setwd("E:/Jun\_GeneExpression/Batch2")

changed to the following on 6/10/2010

setwd("F:/JunLi/Jun\_GeneExpression/Batch2")

sys.load.image(".RData", quiet=F)

acg<-as.matrix(read.delim("acg.txt",header=T,row.names=1,sep="\t"))

norm2.acg<-as.matrix(read.delim("acg\_norm2.txt",header=T,row.names=1,sep="\t"))

amy<-as.matrix(read.delim("amy.txt",header=T,row.names=1,sep="\t"))

norm2.amy<-as.matrix(read.delim("amy\_norm2.txt",header=T,row.names=1,sep="\t"))

dlpfc<-as.matrix(read.delim("dlpfc.txt",header=T,row.names=1,sep="\t"))

norm2.dlpfc<-as.matrix(read.delim("dlpfc\_norm2.txt",header=T,row.names=1,sep="\t"))

cb<-as.matrix(read.delim("cb.txt",header=T,row.names=1,sep="\t"))

norm2.cb<-as.matrix(read.delim("cb\_norm2.txt",header=T,row.names=1,sep="\t"))

hc<-as.matrix(read.delim("hc.txt",header=T,row.names=1,sep="\t"))

norm2.hc<-as.matrix(read.delim("hc\_norm2.txt",header=T,row.names=1,sep="\t"))

nacc<-as.matrix(read.delim("nacc.txt",header=T,row.names=1,sep="\t"))

norm2.nacc<-as.matrix(read.delim("nacc\_norm2.txt",header=T,row.names=1,sep="\t"))

ordered.sample3<-read.delim("sample\_ordered3.txt",header=T,row.names=1,sep="\t") #add chip-order

ordered.sample644<- ordered.sample3[ordered.sample3$remove=="keep",1:16]

total<-cbind(amy, acg, cb, dlpfc, hc, nacc)

total2<-cbind(norm2.amy, norm2.acg, norm2.cb, norm2.dlpfc, norm2.hc, norm2.nacc)

sum(colnames(total)==rownames(ordered.sample644))

Q10

q10<-matrix(NA,644,10)

q10.2<-matrix(NA,644,10)

q10[,1]<-apply(total,2,min)

q10.2[,1]<-apply(total2,2,min)

for (i in 2:9) {

q10[,i]<-apply(total,2,function(x) quantile(x,0.1\*i))

q10.2[,i]<-apply(total2,2,function(x) quantile(x,0.1\*i))

}

q10[,10]<-apply(total,2,max)

q10.2[,10]<-apply(total2,2,max)

par(mfrow=c(2,2))

plot(q10[,1],ylim=c(3,15),type="l",main="By region")

lines(q10[,3],col=2)

lines(q10[,5],col=3)

lines(q10[,7],col=4)

lines(q10[,9],col=5)

lines(q10[,10],col=6)

plot(q10[order(ordered.sample644$chip),1],ylim=c(3,15),type="l",main="By chip")

lines(q10[order(ordered.sample644$chip),3],col=2)

lines(q10[order(ordered.sample644$chip),5],col=3)

lines(q10[order(ordered.sample644$chip),7],col=4)

lines(q10[order(ordered.sample644$chip),9],col=5)

lines(q10[order(ordered.sample644$chip),10],col=6)

plot(q10[order(ordered.sample644$cohort),1],ylim=c(3,15),type="l",main="By cohort")

lines(q10[order(ordered.sample644$cohort),3],col=2)

lines(q10[order(ordered.sample644$cohort),5],col=3)

lines(q10[order(ordered.sample644$cohort),7],col=4)

lines(q10[order(ordered.sample644$cohort),9],col=5)

lines(q10[order(ordered.sample644$cohort),10],col=6)

plot(q10[order(ordered.sample644$position),1],ylim=c(3,15),type="l",main="By position A-G")

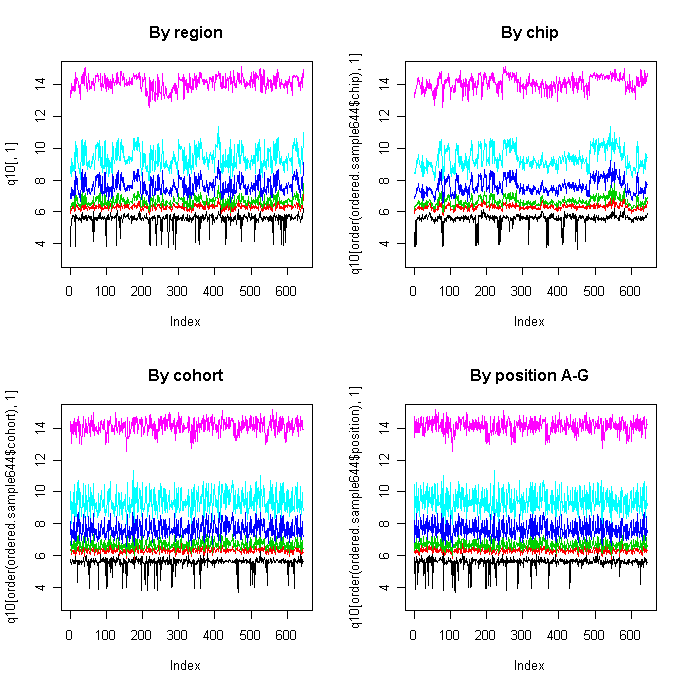
lines(q10[order(ordered.sample644$position),3],col=2)

lines(q10[order(ordered.sample644$position),5],col=3)

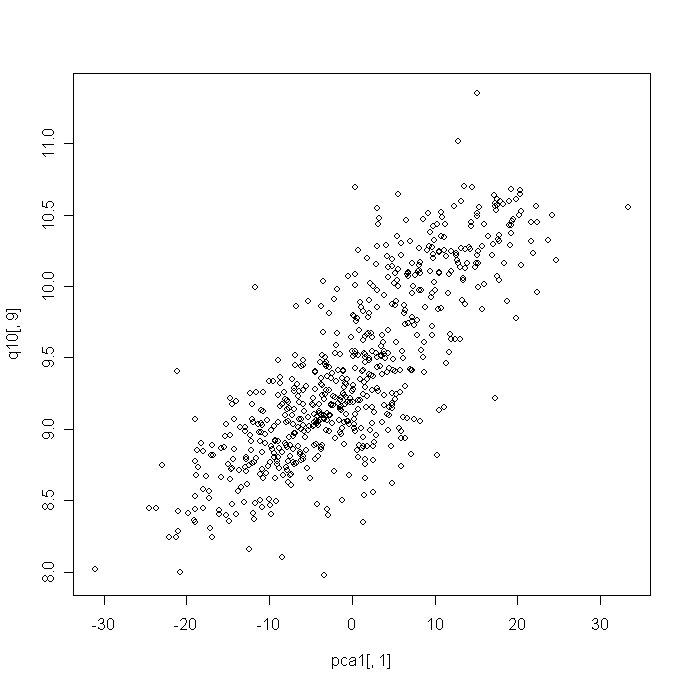
lines(q10[order(ordered.sample644$position),7],col=4)

lines(q10[order(ordered.sample644$position),9],col=5)

lines(q10[order(ordered.sample644$position),10],col=6)



plot(pca1[,1],q10[,9])



plot(q10.2[,1],ylim=c(3,12),type="l")

lines(q10.2[,3],col=2)

lines(q10.2[,5],col=3)

lines(q10.2[,7],col=4)

lines(q10.2[,9],col=5)

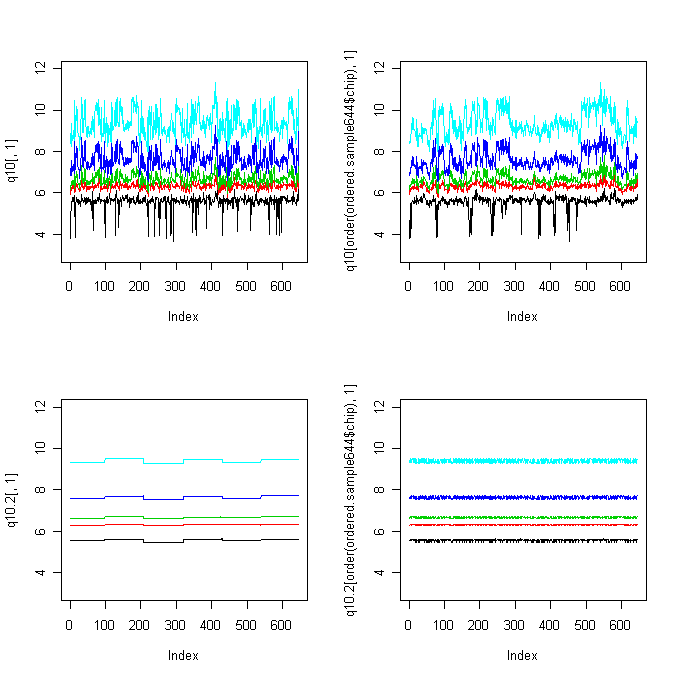
plot(q10.2[order(ordered.sample644$chip),1],ylim=c(3,12),type="l")

lines(q10.2[order(ordered.sample644$chip),3],col=2)

lines(q10.2[order(ordered.sample644$chip),5],col=3)

lines(q10.2[order(ordered.sample644$chip),7],col=4)

lines(q10.2[order(ordered.sample644$chip),9],col=5)



pca.raw<-prcomp(t(total))

pca.raw<-pca.raw$x[,1:6]

plot(pca.raw[,1],pca.raw[,2],pch=19,xlab="PC1-raw",ylab="PC2-raw")

points(pca.raw[ordered.sample644$Region=="ANCG",1],pca.raw[ordered.sample644$Region=="ANCG",2],col=2,pch=19)

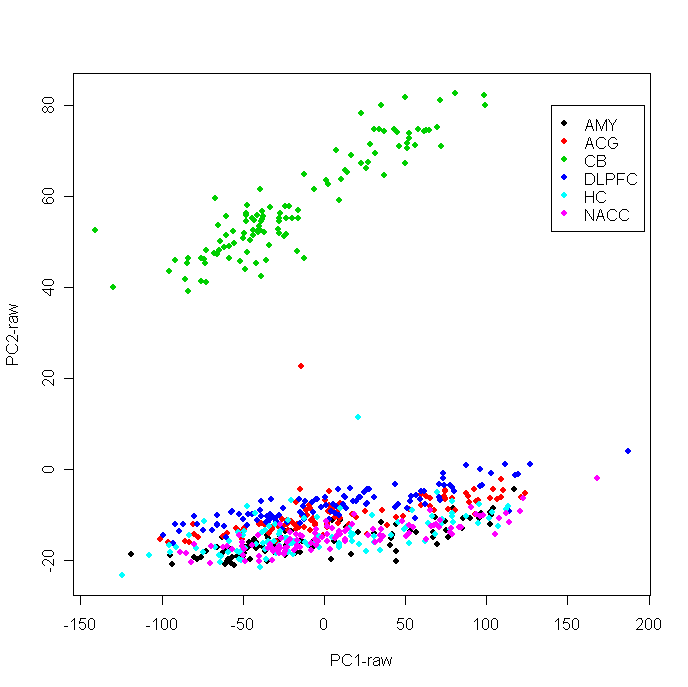
points(pca.raw[ordered.sample644$Region=="CB",1],pca.raw[ordered.sample644$Region=="CB",2],col=3,pch=19)

points(pca.raw[ordered.sample644$Region=="DLPFC",1],pca.raw[ordered.sample644$Region=="DLPFC",2],col=4,pch=19)

points(pca.raw[ordered.sample644$Region=="HC",1],pca.raw[ordered.sample644$Region=="HC",2],col=5,pch=19)

points(pca.raw[ordered.sample644$Region=="NACC",1],pca.raw[ordered.sample644$Region=="NACC",2],col=6,pch=19)

legend(140,80,c("AMY","ACG","CB","DLPFC","HC","NACC"),pch=rep(19,6),col=c(1:6))



plot(pca.raw[,1],pca.raw[,2],pch=19)

