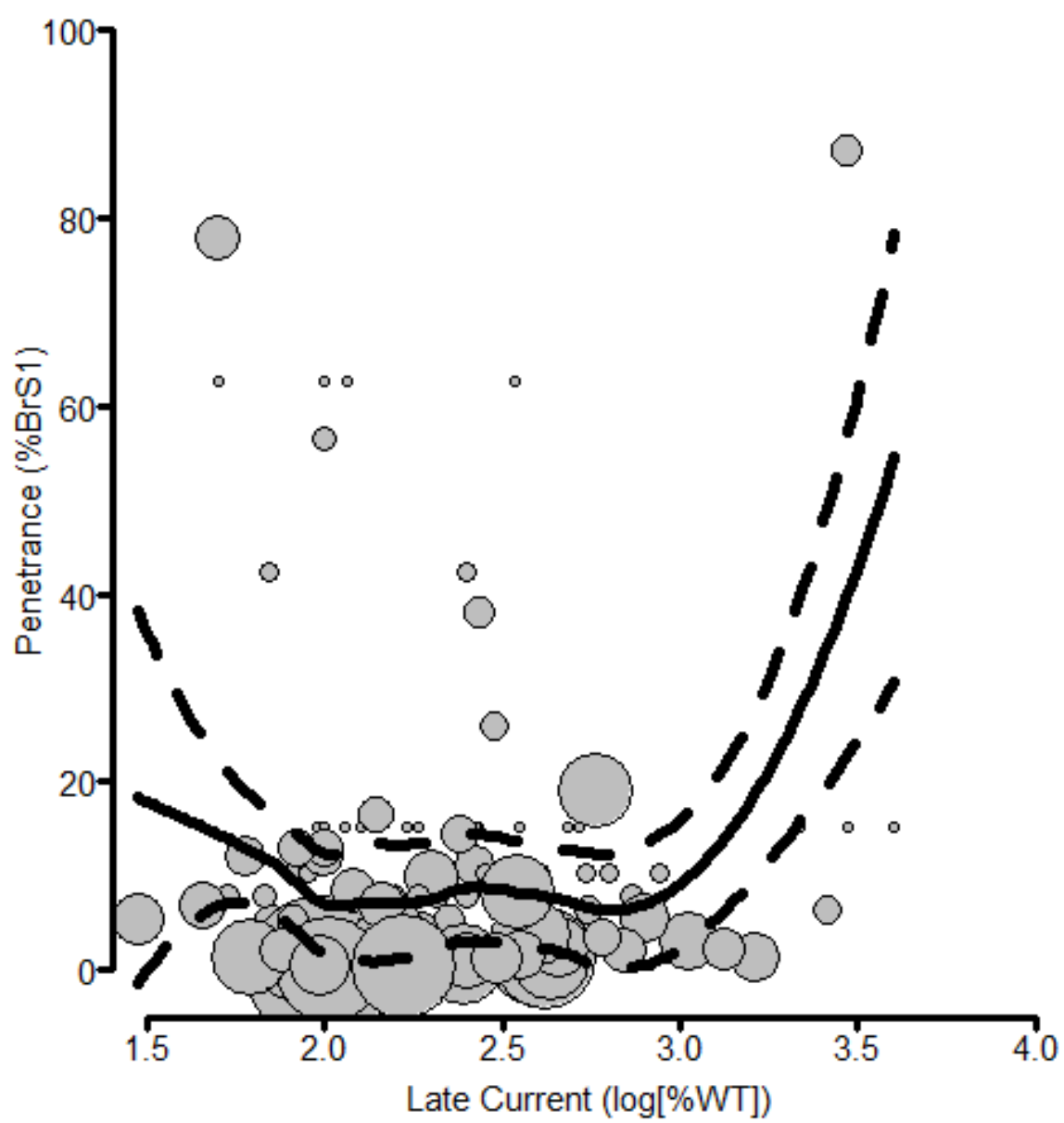


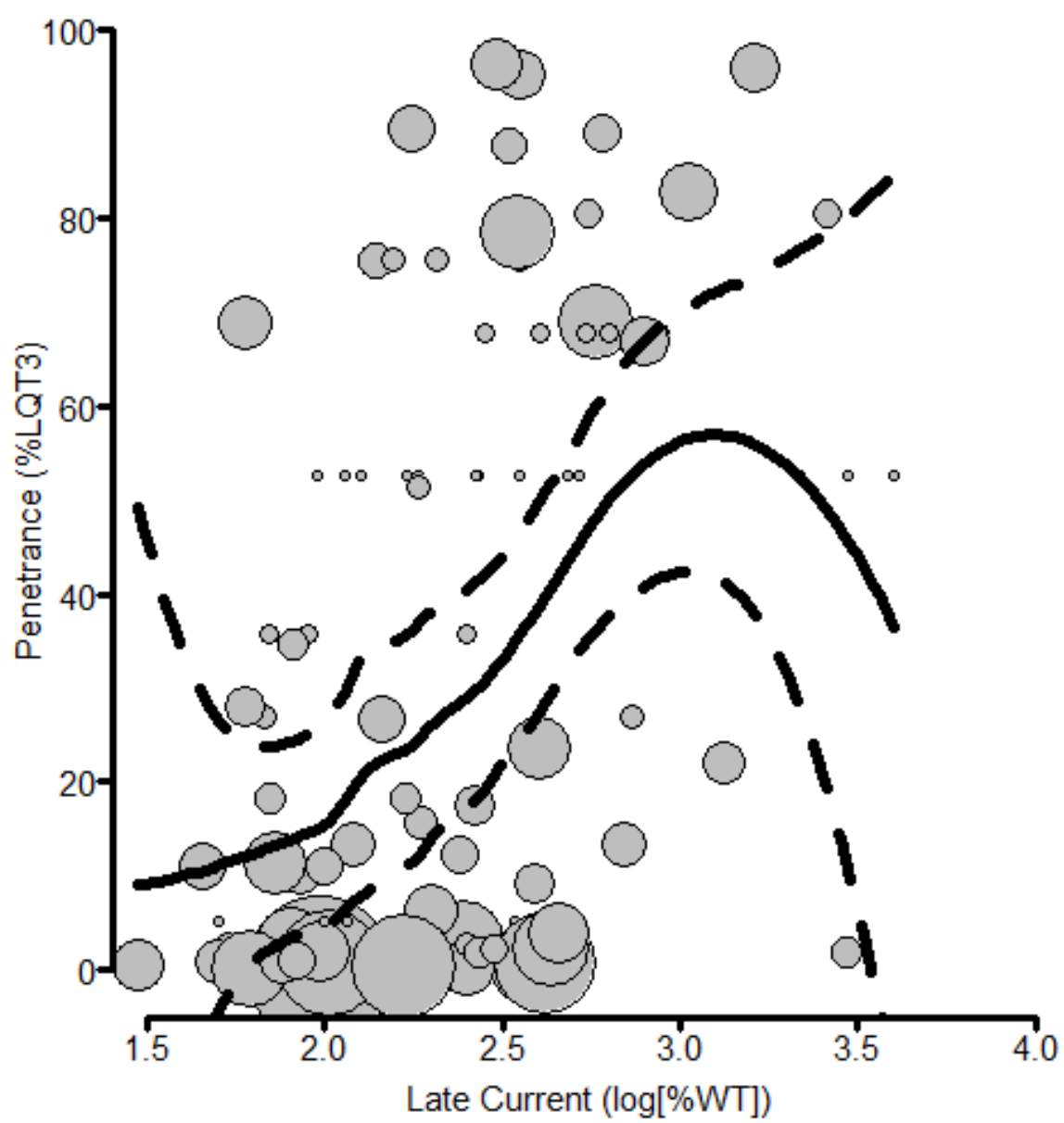


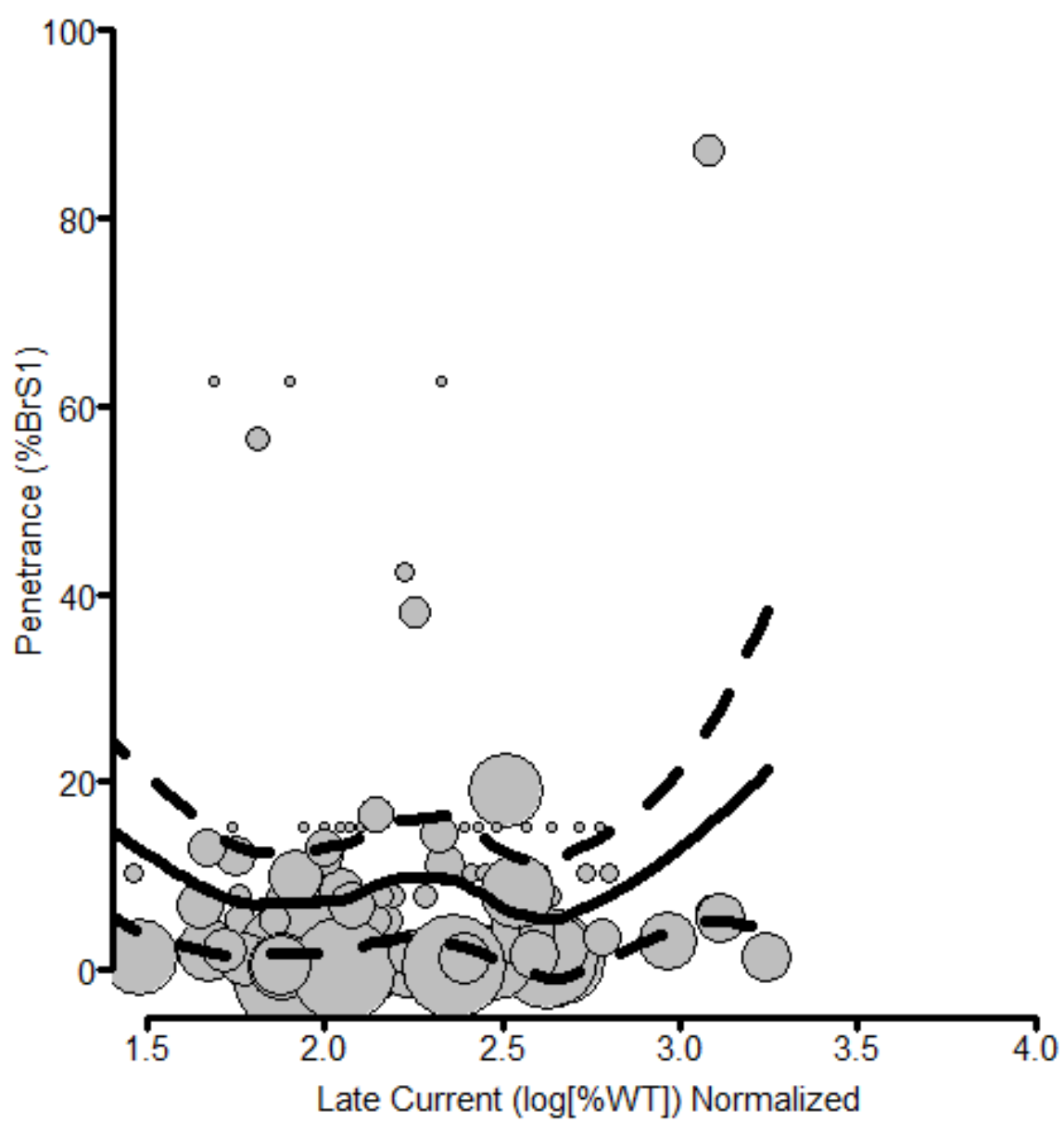


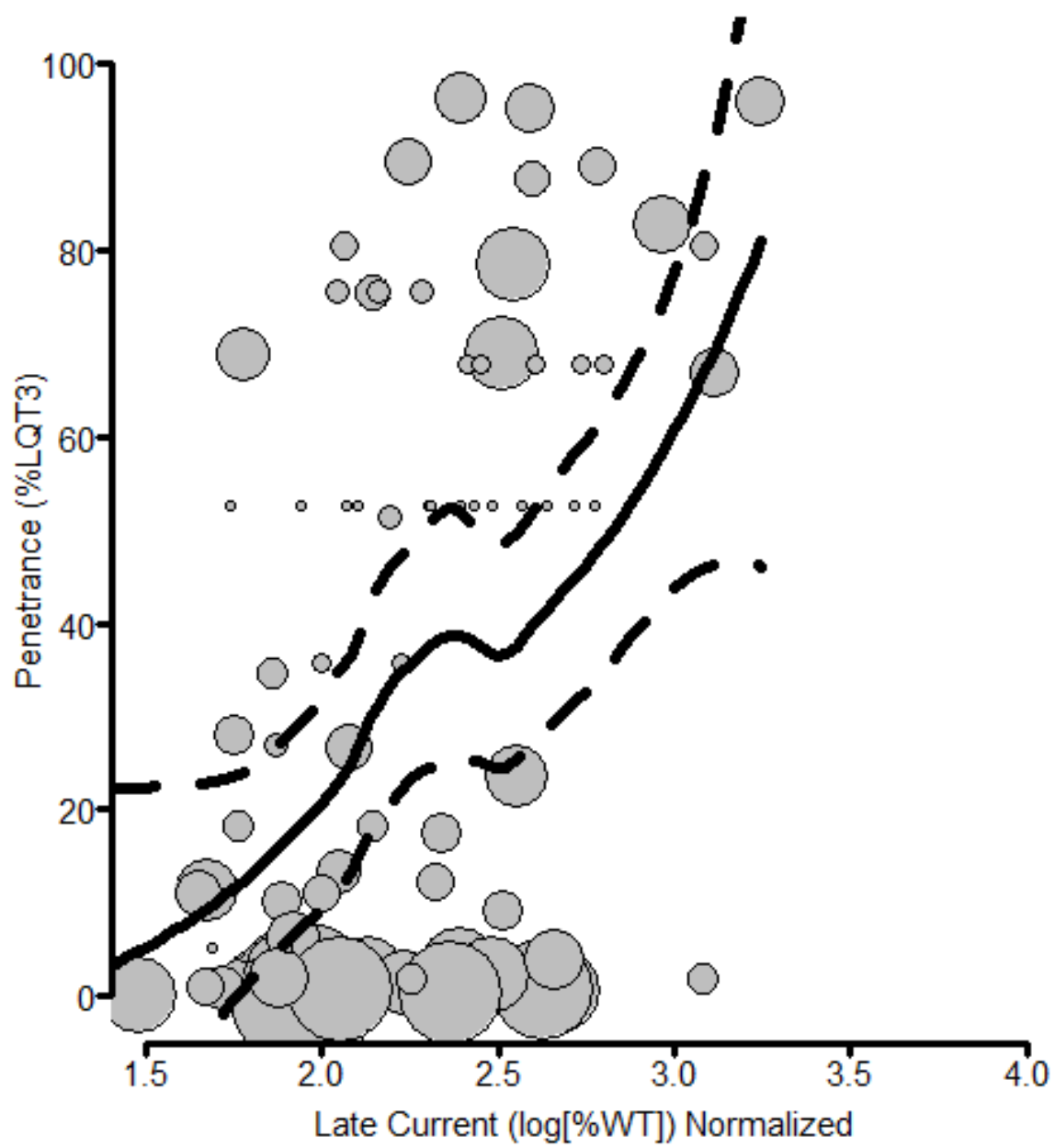


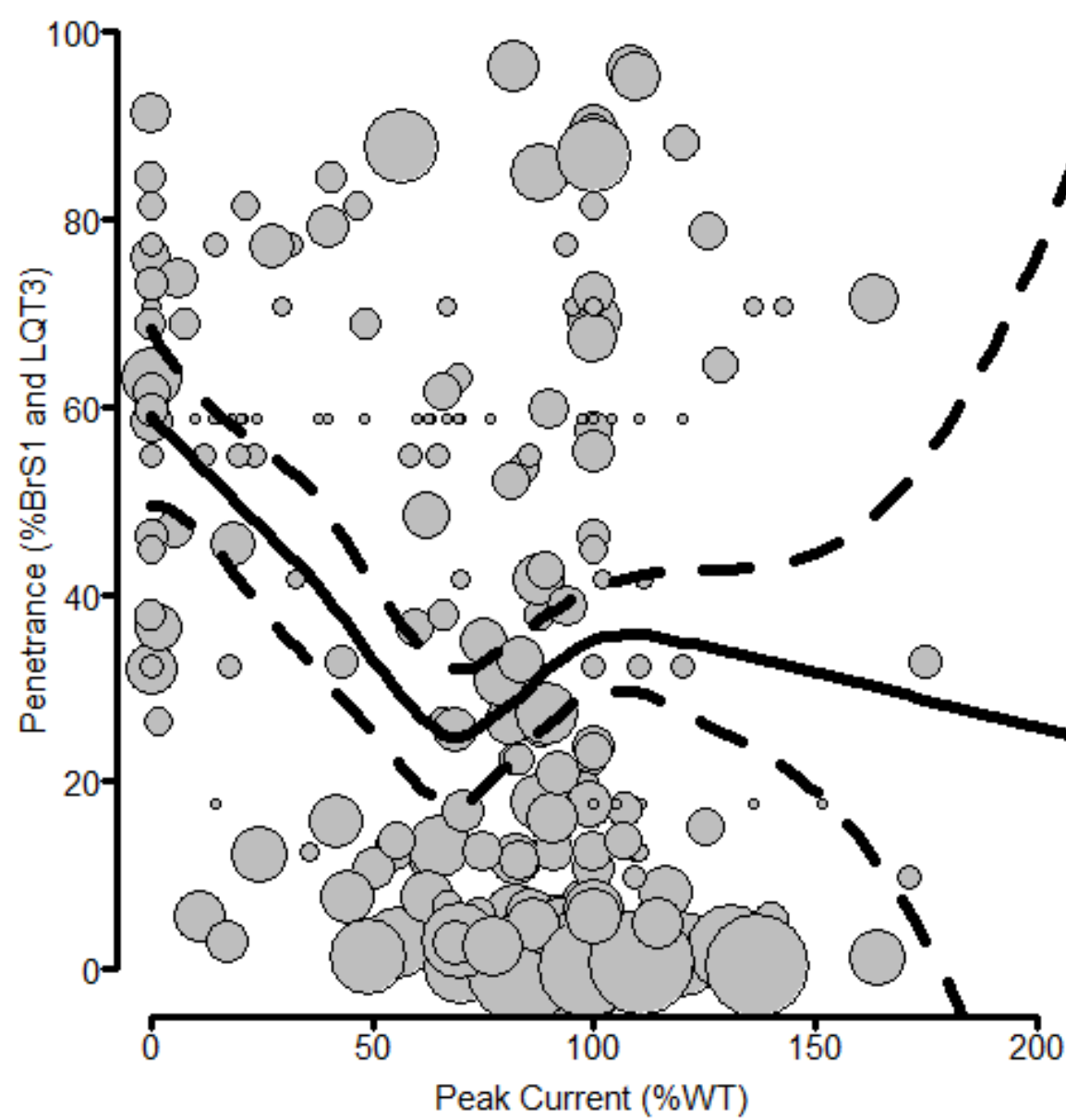


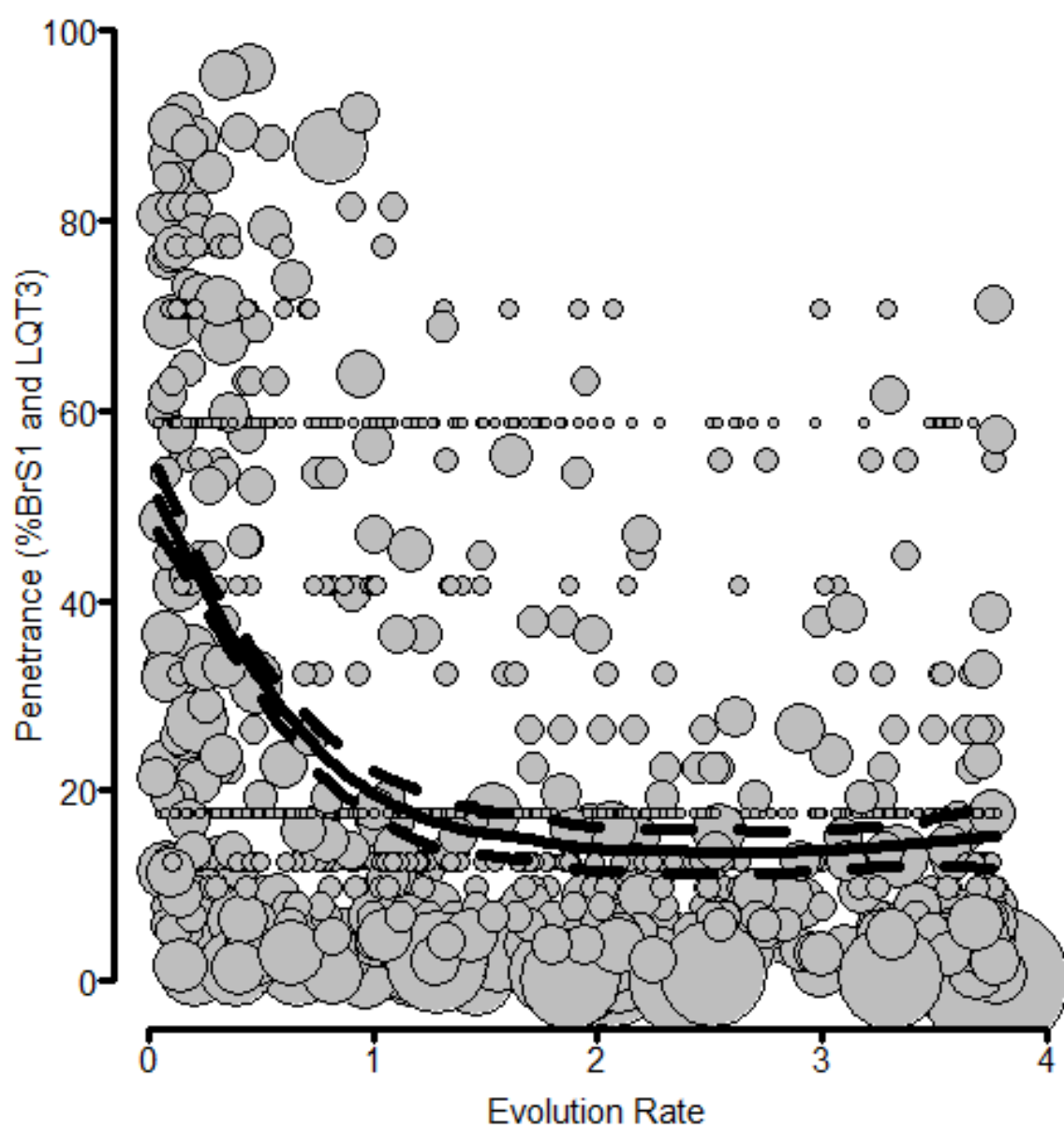