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

points(pca.raw[filter,1],pca.raw[filter,2],pch=19,col=2,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

points(pca.raw[filter,1],pca.raw[filter,2],pch=19,col=4,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

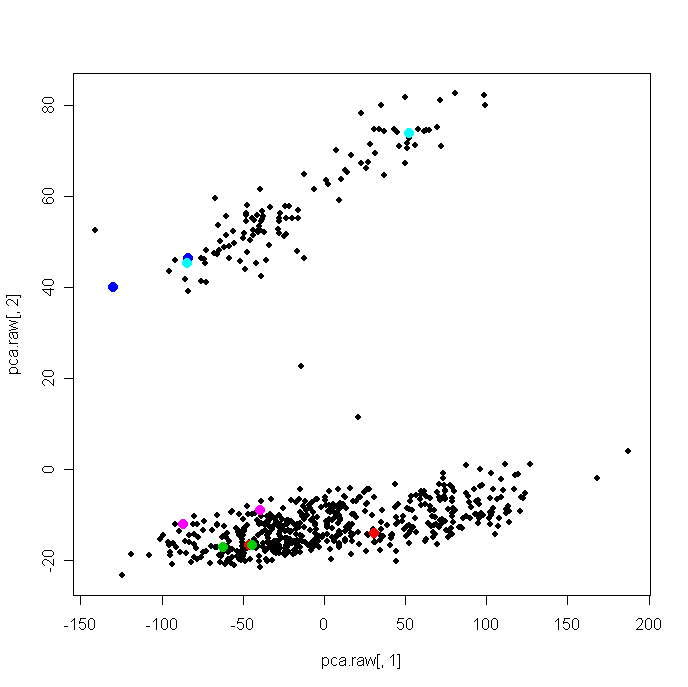
points(pca.raw[filter,1],pca.raw[filter,2],pch=19,col=5,cex=1.5)

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

points(pca.raw[filter,1],pca.raw[filter,2],pch=19,col=6,cex=1.5)

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

points(pca.raw[filter,1],pca.raw[filter,2],pch=19,col=3,cex=1.5)

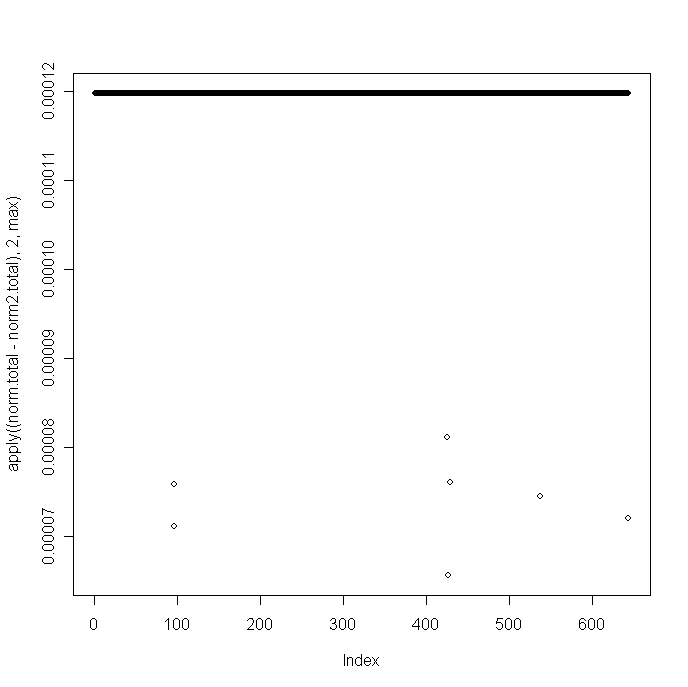


norm

norm.total<-normalize.quantiles(total)

norm2.total<-normalize.quantiles(total2)

plot(apply((norm.total-norm2.total),2,max)) #nearly the same



rm(norm2.total)

pca.norm<-prcomp(t(norm.total))

pca.norm<-pca.norm$x[,1:6]

plot(pca.norm[,1],pca.norm[,2],pch=19,xlab="PC1-norm",ylab="PC2-norm")

points(pca.norm[ordered.sample644$Region=="ANCG",1],pca.norm[ordered.sample644$Region=="ANCG",2],col=2,pch=19)

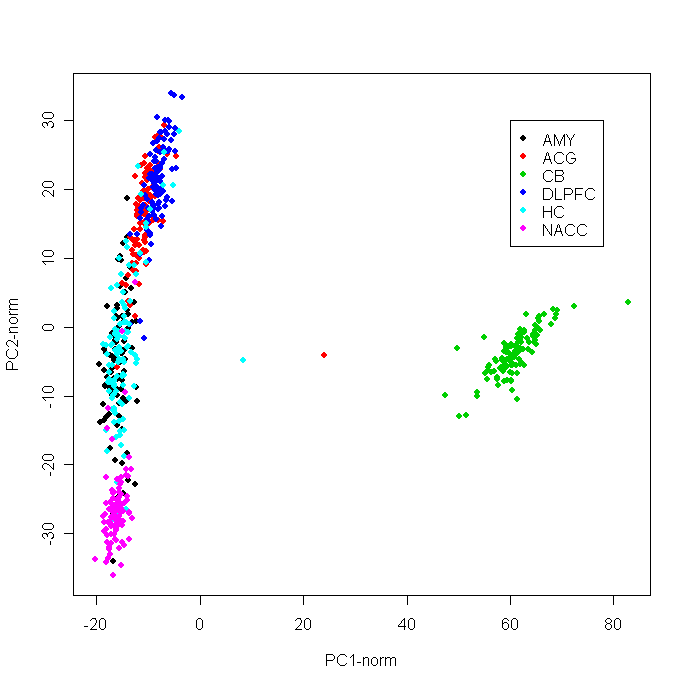
points(pca.norm[ordered.sample644$Region=="CB",1],pca.norm[ordered.sample644$Region=="CB",2],col=3,pch=19)

points(pca.norm[ordered.sample644$Region=="DLPFC",1],pca.norm[ordered.sample644$Region=="DLPFC",2],col=4,pch=19)

points(pca.norm[ordered.sample644$Region=="HC",1],pca.norm[ordered.sample644$Region=="HC",2],col=5,pch=19)

points(pca.norm[ordered.sample644$Region=="NACC",1],pca.norm[ordered.sample644$Region=="NACC",2],col=6,pch=19)

legend(60,30,c("AMY","ACG","CB","DLPFC","HC","NACC"),pch=rep(19,6),col=c(1:6))



plot(pca.norm[,1],pca.norm[,2],pch=19)

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

points(pca.norm[filter,1],pca.norm[filter,2],pch=19,col=2,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

points(pca.norm[filter,1],pca.norm[filter,2],pch=19,col=4,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

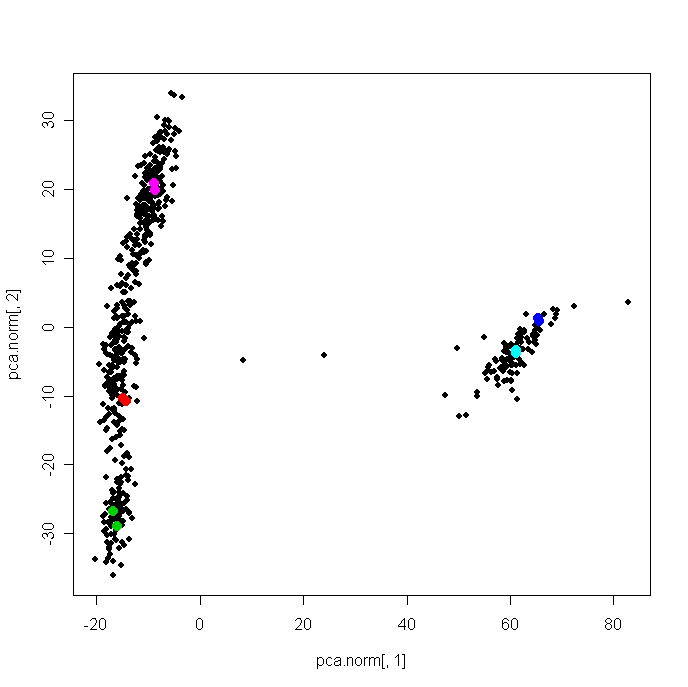
points(pca.norm[filter,1],pca.norm[filter,2],pch=19,col=5,cex=1.5)

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

points(pca.norm[filter,1],pca.norm[filter,2],pch=19,col=6,cex=1.5)

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

points(pca.norm[filter,1],pca.norm[filter,2],pch=19,col=3,cex=1.5)



ANOVA

chip<-matrix(NA,22177,4)

region<-matrix(NA,22177,4)

date()

for (i in 1:22177) {

tmp<-anova (lm(norm.total[i,]~as.factor(ordered.sample644$chip.order)+ as.factor(ordered.sample644$Region)))

chip[i,]<-as.numeric(tmp[1,2:5])

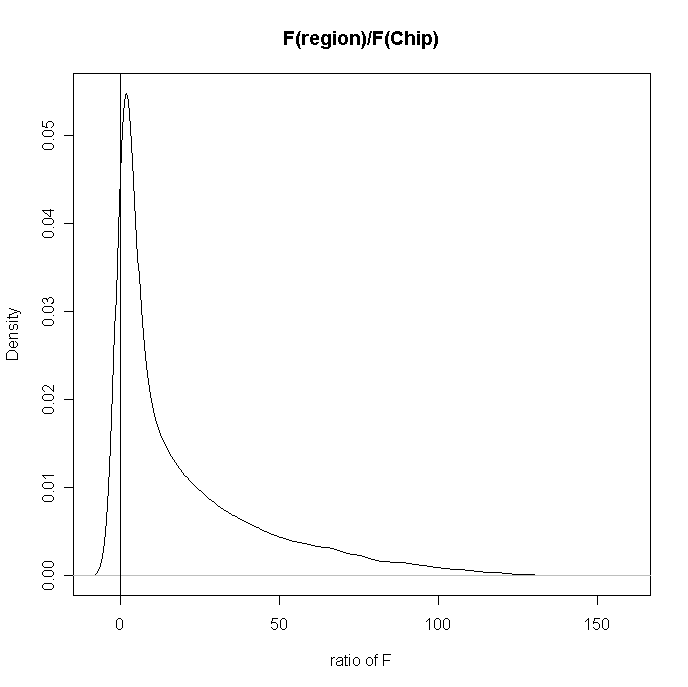
region[i,]<-as.numeric(tmp[2,2:5])

}

date()

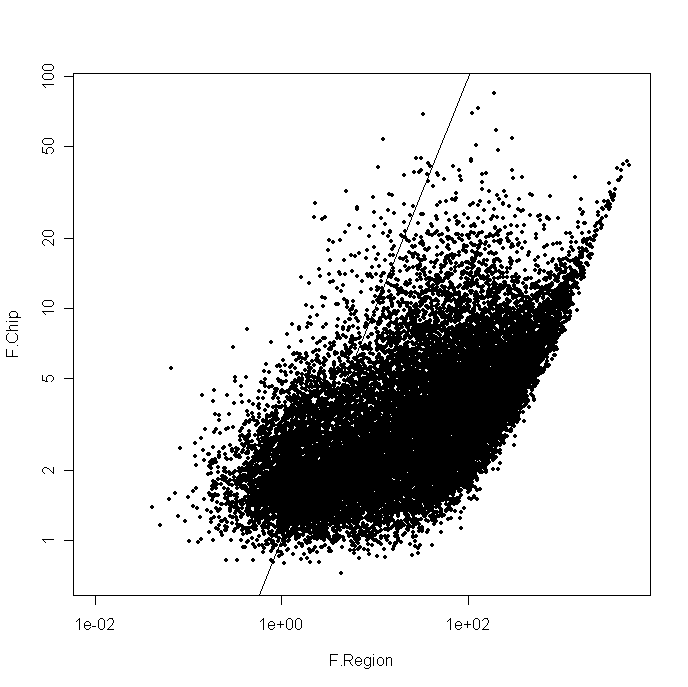
plot(density(region[,3]/chip[,3]),main="F(region)/F(Chip)",xlab="ratio of F")

abline(-1e-5,1e5)



plot(region[,3],chip[,3],xlab=c("F.Region"),ylab=c("F.Chip"),log="xy",xlim=c(0.01,5200),ylim=c(0.7,85),pch=19,cex=0.6)

abline(0,1)



sum(region[,3]>chip[,3])

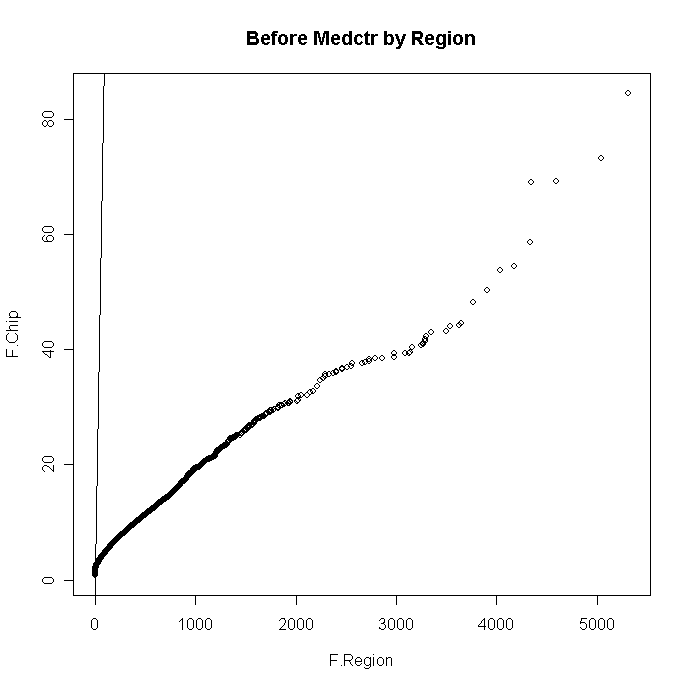
[1] 18819

> 18819/22177

[1] 0.8485819

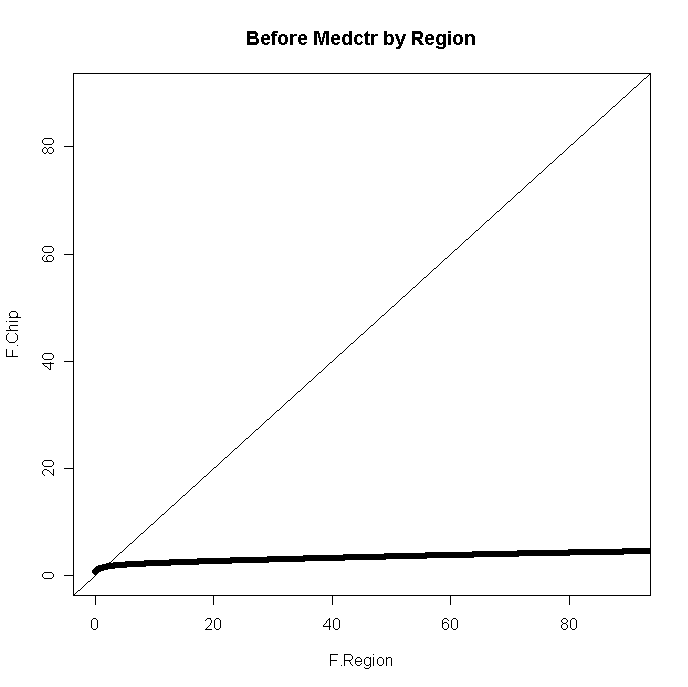
qqplot(region[,3],chip[,3],xlab=c("F.Region"),ylab=c("F.Chip"),main="Before Medctr by Region")

abline(0,1)



qqplot(region[,3],chip[,3],xlab=c("F.Region"),ylab=c("F.Chip"),main="Before Medctr by Region",xlim=c(0,90),ylim=c(0,90))

abline(0,1)



2. medctr

2.1 medctr

norm.amy<-norm.total[,1:98]

norm.acg<-norm.total[,99:207]

norm.cb<-norm.total[,208:320]

norm.dlpfc<-norm.total[,321:429]

norm.hc<-norm.total[,430:538]

norm.nacc<-norm.total[,539:644]

med.amy<-apply(norm.amy,1,median)

norm.medctr.amy<-norm.amy

for (i in 1:dim(norm.amy)[2]) {

norm.medctr.amy [,i]<-norm.amy[,i]-med.amy

}

med.acg<-apply(norm.acg,1,median)

norm.medctr.acg<-norm.acg

for (i in 1:dim(norm.acg)[2]) {

norm.medctr.acg [,i]<-norm.acg[,i]-med.acg

}

med.cb<-apply(norm.cb,1,median)

norm.medctr.cb<-norm.cb

for (i in 1:dim(norm.cb)[2]) {

norm.medctr.cb [,i]<-norm.cb[,i]-med.cb

}

med.dlpfc<-apply(norm.dlpfc,1,median)

norm.medctr.dlpfc<-norm.dlpfc

for (i in 1:dim(norm.dlpfc)[2]) {

norm.medctr.dlpfc [,i]<-norm.dlpfc[,i]-med.dlpfc

}

med.hc<-apply(norm.hc,1,median)

norm.medctr.hc<-norm.hc

for (i in 1:dim(norm.hc)[2]) {

norm.medctr.hc [,i]<-norm.hc[,i]-med.hc

}

med.nacc<-apply(norm.nacc,1,median)

norm.medctr.nacc<-norm.nacc

for (i in 1:dim(norm.nacc)[2]) {

norm.medctr.nacc [,i]<-norm.nacc[,i]-med.nacc

}

rm(norm.amy,norm.acg,norm.cb,norm.dlpfc,norm.hc,norm.nacc)

medctr.6<-cbind(norm.medctr.amy, norm.medctr.acg, norm.medctr.cb, norm.medctr.dlpfc, norm.medctr.hc, norm.medctr.nacc)

2.2

pca1<-prcomp(t(medctr.6))

pca1<-pca1$x[,1:6]

par(mfrow=c(2,2))

plot(pca1[,1],pca1[,2])

points(pca1[ordered.sample644$Region=="CB",1],pca1[ordered.sample644$Region=="CB",2],col=2,pch=19)

plot(pca1[,1],pca1[,3])

points(pca1[ordered.sample644$Region=="CB",1],pca1[ordered.sample644$Region=="CB",3],col=2,pch=19)

plot(pca1[,2],pca1[,3])

points(pca1[ordered.sample644$Region=="CB",2],pca1[ordered.sample644$Region=="CB",3],col=2,pch=19)

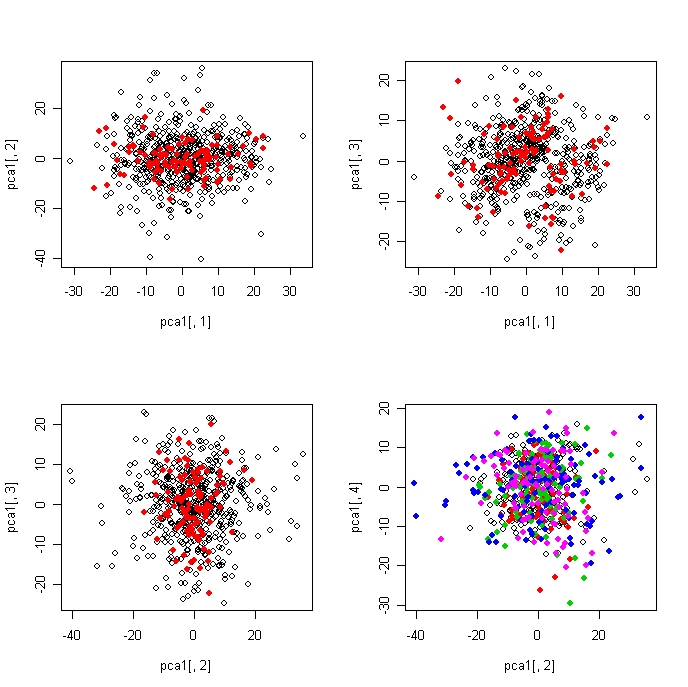
plot(pca1[,2],pca1[,4])

points(pca1[ordered.sample644$Region=="CB",2],pca1[ordered.sample644$Region=="CB",4],col=2,pch=19)

points(pca1[ordered.sample644$Region=="NACC",2],pca1[ordered.sample644$Region=="NACC",4],col=3,pch=19)

points(pca1[ordered.sample644$Region=="HC",2],pca1[ordered.sample644$Region=="HC",4],col=4,pch=19)

points(pca1[ordered.sample644$Region=="AMY",2],pca1[ordered.sample644$Region=="AMY",4],col=6,pch=19)



2.3

q10.3<-matrix(NA,644,10)

q10.3[,1]<-apply(medctr.6,2,min)

for (i in 2:9) {

q10.3[,i]<-apply(medctr.6,2,function(x) quantile(x,0.1\*i))

}

q10.3[,10]<-apply(medctr.6,2,max)

par(mfrow=c(1,2))

plot(q10.3[,2],ylim=c(-0.5,0.5),type="l")

lines(q10.3[,3],col=2)

lines(q10.3[,5],col=3)

lines(q10.3[,7],col=4)

lines(q10.3[,9],col=5)

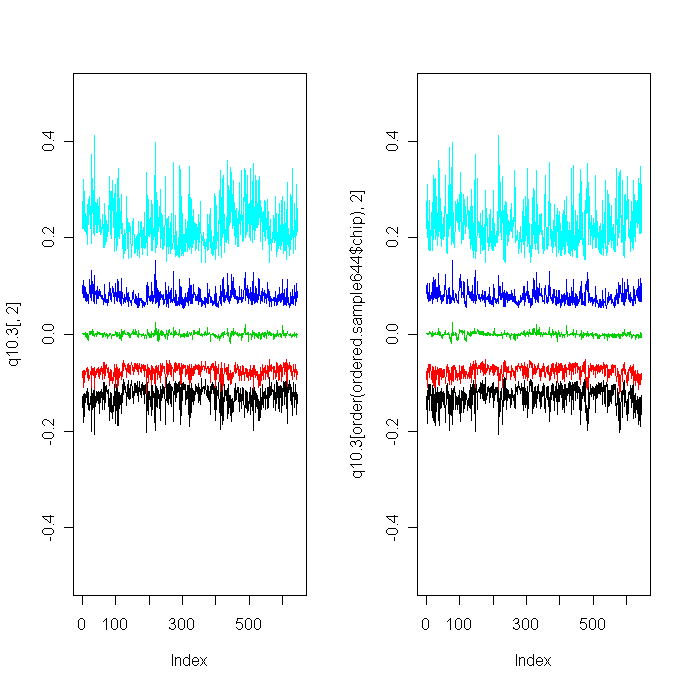
plot(q10.3[order(ordered.sample644$chip),2],ylim=c(-0.5,0.5),type="l")

lines(q10.3[order(ordered.sample644$chip),3],col=2)

lines(q10.3[order(ordered.sample644$chip),5],col=3)

lines(q10.3[order(ordered.sample644$chip),7],col=4)

lines(q10.3[order(ordered.sample644$chip),9],col=5)