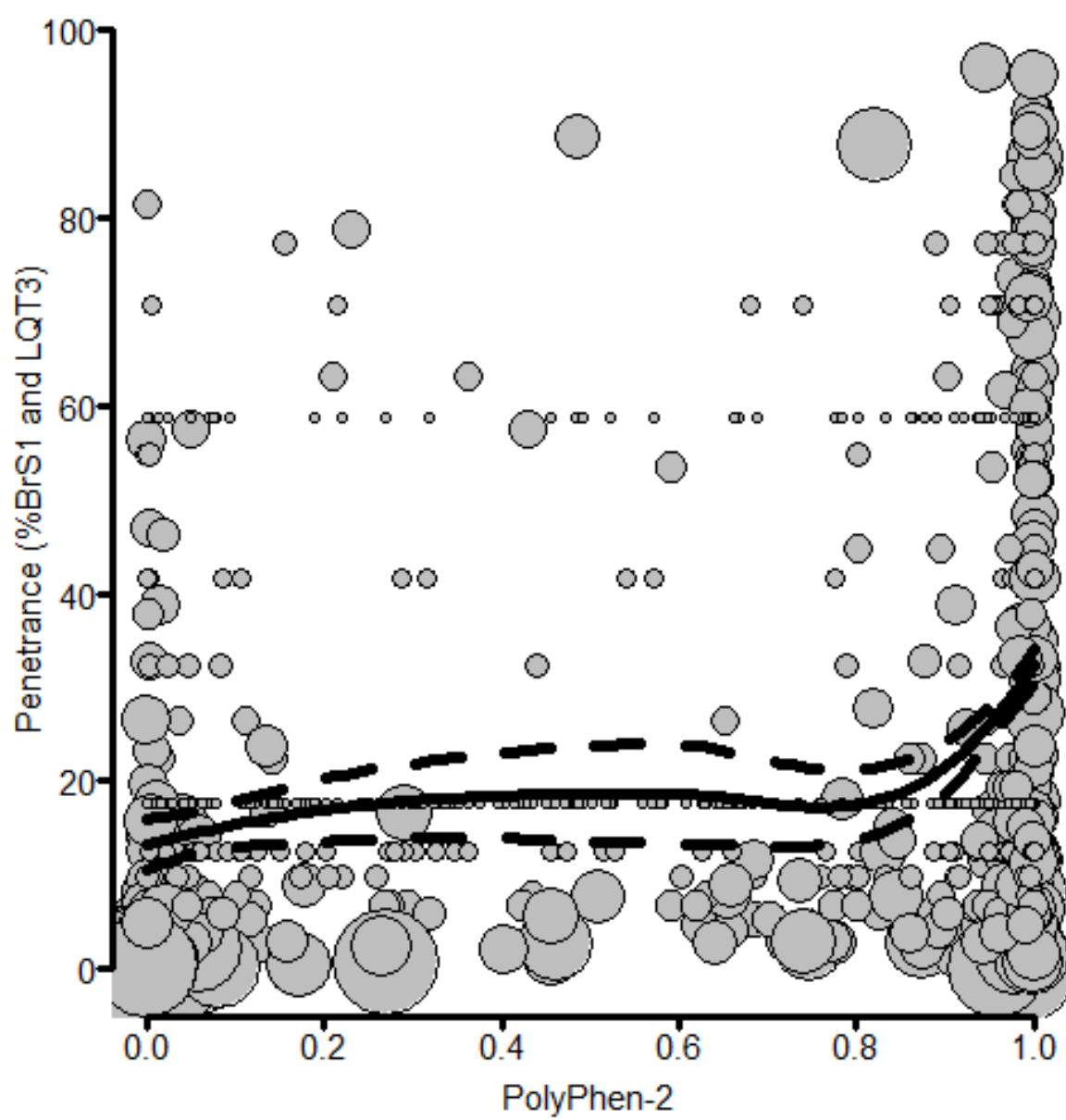


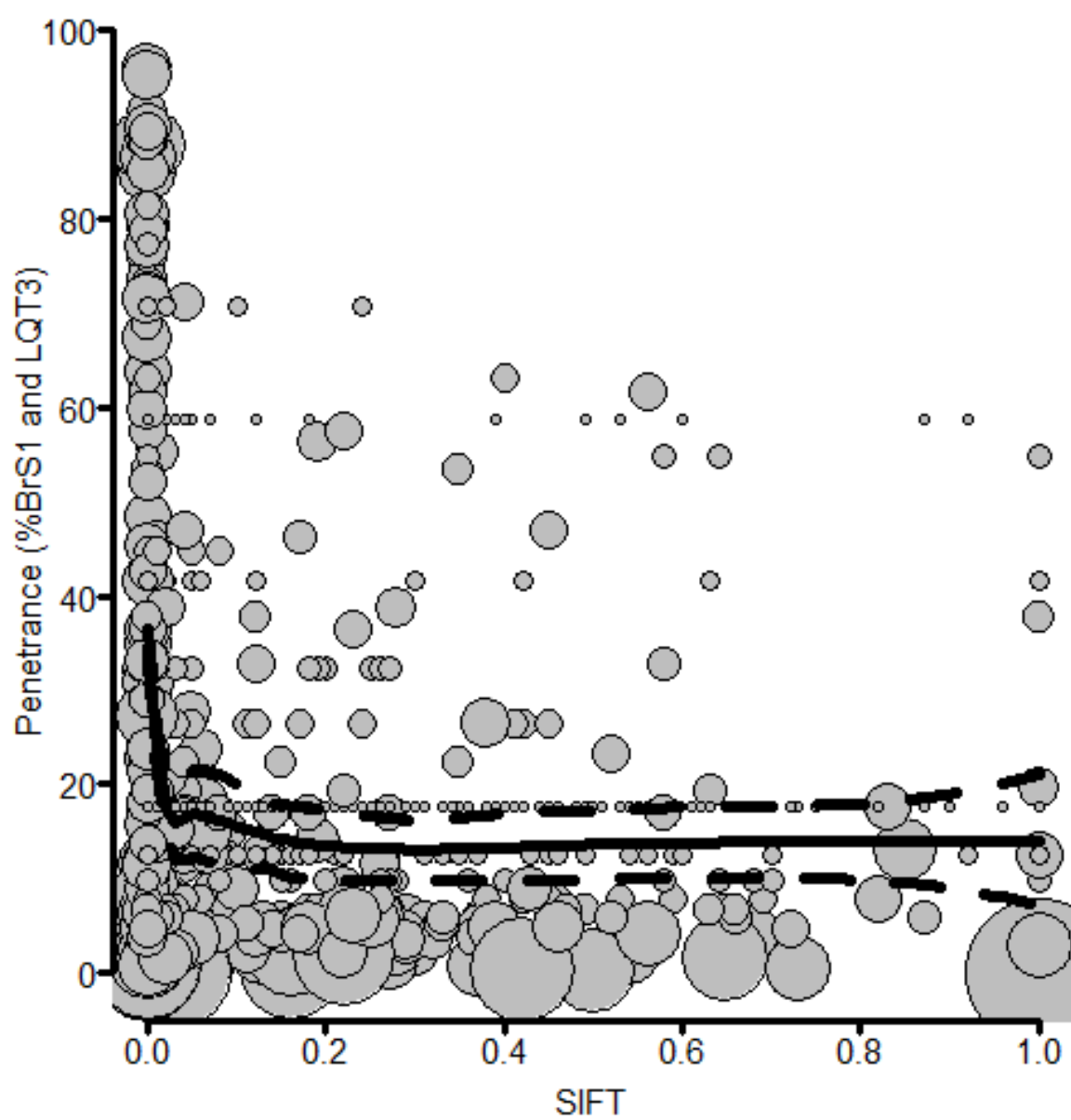


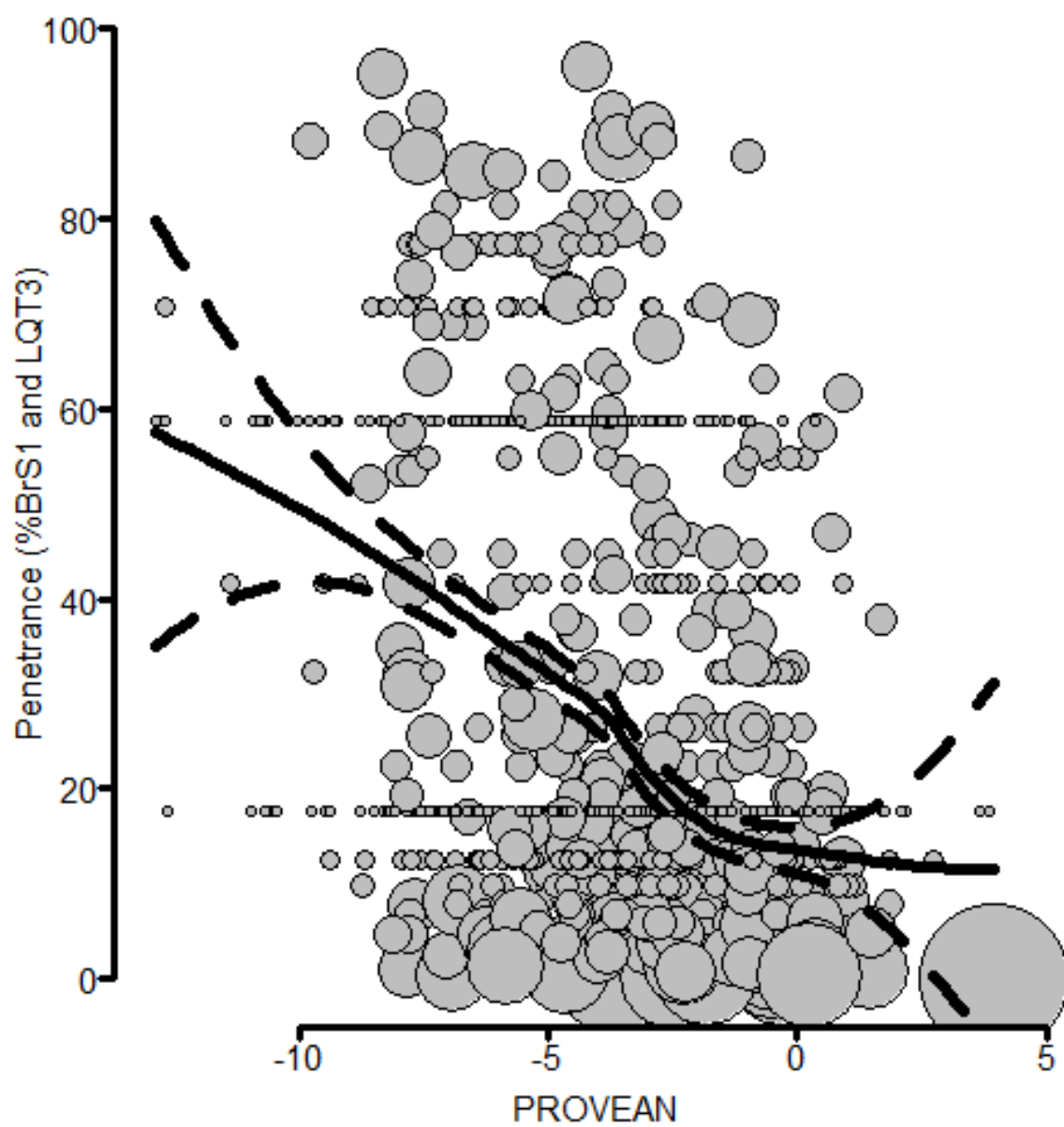


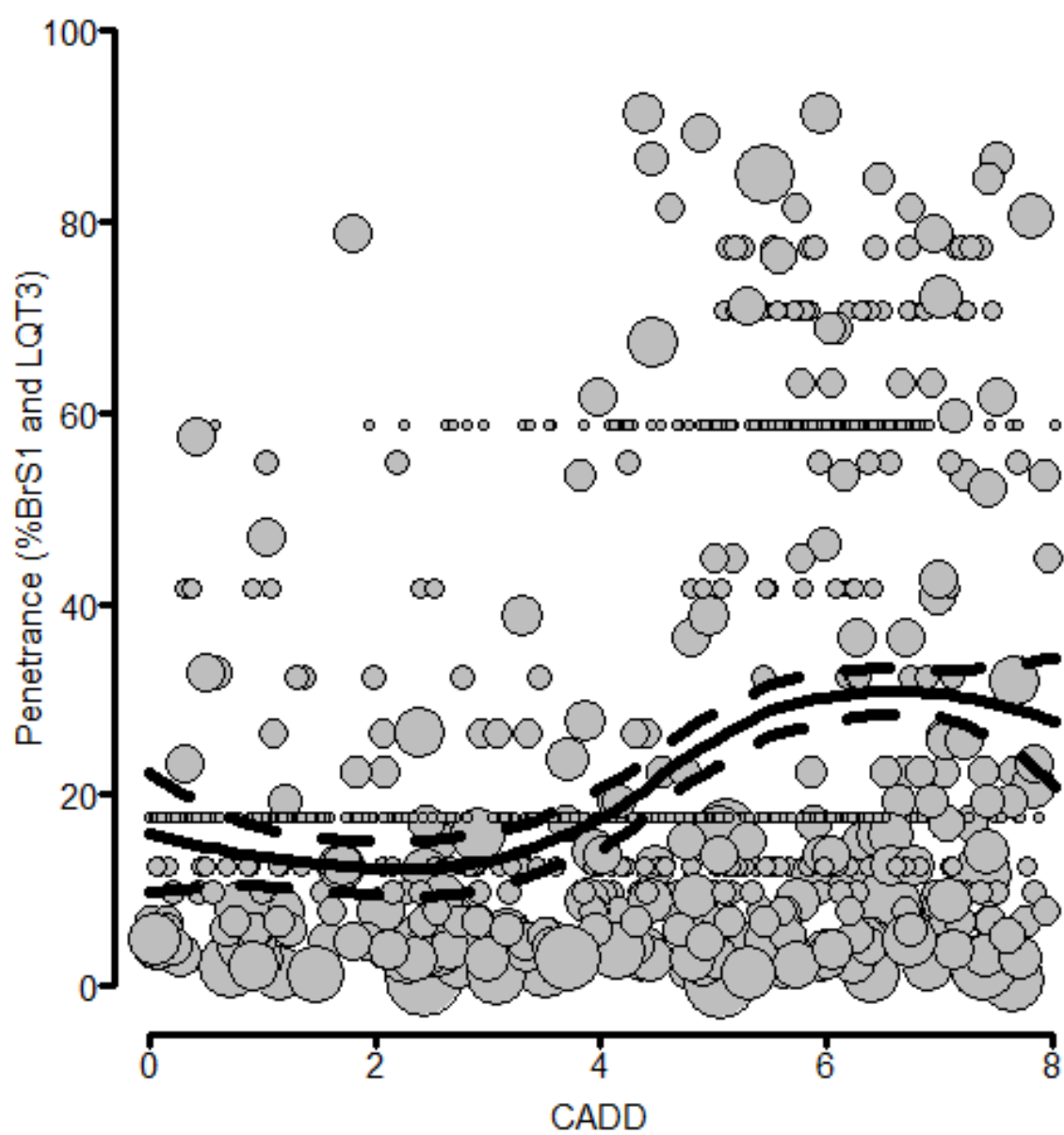


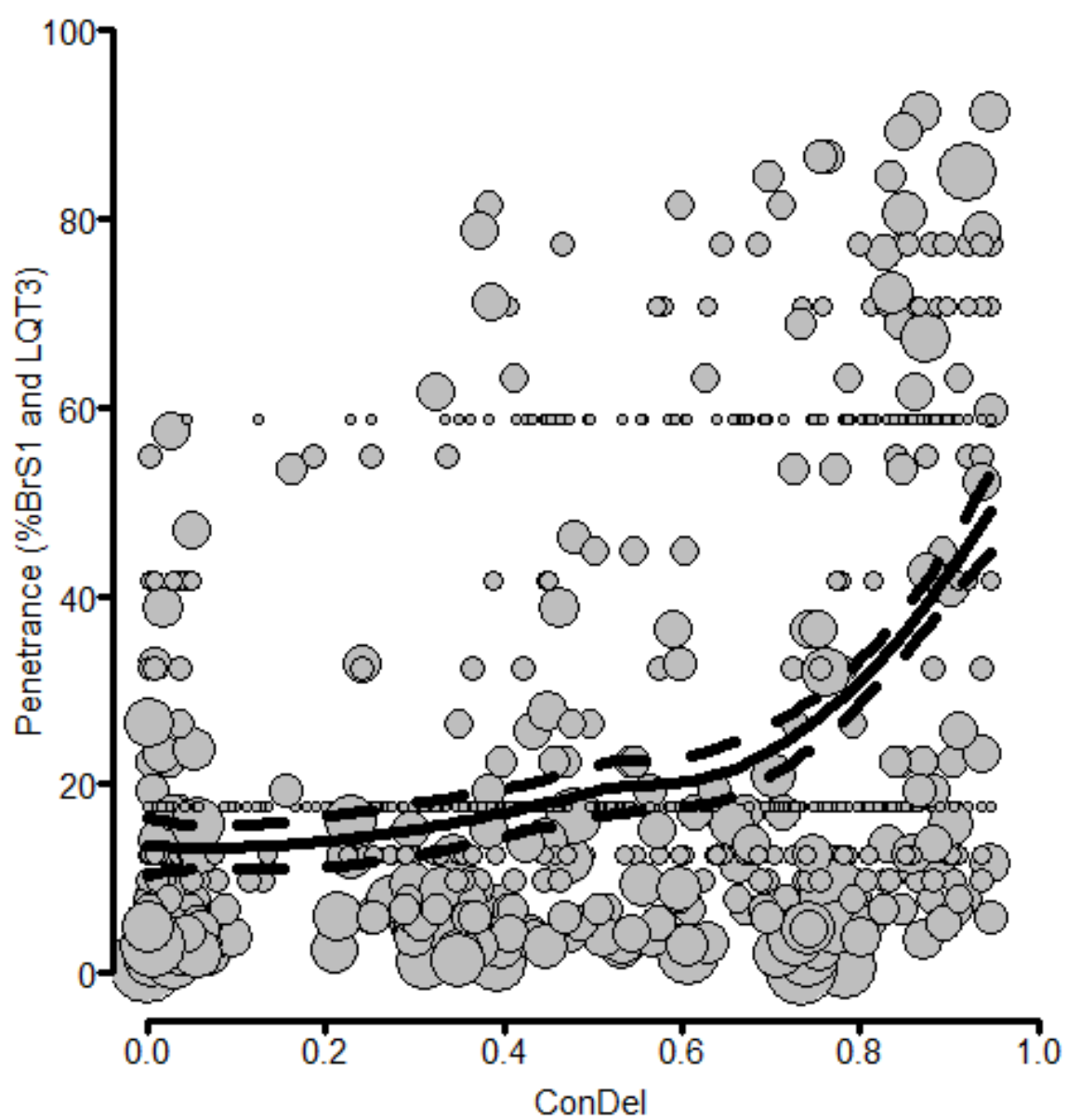


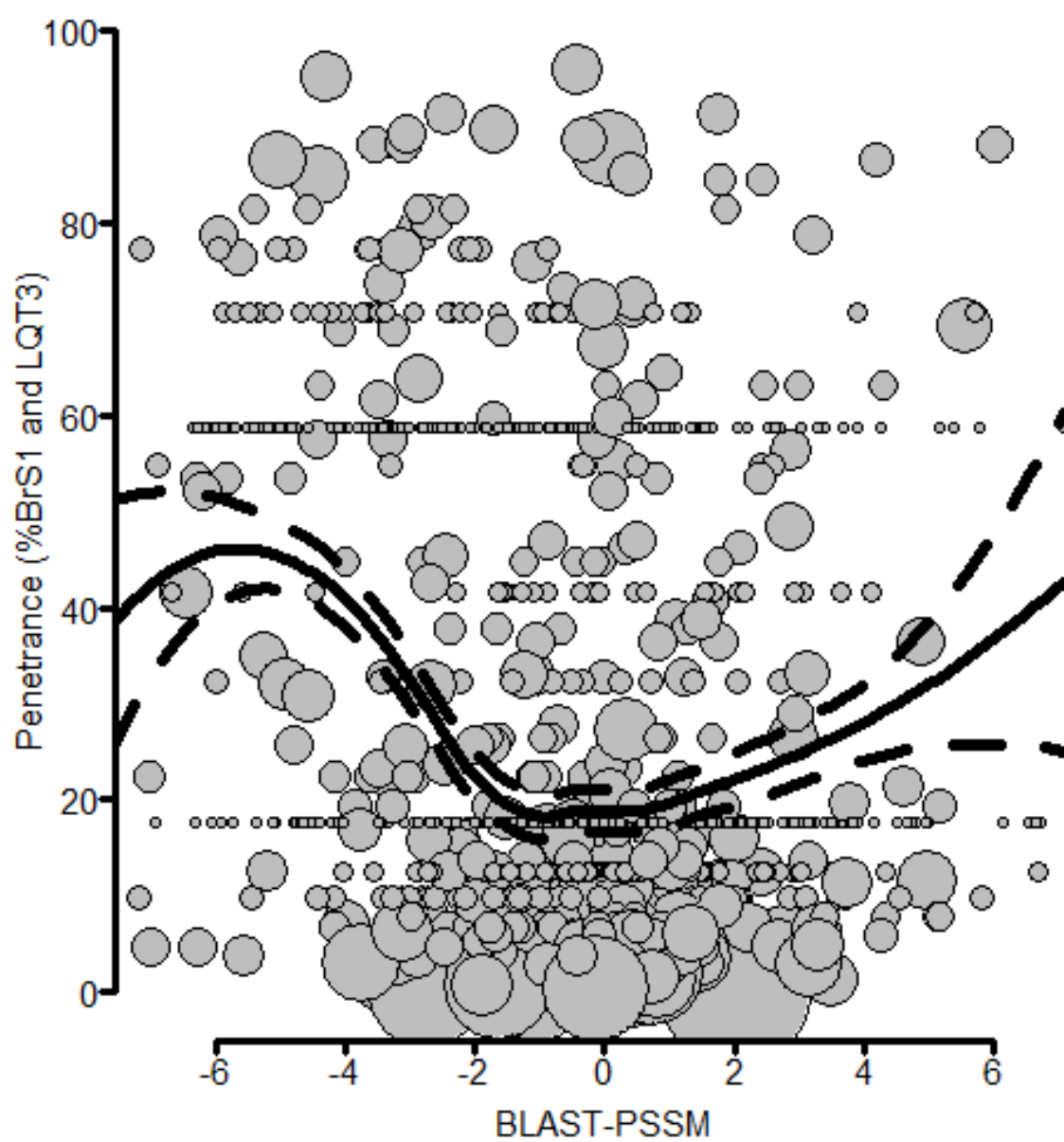


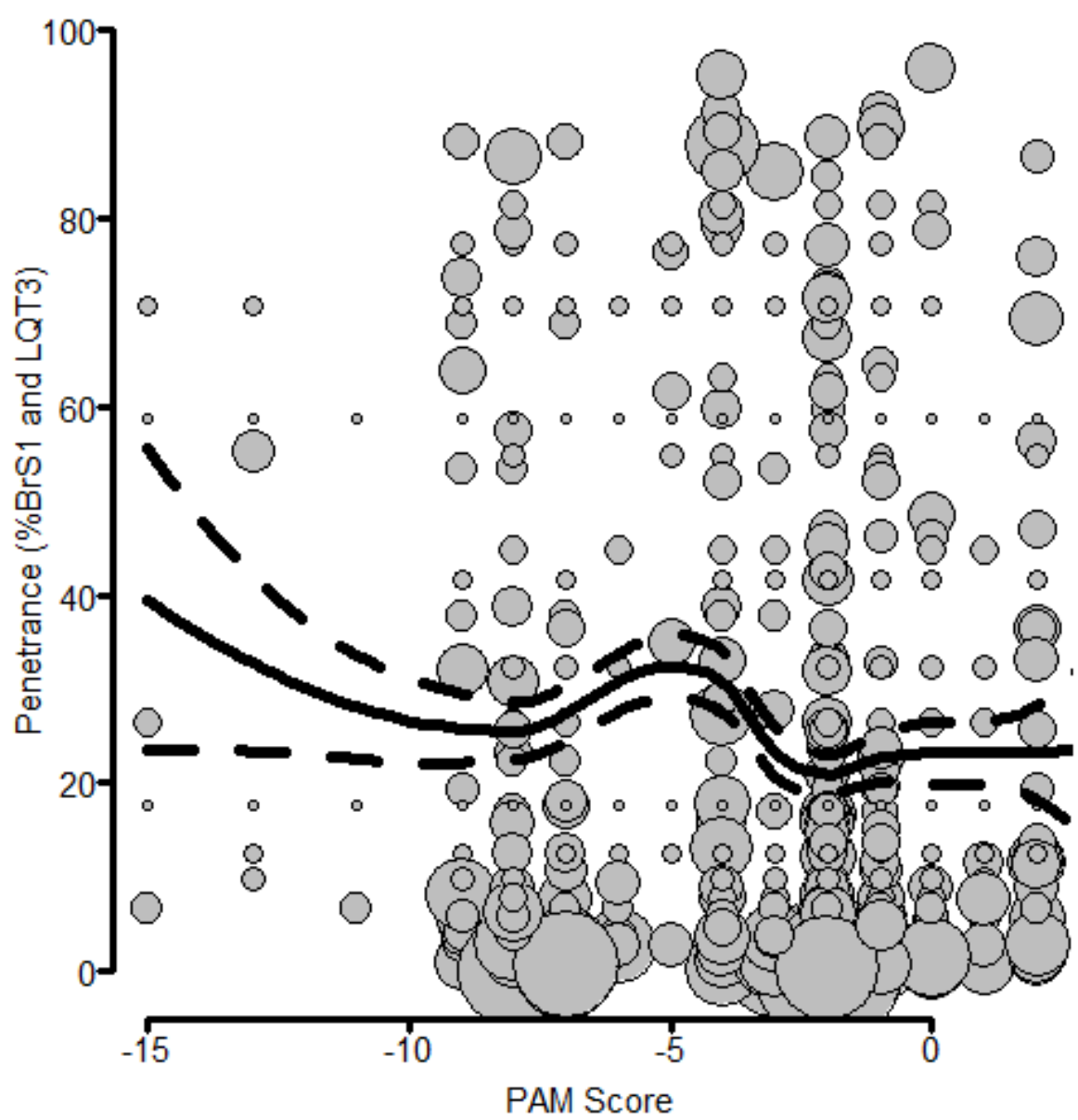


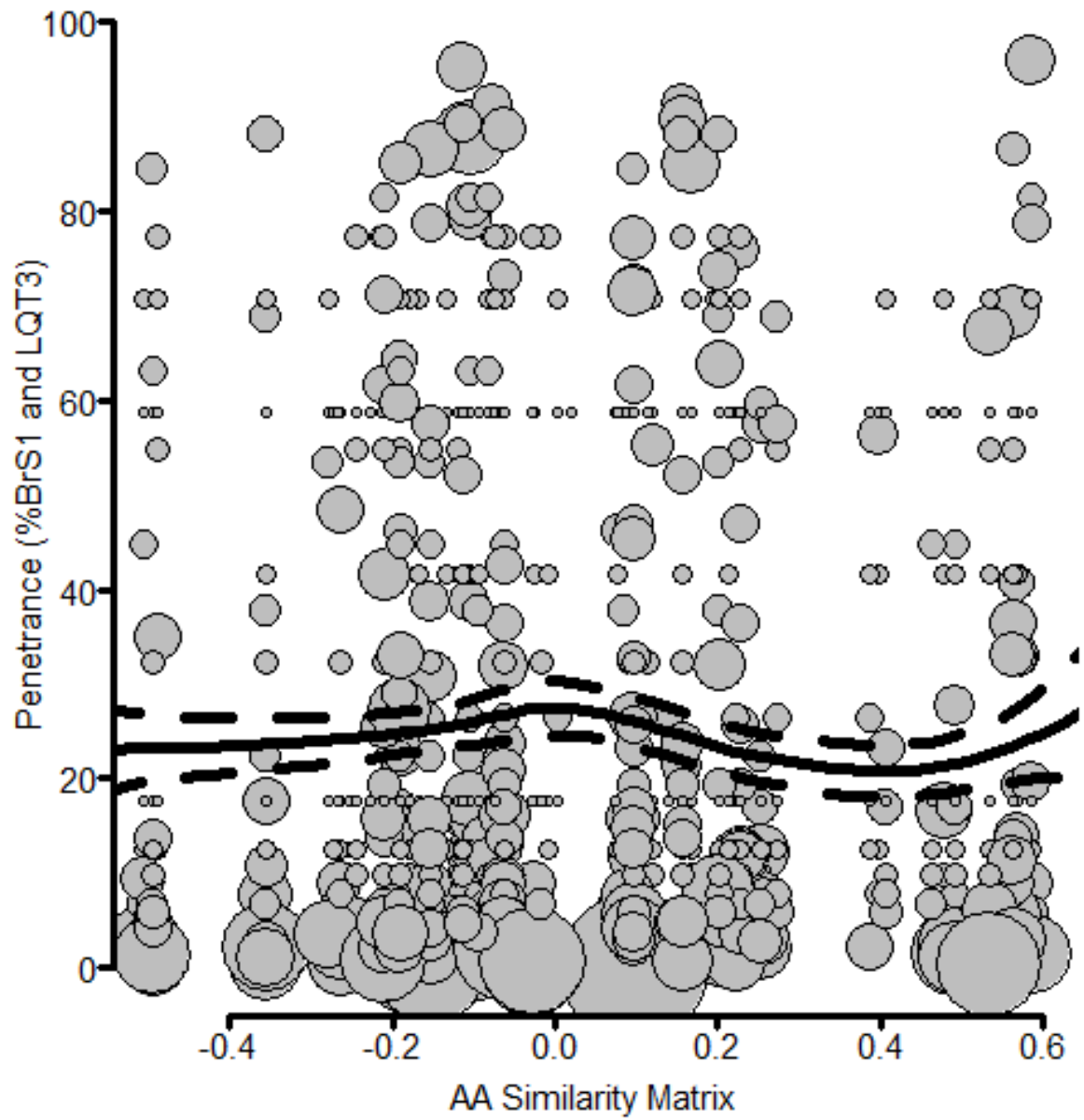










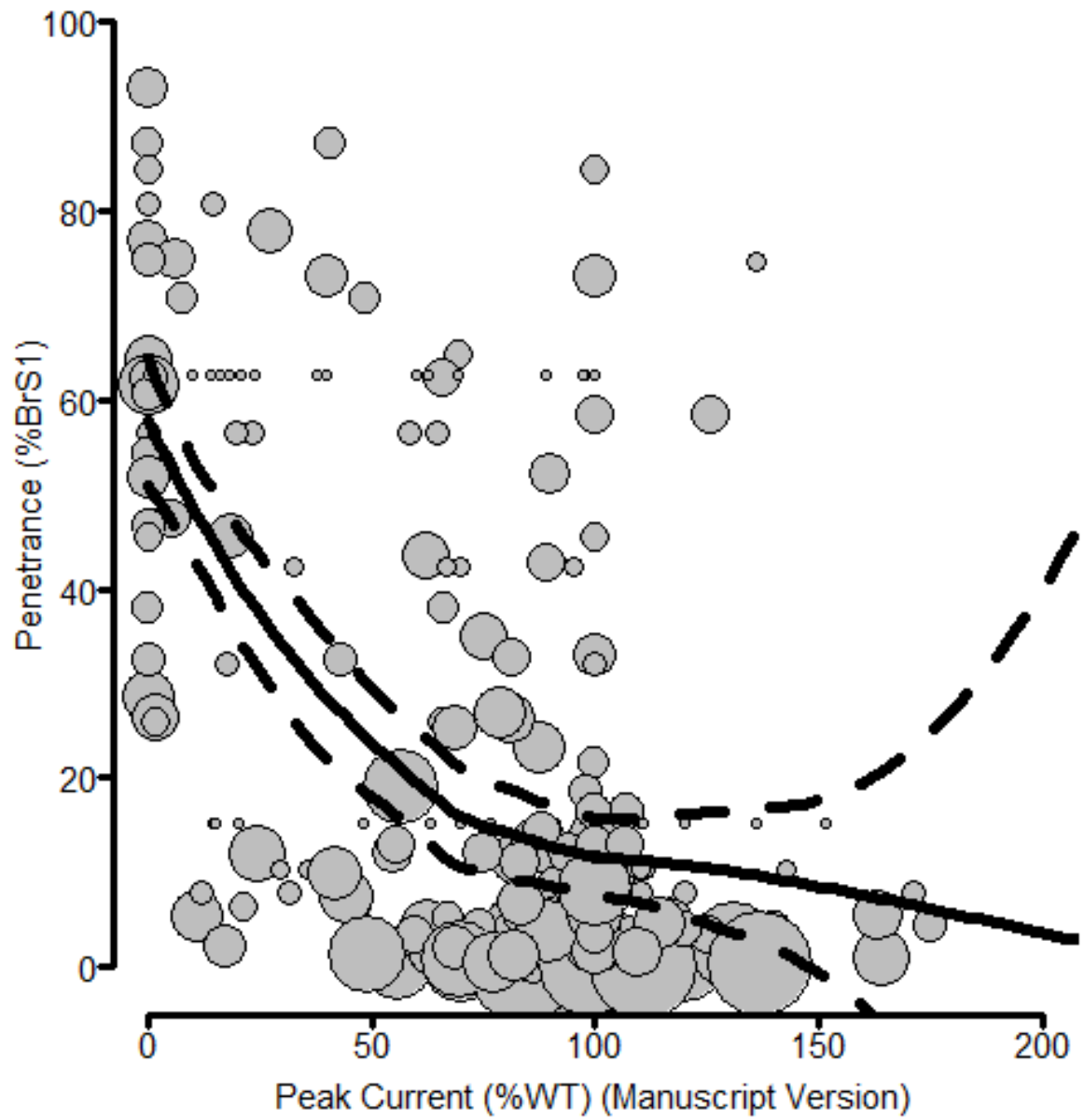


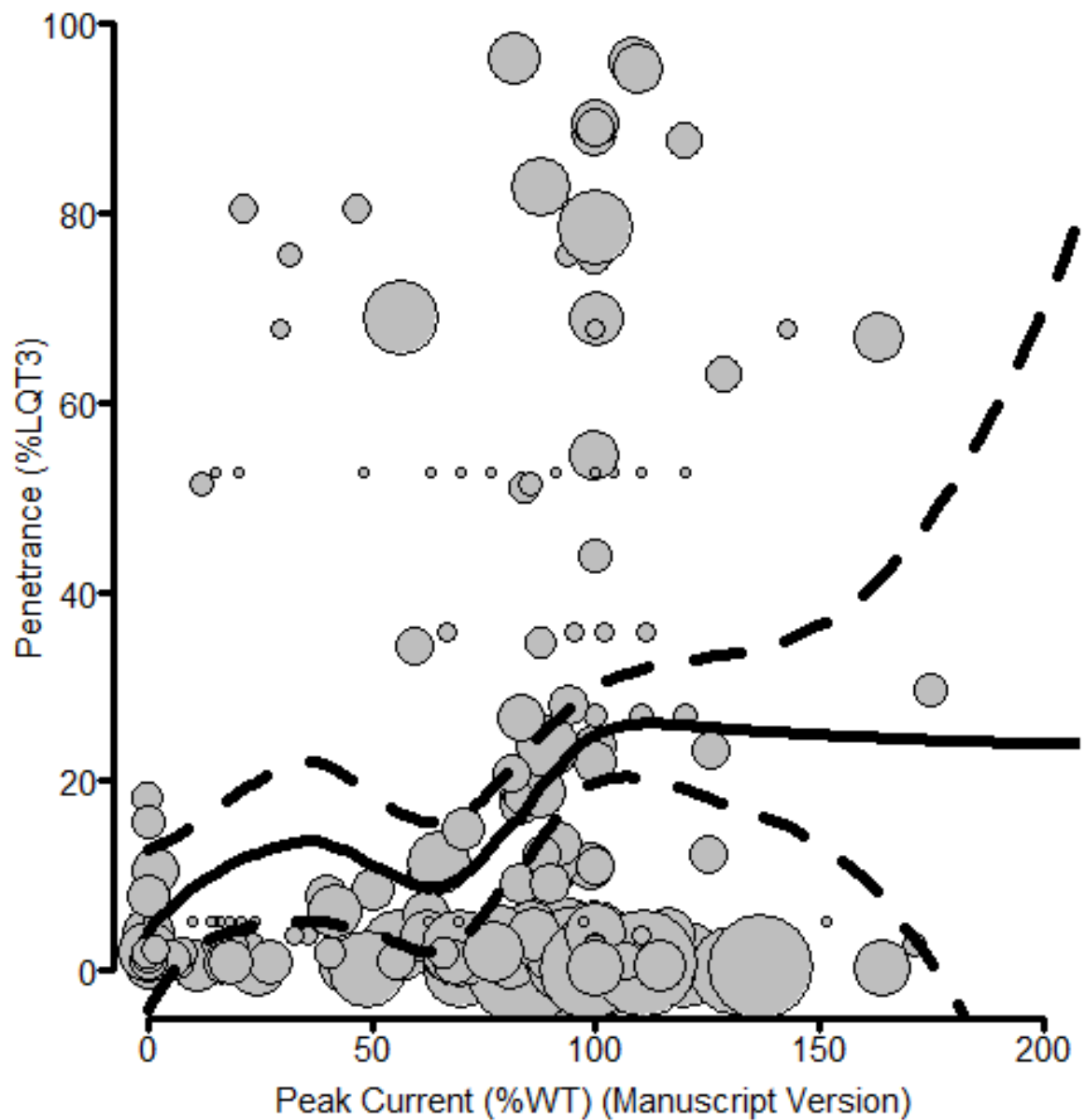
Sensitivity Analysis

Empirical Bayes (used in manuscript)

```
# Penetrance calculation used (Manuscript Version)
abrs0=0.32