2.4 ANOVA of medctr

chip2<-matrix(NA,22177,4)

region2<-matrix(NA,22177,4)

date()

for (i in 1:22177) {

tmp<-anova (lm(medctr.6[i,]~as.factor(ordered.sample644$chip.order)+ as.factor(ordered.sample644$Region)))

chip2[i,]<-as.numeric(tmp[1,2:5])

region2[i,]<-as.numeric(tmp[2,2:5])

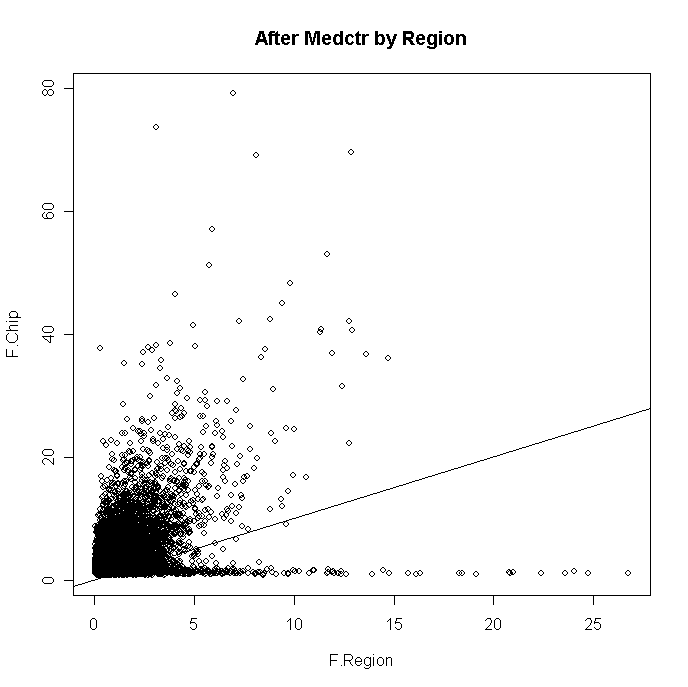
}

date()

Sum Sq, Mean Sq, F value, Pr(>F),

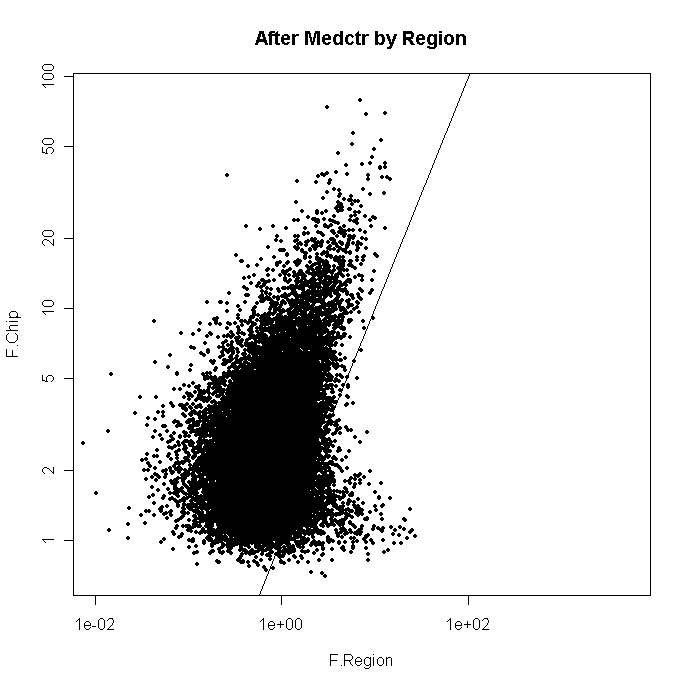
plot(region2[,3],chip2[,3],xlab=c("F.Region"),ylab=c("F.Chip"),main="After Medctr by Region")

abline(0,1)



plot(region2[,3],chip2[,3],xlab=c("F.Region"),ylab=c("F.Chip"), main="After Medctr by Region",log="xy",xlim=c(0.01,5200),ylim=c(0.7,85),pch=19,cex=0.6)

abline(0,1)



sum(region2[,3]>chip2[,3])

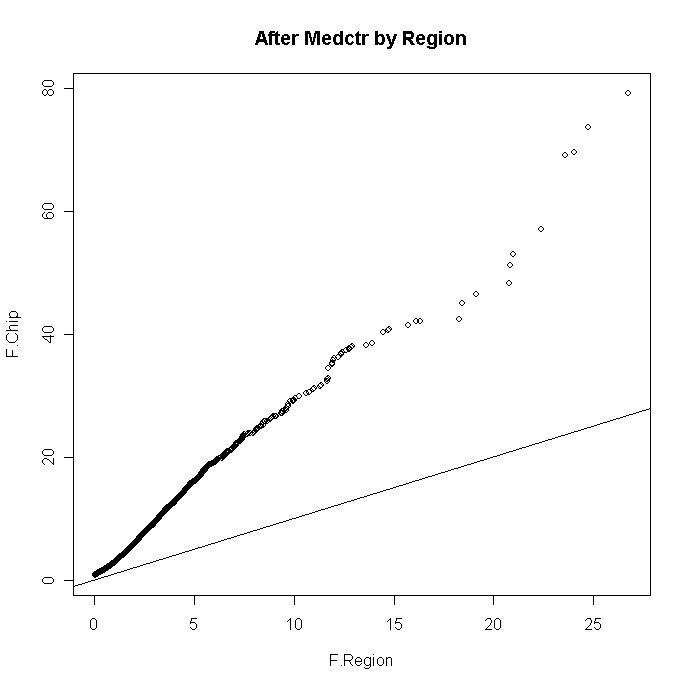
[1] 1292

> 1292/22177

[1] 0.05825856

qqplot(region2[,3],chip2[,3],xlab=c("F.Region"),ylab=c("F.Chip"),main="After Medctr by Region")

abline(0,1)



2.5 plot against chip

par(mfrow=c(2,2))

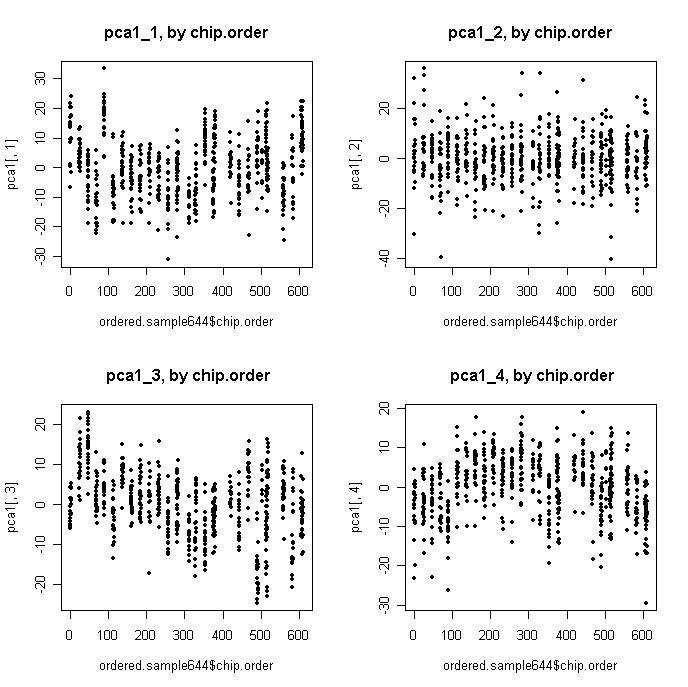
plot(ordered.sample644$chip.order,pca1[,1], main="pca1\_1, by chip.order",pch=19,cex=0.7)

plot(ordered.sample644$chip.order,pca1[,2] , main="pca1\_2, by chip.order",pch=19,cex=0.7)

plot(ordered.sample644$chip.order,pca1[,3] , main=" pca1\_3, by chip.order",pch=19,cex=0.7)

plot(ordered.sample644$chip.order,pca1[,4] , main=" pca1\_4, by chip.order",pch=19,cex=0.7)

par(mfrow=c(1,1))



par(mfrow=c(2,2))

plot(ordered.sample644$chip.order,pca1[,1], main="pca1\_1, by chip.order",pch=19,cex=0.7)

points(ordered.sample644$chip.order[ordered.sample644$Region=="CB"],pca1[ordered.sample644$Region=="CB",1],pch=19,cex=0.7,col=3)

plot(ordered.sample644$chip.order,pca1[,2] , main="pca1\_2, by chip.order",pch=19,cex=0.7)

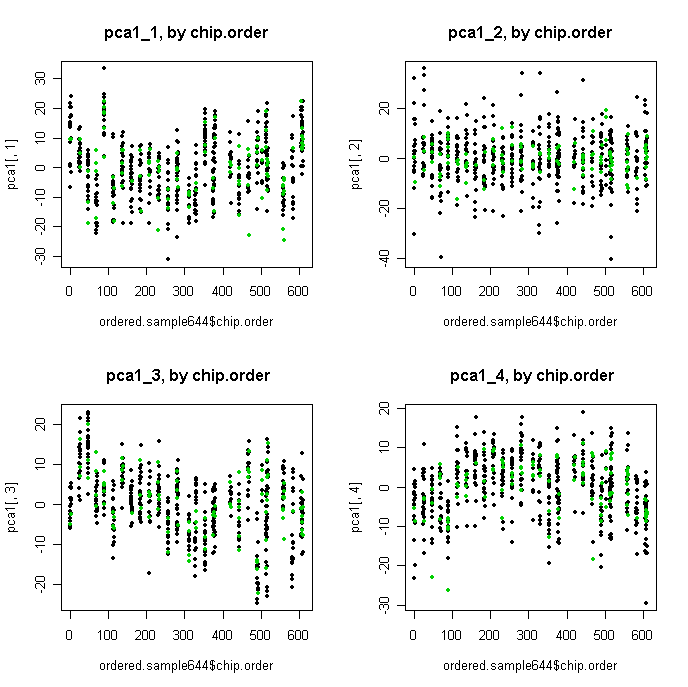
points(ordered.sample644$chip.order[ordered.sample644$Region=="CB"],pca1[ordered.sample644$Region=="CB",2],pch=19,cex=0.7,col=3)

plot(ordered.sample644$chip.order,pca1[,3] , main=" pca1\_3, by chip.order",pch=19,cex=0.7)

points(ordered.sample644$chip.order[ordered.sample644$Region=="CB"],pca1[ordered.sample644$Region=="CB",3],pch=19,cex=0.7,col=3)

plot(ordered.sample644$chip.order,pca1[,4] , main=" pca1\_4, by chip.order",pch=19,cex=0.7)

points(ordered.sample644$chip.order[ordered.sample644$Region=="CB"],pca1[ordered.sample644$Region=="CB",4],pch=19,cex=0.7,col=3)



plot(ordered.sample644$chip.order,pca1[,1], main="pca1\_1, by chip.order",pch=19,cex=0.7)