alqt0=0.11
beta0=1
d$LQT_penetranceBayesian<-(d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0)
d$BrS_penetranceBayesian<-(d$brs1+abrs0)/(d$total_carriers+1+alqt0)
d$all_penetranceBayesian<-(d$brs1+abrs0+d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0+abrs0)
```

```
plt.disease("ipeak","Peak Current (%WT) (Manuscript Version)", "BrS_penetranceBayesian")
plt.disease("ipeak","Peak Current (%WT) (Manuscript Version)", "LQT_penetranceBayesian")
```

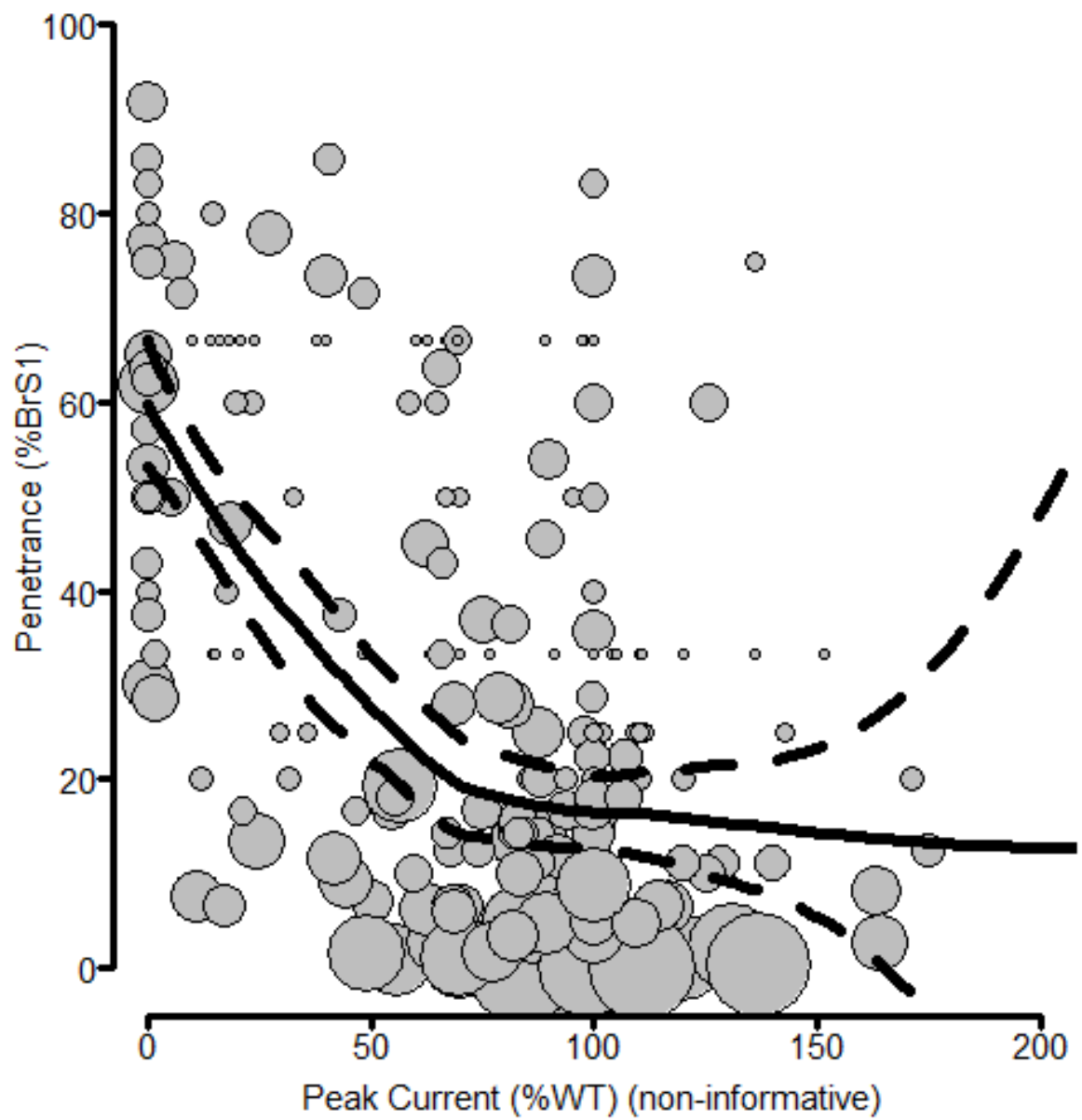


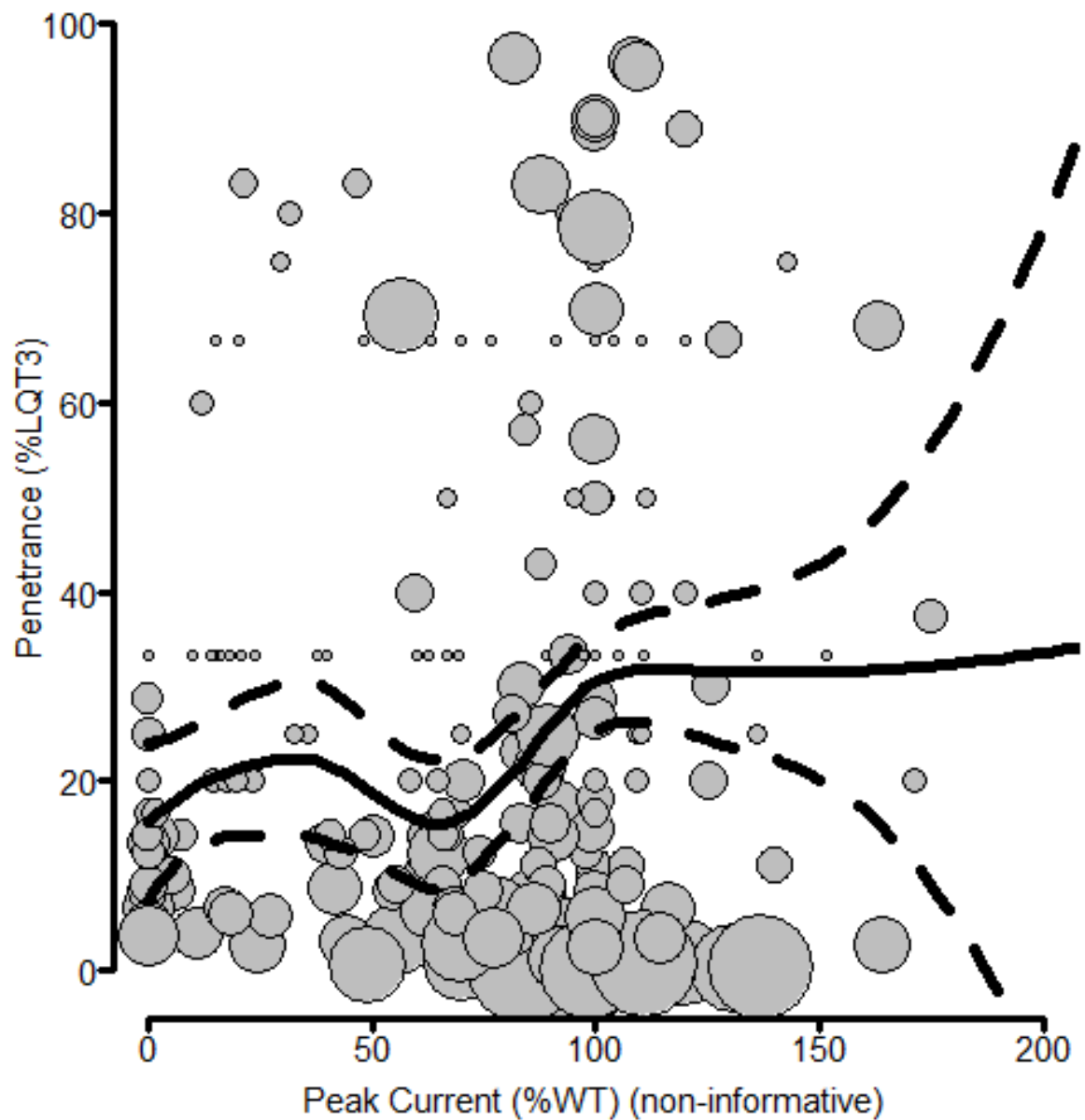


Uninformative Prior

```
# Changing penetrance calculation to uninformative prior
abrs0=1
alqt0=1
beta0=1
d$LQT_penetranceBayesian<-(d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0)
d$BrS_penetranceBayesian<-(d$brs1+abrs0)/(d$total_carriers+1+alqt0)
d$all_penetranceBayesian<-(d$brs1+abrs0+d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0+abrs0)

plt.disease("ipeak","Peak Current (%WT) (non-informative)", "BrS_penetranceBayesian")
plt.disease("ipeak","Peak Current (%WT) (non-informative)", "LQT_penetranceBayesian")
```