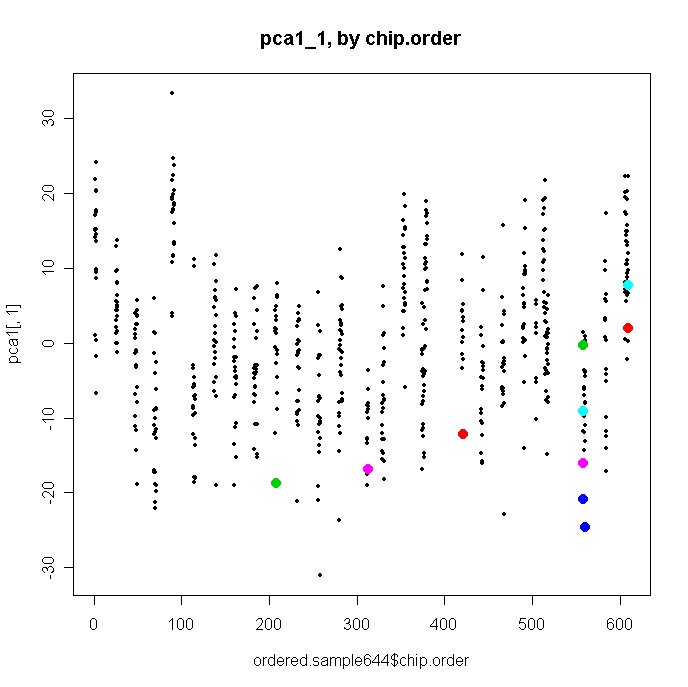
points(ordered.sample644$chip.order[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")],pca1[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445"),1],pch=19,cex=1.5,col=2)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")],pca1[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667"),1],pch=19,cex=1.5,col=4)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")],pca1[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196"),1],pch=19,cex=1.5,col=5)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")],pca1[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273"),1],pch=19,cex=1.5,col=6)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")],pca1[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228"),1],pch=19,cex=1.5,col=3)



par(mfrow=c(2,2))

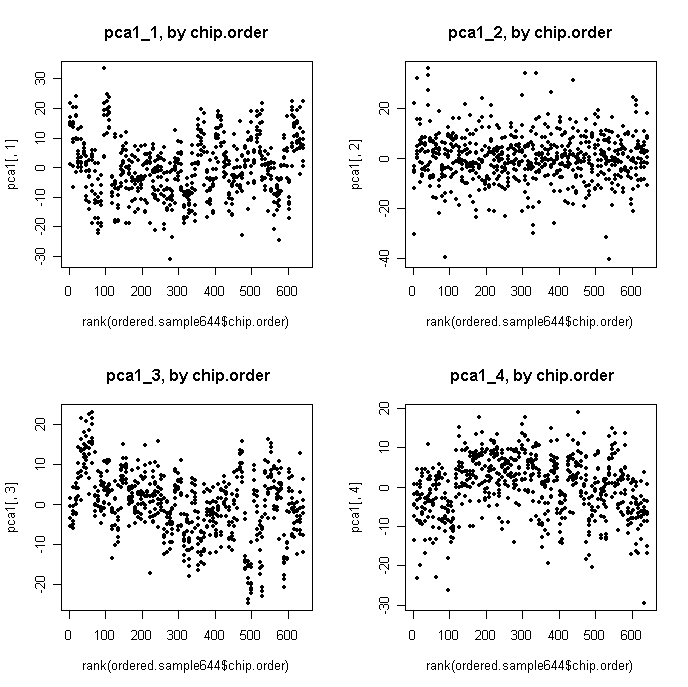
plot(rank(ordered.sample644$chip.order),pca1[,1], main="pca1\_1, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca1[,2] , main="pca1\_2, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca1[,3] , main=" pca1\_3, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca1[,4] , main=" pca1\_4, by chip.order",pch=19,cex=0.7)

par(mfrow=c(1,1))



3. medctr by chip

chip.levels<-levels(factor(ordered.sample644$chip.order))

medctr.6.85<-matrix(NA,22177,644)

for (i in 4:85) {

tmp<-medctr.6[,ordered.sample644$chip.order==chip.levels[i]]

med.chip<-apply(tmp,1,median)

medctr.6.85[,ordered.sample644$chip.order==chip.levels[i]]<-medctr.6[,ordered.sample644$chip.order==chip.levels[i]]-med.chip

}

3.2

pca2<-prcomp(t(medctr.6.85))

pca2<-pca2$x[,1:6]

par(mfrow=c(2,2))

plot(pca2[,1],pca2[,2])

points(pca2[ordered.sample644$Region=="CB",1],pca2[ordered.sample644$Region=="CB",2],col=2,pch=19)

plot(pca2[,1],pca2[,3])

points(pca2[ordered.sample644$Region=="CB",1],pca2[ordered.sample644$Region=="CB",3],col=2,pch=19)

plot(pca2[,2],pca2[,3])

points(pca2[ordered.sample644$Region=="CB",2],pca2[ordered.sample644$Region=="CB",3],col=2,pch=19)

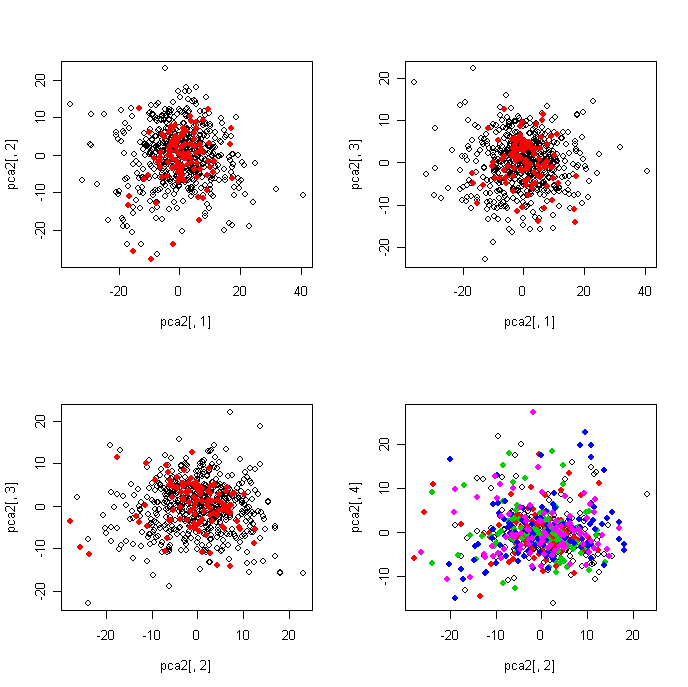
plot(pca2[,2],pca2[,4])

points(pca2[ordered.sample644$Region=="CB",2],pca2[ordered.sample644$Region=="CB",4],col=2,pch=19)

points(pca2[ordered.sample644$Region=="NACC",2],pca2[ordered.sample644$Region=="NACC",4],col=3,pch=19)

points(pca2[ordered.sample644$Region=="HC",2],pca2[ordered.sample644$Region=="HC",4],col=4,pch=19)

points(pca2[ordered.sample644$Region=="AMY",2],pca2[ordered.sample644$Region=="AMY",4],col=6,pch=19)



plot(pca2[,1],pca2[,2] ,xlab="PC1.freeze3", ylab="PC2.freeze3",pch=19)

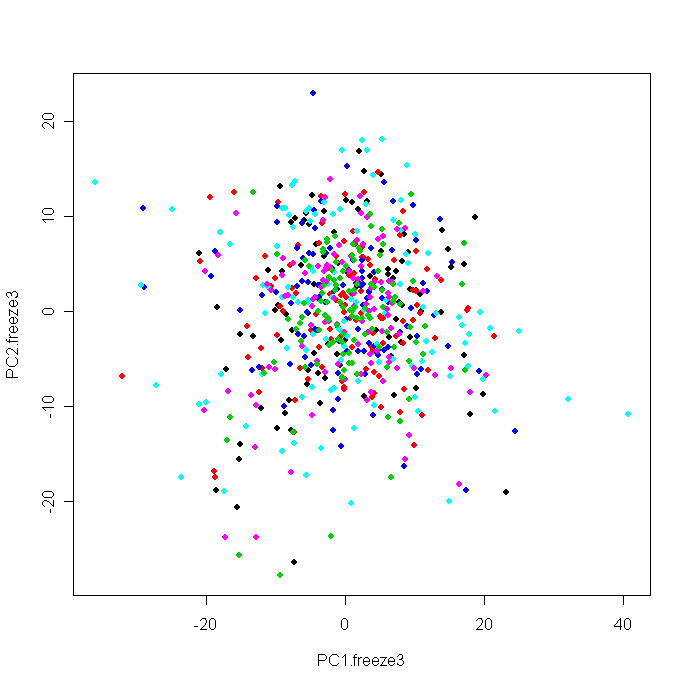
points(pca2[ordered.sample644$Region=="ANCG",1],pca2[ordered.sample644$Region=="ANCG",2],col=2,pch=19)

points(pca2[ordered.sample644$Region=="DLPFC",1],pca2[ordered.sample644$Region=="DLPFC",2],col=4,pch=19)

points(pca2[ordered.sample644$Region=="HC",1],pca2[ordered.sample644$Region=="HC",2],col=5,pch=19)

points(pca2[ordered.sample644$Region=="NACC",1],pca2[ordered.sample644$Region=="NACC",2],col=6,pch=19)

points(pca2[ordered.sample644$Region=="CB",1],pca2[ordered.sample644$Region=="CB",2],col=3,pch=19)



3.3

q10.4<-matrix(NA,644,10)

q10.4[,1]<-apply(medctr.6.85,2,min)

for (i in 2:9) {

q10.4[,i]<-apply(medctr.6.85,2,function(x) quantile(x,0.1\*i))

}

q10.4[,10]<-apply(medctr.6.85,2,max)

par(mfrow=c(1,2))

plot(q10.4[,2],ylim=c(-0.3,0.3),type="l")

lines(q10.4[,3],col=2)

lines(q10.4[,5],col=3)

lines(q10.4[,7],col=4)

lines(q10.4[,9],col=5)

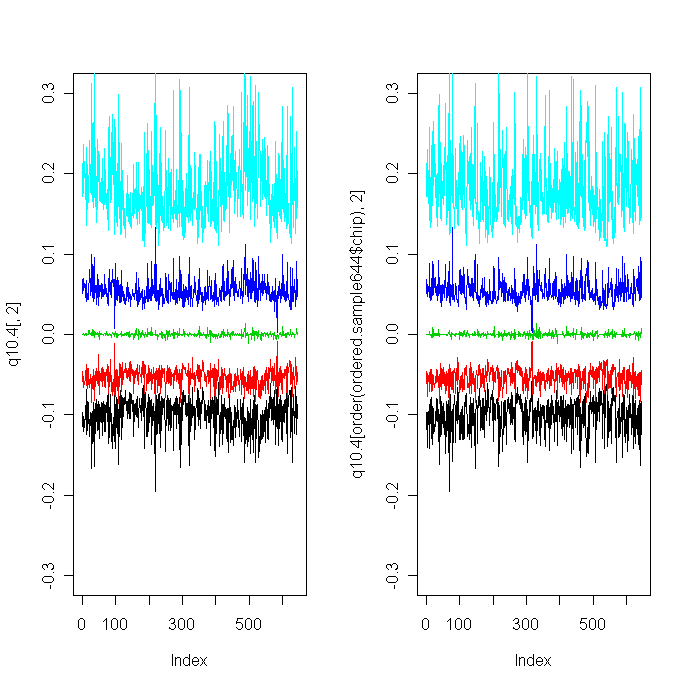
plot(q10.4[order(ordered.sample644$chip),2],ylim=c(-0.3,0.3),type="l")

lines(q10.4[order(ordered.sample644$chip),3],col=2)

lines(q10.4[order(ordered.sample644$chip),5],col=3)

lines(q10.4[order(ordered.sample644$chip),7],col=4)

lines(q10.4[order(ordered.sample644$chip),9],col=5)



3.4 ANOVA of medctr by chip

chip3<-matrix(NA,22177,4)

region3<-matrix(NA,22177,4)

date()

for (i in 1:22177) {

tmp<-anova (lm(medctr.6.85[i,]~as.factor(ordered.sample644$chip.order)+ as.factor(ordered.sample644$Region)))

chip3[i,]<-as.numeric(tmp[1,2:5])

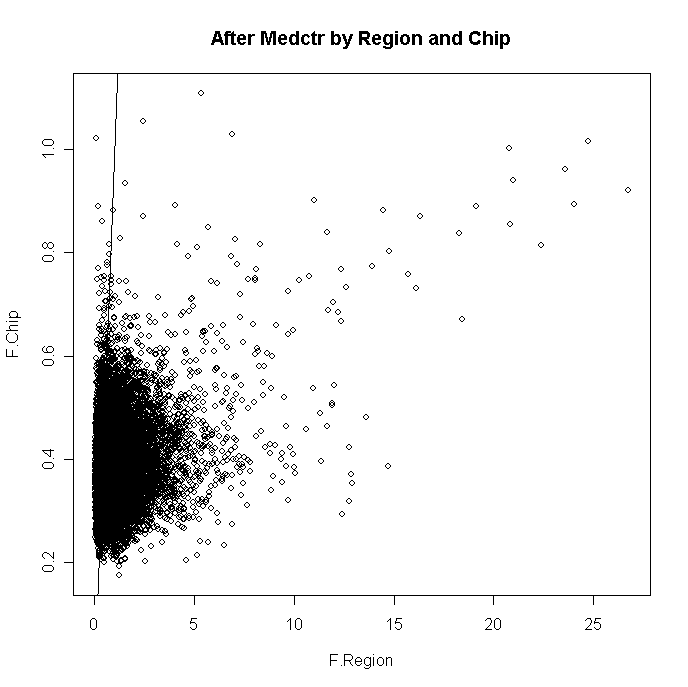
region3[i,]<-as.numeric(tmp[2,2:5])

}

date()

plot(region3[,3],chip3[,3],xlab=c("F.Region"),ylab=c("F.Chip"),main="After Medctr by Region and Chip")

abline(0,1)

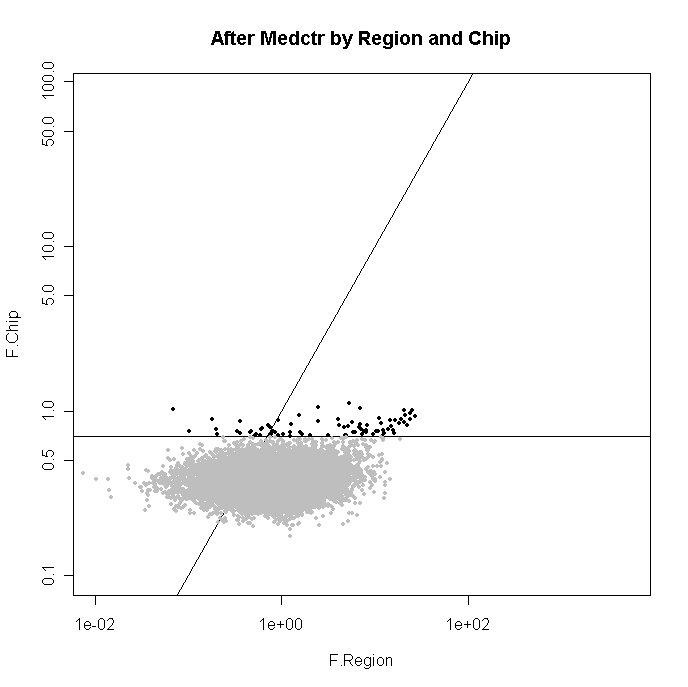


plot(region3[,3],chip3[,3],xlab=c("F.Region"),ylab=c("F.Chip"), main="After Medctr by Region and Chip",log="xy",xlim=c(0.01,5200),ylim=c(0.1,85),pch=19,cex=0.6)

abline(0,1)

abline(-0.1549,0)

points(region3[chip3[,3]<0.7,3],chip3[chip3[,3]<0.7,3],xlab=c("F.Region"),ylab=c("F.Chip"), pch=19,col="gray",cex=0.6)



sum(region3[,3]>chip3[,3])

[1] 18215

> 18215/22177

[1] 0.8213464

chip2 is smaller

3.5 plot against chip

par(mfrow=c(2,2))

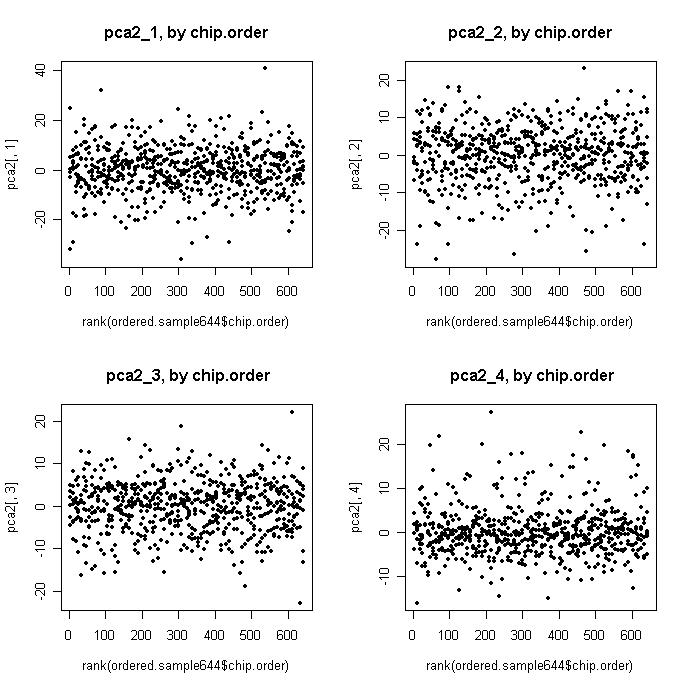
plot(rank(ordered.sample644$chip.order),pca2[,1], main="pca2\_1, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca2[,2] , main="pca2\_2, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca2[,3] , main=" pca2\_3, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca2[,4] , main=" pca2\_4, by chip.order",pch=19,cex=0.7)

par(mfrow=c(1,1))



plot(ordered.sample644$chip.order,pca2[,2], main="pca2\_2, by chip.order",pch=19,cex=0.7)

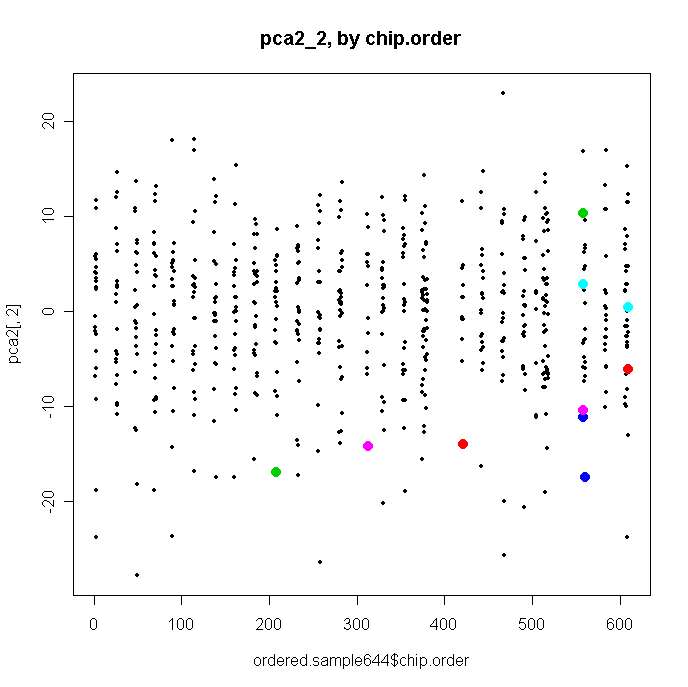
points(ordered.sample644$chip.order[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")],pca2[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445"),2],pch=19,cex=1.5,col=2)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")],pca2[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667"),2],pch=19,cex=1.5,col=4)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")],pca2[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196"),2],pch=19,cex=1.5,col=5)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")],pca2[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273"),2],pch=19,cex=1.5,col=6)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")],pca2[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228"),2],pch=19,cex=1.5,col=3)



export medctr.6.85

colnames(medctr.6.85)<-colnames(total)

rownames(medctr.6.85)<-rownames(total)

write.table(medctr.6.85,"medctr\_region\_chip.txt",sep="\t")

save.image()

4. norm the medctr.6

norm.medctr6<-normalize.quantiles(medctr.6)

4.2

pca3<-prcomp(t(norm.medctr6))

pca3<-pca3$x[,1:6]

par(mfrow=c(2,2))

plot(pca3[,1],pca3[,2])

points(pca3[ordered.sample644$Region=="CB",1],pca3[ordered.sample644$Region=="CB",2],col=2,pch=19)

plot(pca3[,1],pca3[,3])

points(pca3[ordered.sample644$Region=="CB",1],pca3[ordered.sample644$Region=="CB",3],col=2,pch=19)

plot(pca3[,2],pca3[,3])

points(pca3[ordered.sample644$Region=="CB",2],pca3[ordered.sample644$Region=="CB",3],col=2,pch=19)

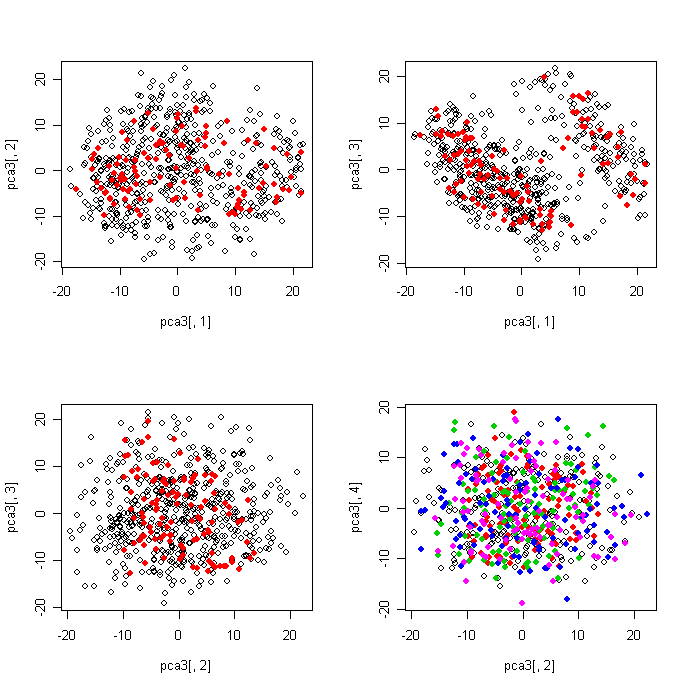
plot(pca3[,2],pca3[,4])

points(pca3[ordered.sample644$Region=="CB",2],pca3[ordered.sample644$Region=="CB",4],col=2,pch=19)

points(pca3[ordered.sample644$Region=="NACC",2],pca3[ordered.sample644$Region=="NACC",4],col=3,pch=19)

points(pca3[ordered.sample644$Region=="HC",2],pca3[ordered.sample644$Region=="HC",4],col=4,pch=19)

points(pca3[ordered.sample644$Region=="AMY",2],pca3[ordered.sample644$Region=="AMY",4],col=6,pch=19)