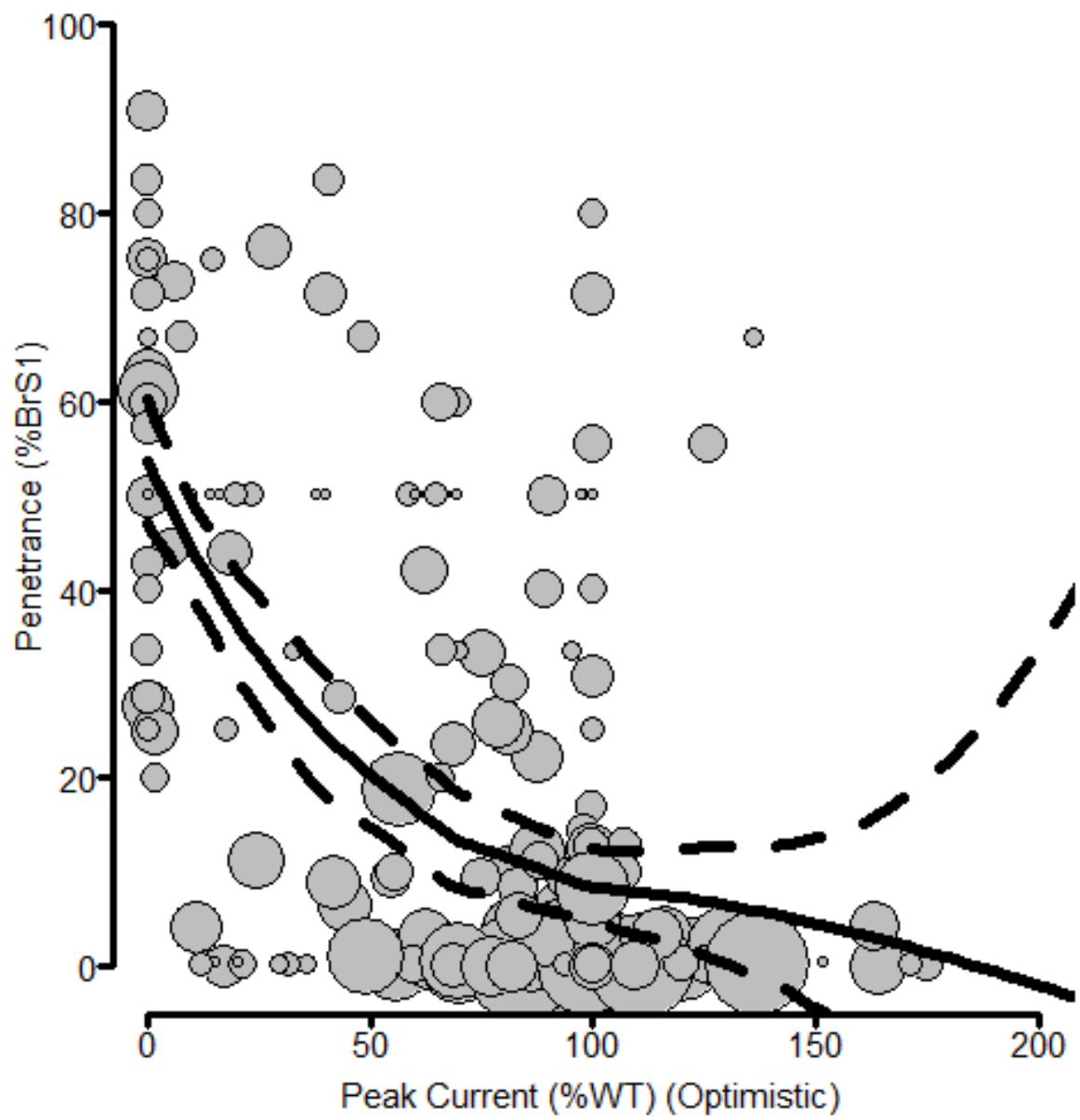


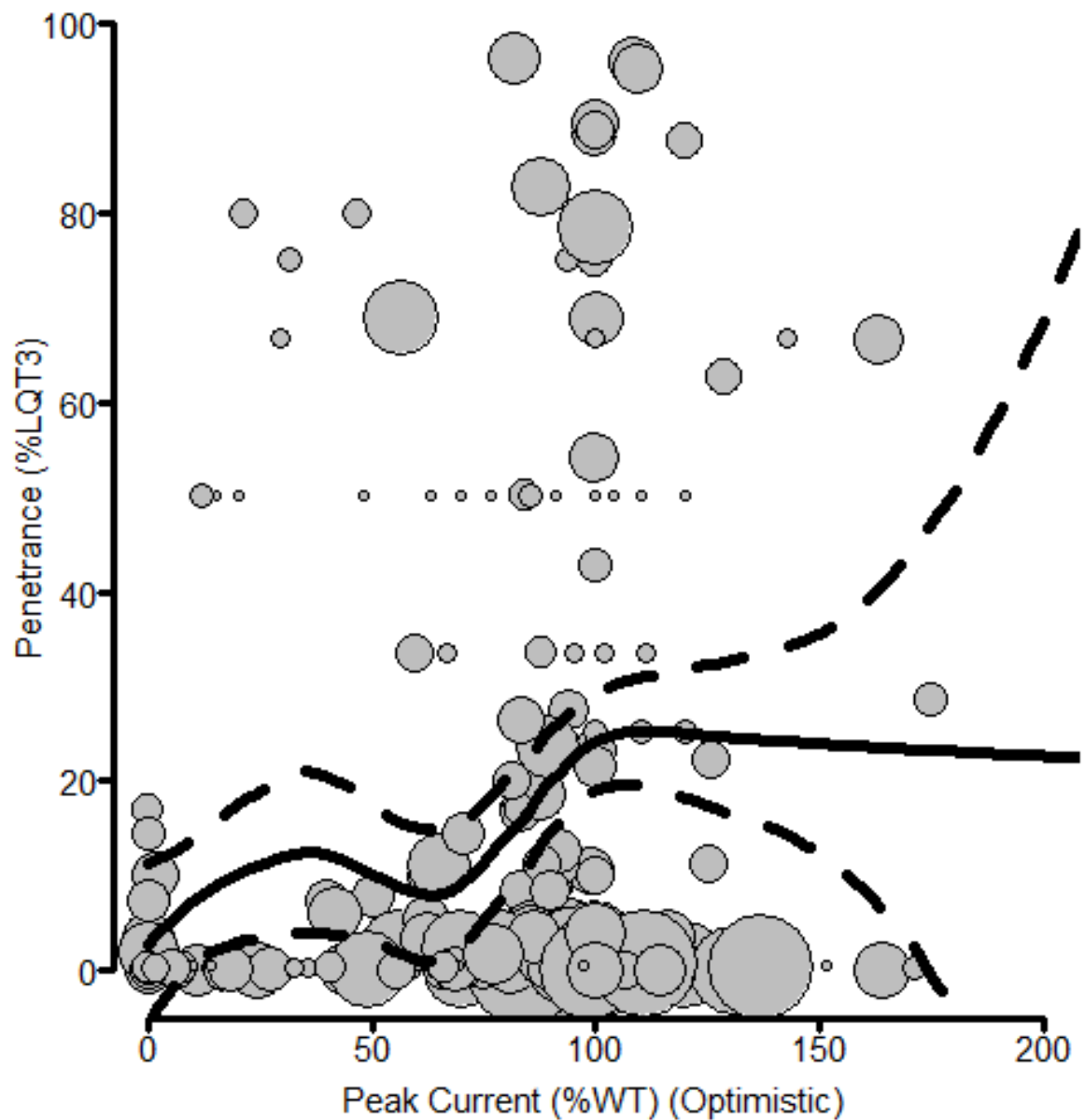


Optimistic Prior

```
# Changing penetrance calculation to optimistic (no affected carriers)
abrs0=0.01
alqt0=0.01
beta0=1
d$LQT_penetranceBayesian<-(d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0)
d$BrS_penetranceBayesian<-(d$brs1+abrs0)/(d$total_carriers+1+alqt0)
d$all_penetranceBayesian<-(d$brs1+abrs0+d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0+abrs0)

plt.disease("ipeak","Peak Current (%WT) (Optimistic)", "BrS_penetranceBayesian")
plt.disease("ipeak","Peak Current (%WT) (Optimistic)", "LQT_penetranceBayesian")
```





Pessimistic Prior

```
# Changing penetrance calculation to pessimistic (one affected carrier)
abrs0=1
alqt0=1
beta0=0.01
d$LQT_penetranceBayesian<-(d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0)
d$BrS_penetranceBayesian<-(d$brs1+abrs0)/(d$total_carriers+1+alqt0)
d$all_penetranceBayesian<-(d$brs1+abrs0+d$lqt3+alqt0)/(d$total_carriers+beta0+alqt0+abrs0)

plt.disease("ipeak","Peak Current (%WT) (Pessimistic)", "BrS_penetranceBayesian")
plt.disease("ipeak","Peak Current (%WT) (Pessimistic)", "LQT_penetranceBayesian")
```