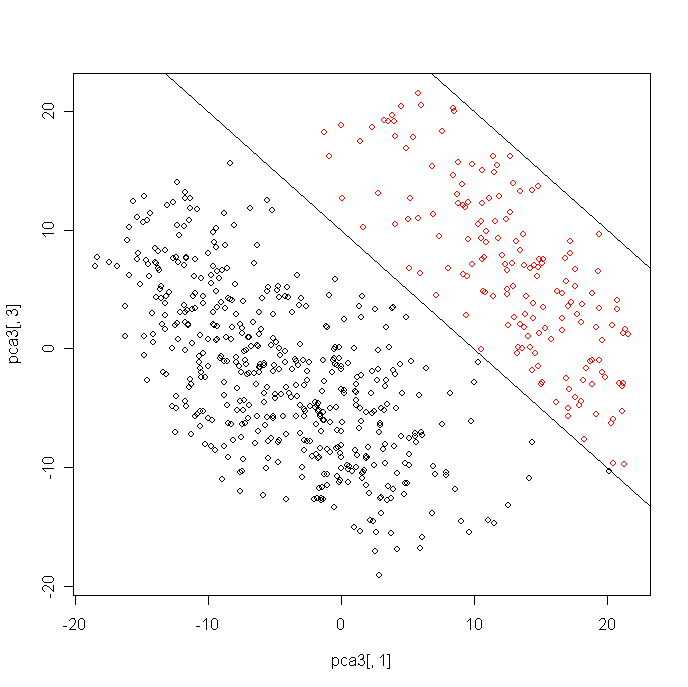
plot(pca3[,1],pca3[,3])

> filter<-((pca3[,1]+pca3[,3])>10)

> sum(filter)

[1] 166

> points(pca3[filter,1],pca3[filter,3],col=2)



4.3

q10.5<-matrix(NA,644,10)

q10.5[,1]<-apply(norm.medctr6,2,min)

for (i in 2:9) {

q10.5[,i]<-apply(norm.medctr6,2,function(x) quantile(x,0.1\*i))

}

q10.5[,10]<-apply(norm.medctr6,2,max)

par(mfrow=c(1,2))

plot(q10.5[,2],ylim=c(-0.3,0.3),type="l")

lines(q10.5[,3],col=2)

lines(q10.5[,5],col=3)

lines(q10.5[,7],col=4)

lines(q10.5[,9],col=5)

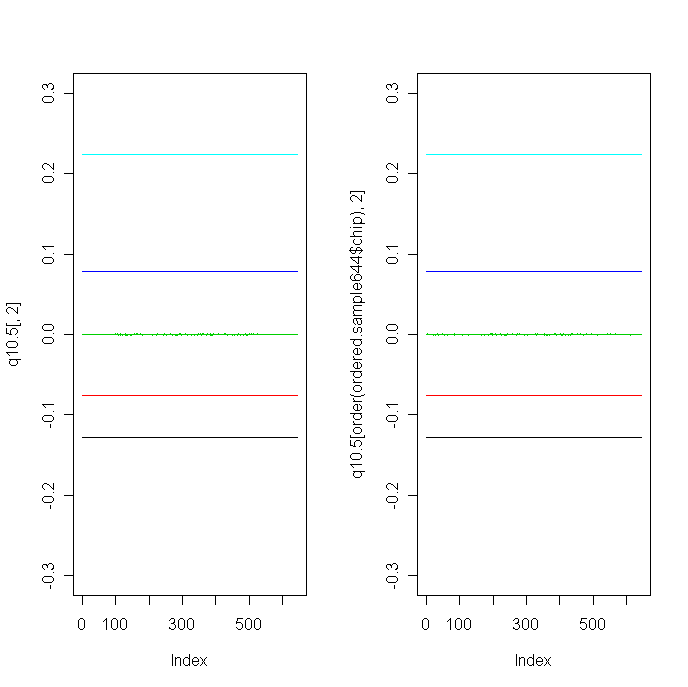
plot(q10.5[order(ordered.sample644$chip),2],ylim=c(-0.3,0.3),type="l")

lines(q10.5[order(ordered.sample644$chip),3],col=2)

lines(q10.5[order(ordered.sample644$chip),5],col=3)

lines(q10.5[order(ordered.sample644$chip),7],col=4)

lines(q10.5[order(ordered.sample644$chip),9],col=5)



4.4 ANOVA of medctr by chip

chip4<-matrix(NA,22177,4)

region4<-matrix(NA,22177,4)

date()

for (i in 1:22177) {

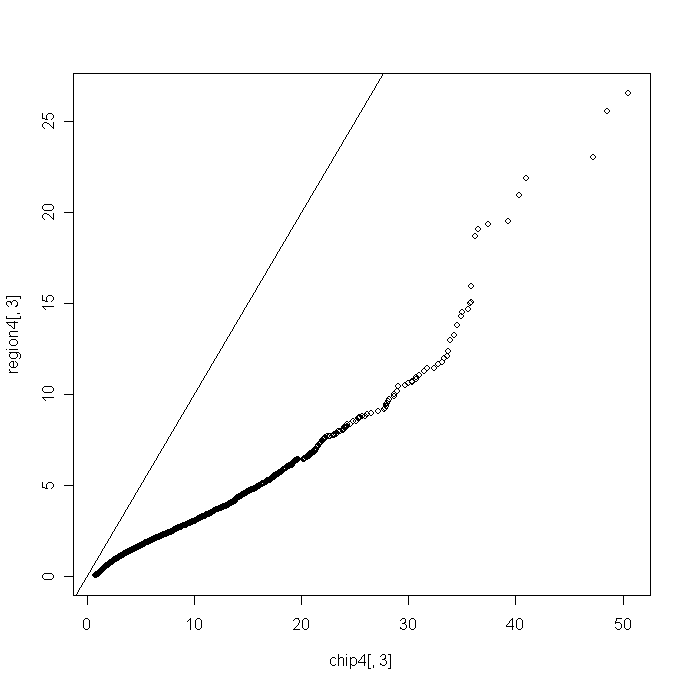
tmp<-anova (lm(norm.medctr6[i,]~as.factor(ordered.sample644$chip.order)+ as.factor(ordered.sample644$Region)))

chip4[i,]<-as.numeric(tmp[1,2:5])

region4[i,]<-as.numeric(tmp[2,2:5])

}

date()



sum(region4[,3]>chip4[,3])

[1] 1147

> 1147/22177

[1] 0.05172025

4.5 plot against chip

par(mfrow=c(2,2))

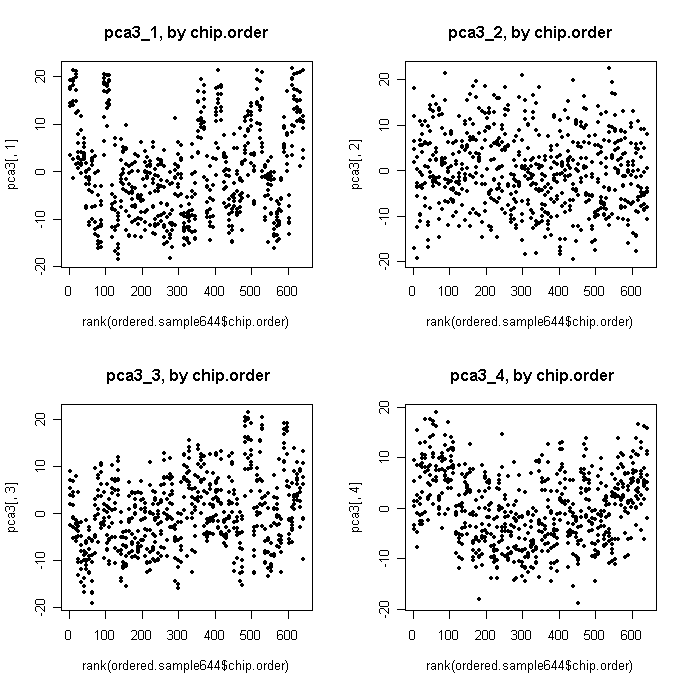
plot(rank(ordered.sample644$chip.order),pca3[,1], main="pca3\_1, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca3[,2] , main="pca3\_2, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca3[,3] , main=" pca3\_3, by chip.order",pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca3[,4] , main=" pca3\_4, by chip.order",pch=19,cex=0.7)

par(mfrow=c(1,1))



par(mfrow=c(2,2))

plot(rank(ordered.sample644$chip.order),pca3[,1], main="pca3\_1, by chip.order",pch=19,cex=0.7)

points(rank(ordered.sample644$chip.order)[filter],pca3[filter,1], col=2,pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca3[,2] , main="pca3\_2, by chip.order",pch=19,cex=0.7)

points(rank(ordered.sample644$chip.order)[filter],pca3[filter,2], col=2,pch=19,cex=0.7)

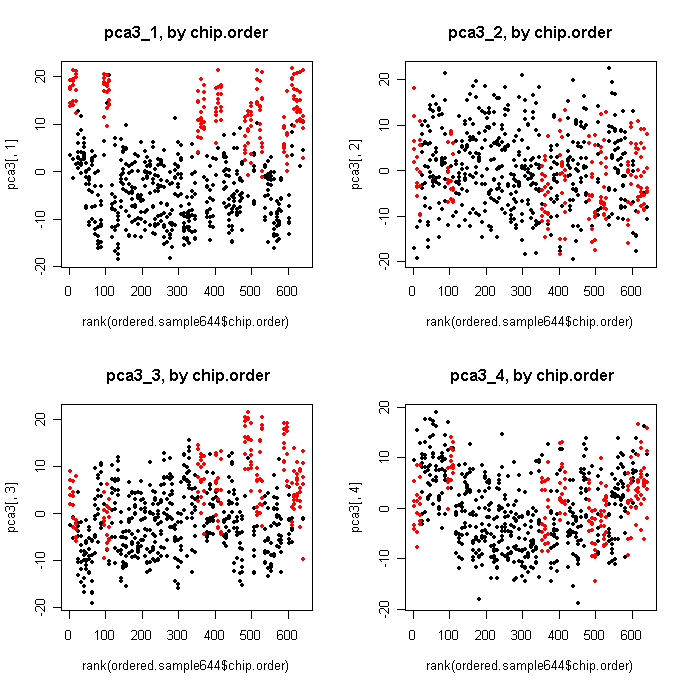
plot(rank(ordered.sample644$chip.order),pca3[,3] , main=" pca3\_3, by chip.order",pch=19,cex=0.7)

points(rank(ordered.sample644$chip.order) [filter],pca3[filter,3] , col=2,pch=19,cex=0.7)

plot(rank(ordered.sample644$chip.order),pca3[,4] , main=" pca3\_4, by chip.order",pch=19,cex=0.7)

points(rank(ordered.sample644$chip.order) [filter],pca3[filter,4] , col=2,pch=19,cex=0.7)

par(mfrow=c(1,1))



par(mfrow=c(2,2))

plot( (ordered.sample644$chip.order),pca3[,1], main="pca3\_1, by chip.order",pch=19,cex=0.7)

points( (ordered.sample644$chip.order)[filter],pca3[filter,1], col=2,pch=19,cex=0.7)

plot( (ordered.sample644$chip.order),pca3[,2] , main="pca3\_2, by chip.order",pch=19,cex=0.7)

points( (ordered.sample644$chip.order)[filter],pca3[filter,2], col=2,pch=19,cex=0.7)

plot( (ordered.sample644$chip.order),pca3[,3] , main=" pca3\_3, by chip.order",pch=19,cex=0.7)

points( (ordered.sample644$chip.order) [filter],pca3[filter,3] , col=2,pch=19,cex=0.7)

plot( (ordered.sample644$chip.order),pca3[,4] , main=" pca3\_4, by chip.order",pch=19,cex=0.7)

points( (ordered.sample644$chip.order) [filter],pca3[filter,4] , col=2,pch=19,cex=0.7)

par(mfrow=c(1,1))

5. case-control comparison

ttest.disease<-function(data,sample,x) {

tmp.mat<-data[,(sample$Disease=="C")|(sample$Disease==x)]

tmp.id<- sample$Disease[(sample$Disease=="C")|(sample$Disease==x)]

tmp.cl <- ifelse(tmp.id=="C", 0, 1)

tmp.test <- mt.teststat(tmp.mat, tmp.cl)

tmp.fc<-apply(tmp.mat[,(tmp.cl==1)],1,mean)-apply(tmp.mat[,(tmp.cl==0)],1,mean)

tmp.combine<-cbind(tmp.test,tmp.fc)

return(tmp.combine)

}

5.1 based on medctr.6

norm.amy<-medctr.6[,1:98]

amy.sample<-ordered.sample644[1:98,]

t.amy.bp<-ttest.disease(norm.amy,amy.sample,"BP")

t.amy.md<-ttest.disease(norm.amy,amy.sample,"MD")

t.amy.sc<-ttest.disease(norm.amy,amy.sample,"SC")

tb.amy <-read.delim("t2\_fc\_amy.txt",header=1,row.names=1,sep="\t")#confirmed they are almost the same

5.2 based on medctr.6.85

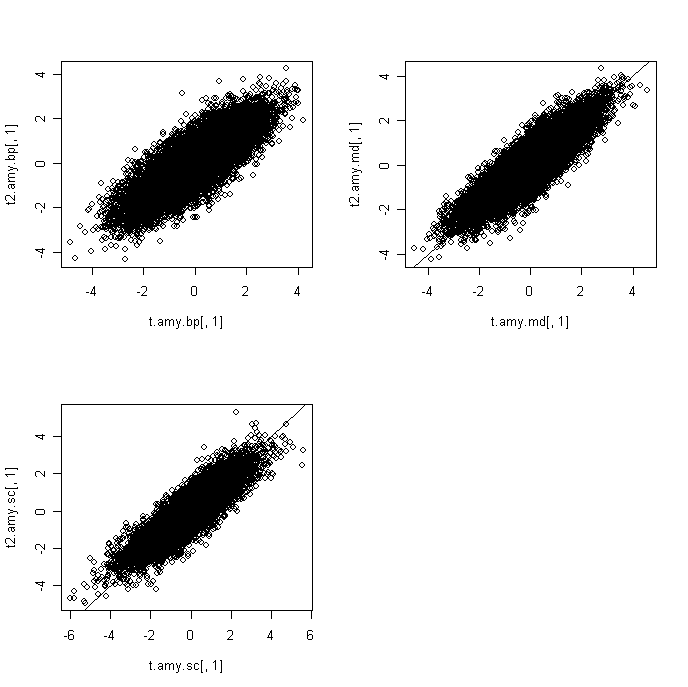
norm.amy<-medctr.6.85[,1:98]

amy.sample<-ordered.sample644[1:98,]

t2.amy.bp<-ttest.disease(norm.amy,amy.sample,"BP")

t2.amy.md<-ttest.disease(norm.amy,amy.sample,"MD")

t2.amy.sc<-ttest.disease(norm.amy,amy.sample,"SC")



turn t into rawp

rawp2.amy.bp<-2\*(1-pt(abs(t2.amy.bp[,1]),53))

rawp2.amy.md<-2\*(1-pt(abs(t2.amy.md[,1]),74))

rawp2.amy.sc<-2\*(1-pt(abs(t2.amy.sc[,1]),51))

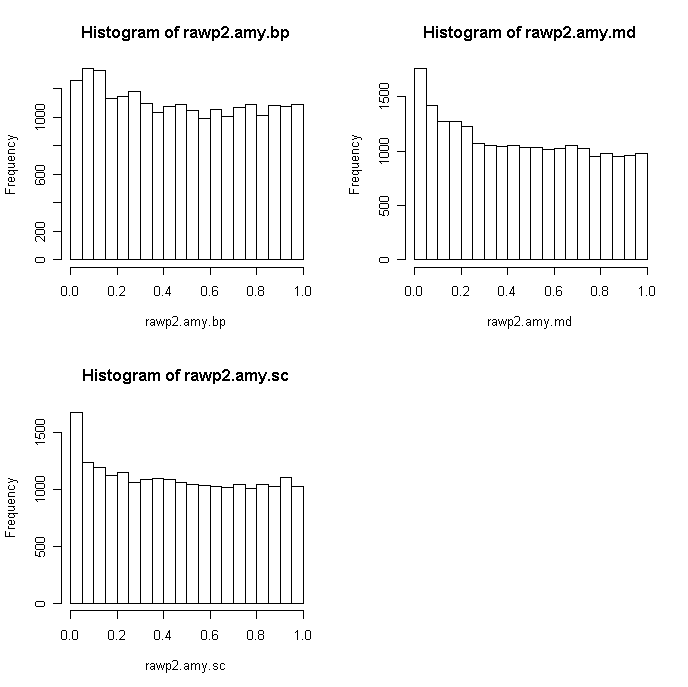
par(mfrow=c(2,2))

hist(rawp2.amy.bp)

hist(rawp2.amy.md)

hist(rawp2.amy.sc)

par(mfrow=c(1,1))



QQplot: 53 74 51

par(mfrow=c(2,2))

tmp<- rt(22177, df = 53)

qqplot(tmp,t2.amy.bp[,1],xlab="Expected t", ylab="Observed t (Amy-BP)")

abline(0,1)

tmp<- rt(22177, df = 74)

qqplot(tmp,t2.amy.md[,1] ,xlab="Expected t", ylab="Observed t (Amy-MD)")

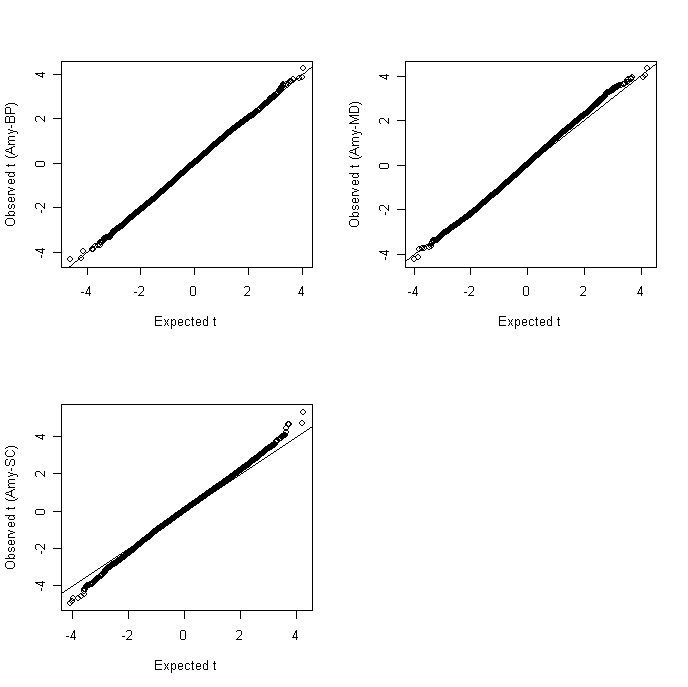
abline(0,1)

tmp<- rt(22177, df = 51)

qqplot(tmp,t2.amy.sc[,1] ,xlab="Expected t", ylab="Observed t (Amy-SC)")

abline(0,1)

par(mfrow=c(1,1))



FDR

procs<-c("BH","BY")

res2.amy.bp<-mt.rawp2adjp(rawp2.amy.bp,procs)

adjp2.amy.bp<-res2.amy.bp$adjp[order(res2.amy.bp$index),]

res2.amy.md<-mt.rawp2adjp(rawp2.amy.md,procs)

adjp2.amy.md<-res2.amy.md$adjp[order(res2.amy.md$index),]

res2.amy.sc<-mt.rawp2adjp(rawp2.amy.sc,procs)

adjp2.amy.sc<-res2.amy.sc$adjp[order(res2.amy.sc$index),]

par(mfrow=c(2,2))

plot(rawp2.amy.bp,adjp2.amy.bp[,2],log="xy",main="BH(amy.bp)")

points(rawp2.amy.bp[med.p.amy<0.05],adjp2.amy.bp[med.p.amy<0.05,2],col=2)

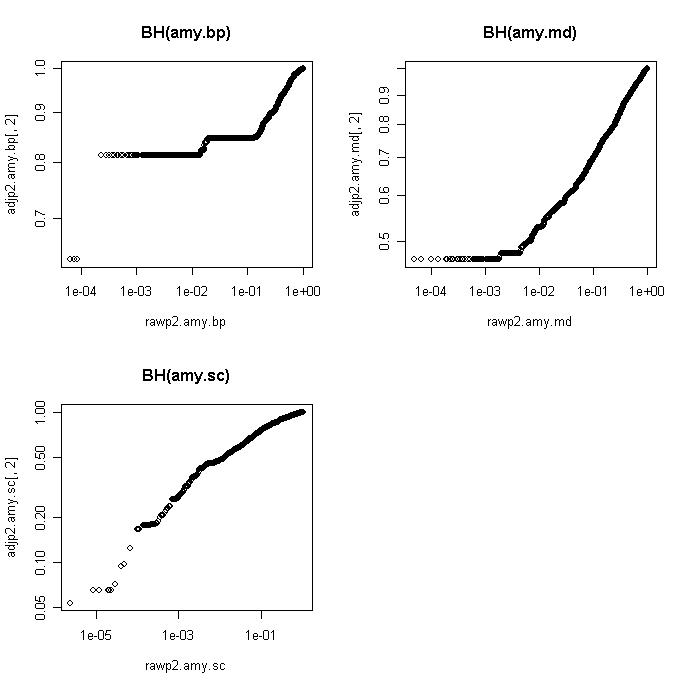
plot(rawp2.amy.md,adjp2.amy.md[,2],log="xy",main="BH(amy.md)")

points(rawp2.amy.md[med.p.amy<0.05],adjp2.amy.md[med.p.amy<0.05,2], col=2)

plot(rawp2.amy.sc,adjp2.amy.sc[,2],log="xy",main="BH(amy.sc)")

points(rawp2.amy.sc[med.p.amy<0.05],adjp2.amy.sc[med.p.amy<0.05,2], col=2)

par(mfrow=c(1,1))



export data

t

t.amy<-as.matrix(cbind(t2.amy.bp,t2.amy.md,t2.amy.sc))

colnames(t.amy)<-c("t.amy.bp","fc.amy.bp","t.amy..md","fc.amy.md","t.amy.sc","fc.amy.sc")

write.table(t.amy,"t\_fc\_amy.txt",sep="\t")

rawp

rawp.amy<-as.matrix(cbind(rawp2.amy.bp,rawp2.amy.md,rawp2.amy.sc))

colnames(rawp.amy)<-c("rawp.amy.bp","rawp.amy..md","rawp.amy.sc")

write.table(rawp.amy,"rawp\_amy.txt",sep="\t")

bh.fdr.amy<-as.matrix(cbind(adjp2.amy.bp[,2], adjp2.amy.md[,2],

adjp2.amy.sc[,2]))

rownames(bh.fdr.amy)<-rownames(norm.amy)

colnames(bh.fdr.amy)<-c("fdr.bp","fdr.md","fdr.sc")

write.table(bh.fdr.amy,"BH \_fdr\_amy.txt",sep="\t")

norm.acg<-medctr.6[,99:207]

acg.sample<-ordered.sample644[99:207,]

t.acg.bp<-ttest.disease(norm.acg,acg.sample,"BP")

t.acg.md<-ttest.disease(norm.acg,acg.sample,"MD")

t.acg.sc<-ttest.disease(norm.acg,acg.sample,"SC")

norm.acg<-medctr.6.85[,99:207]

t2.acg.bp<-ttest.disease(norm.acg,acg.sample,"BP")

t2.acg.md<-ttest.disease(norm.acg,acg.sample,"MD")

t2.acg.sc<-ttest.disease(norm.acg,acg.sample,"SC")

tb.acg <-read.delim("t2\_fc\_acg.txt",header=1,row.names=1,sep="\t")

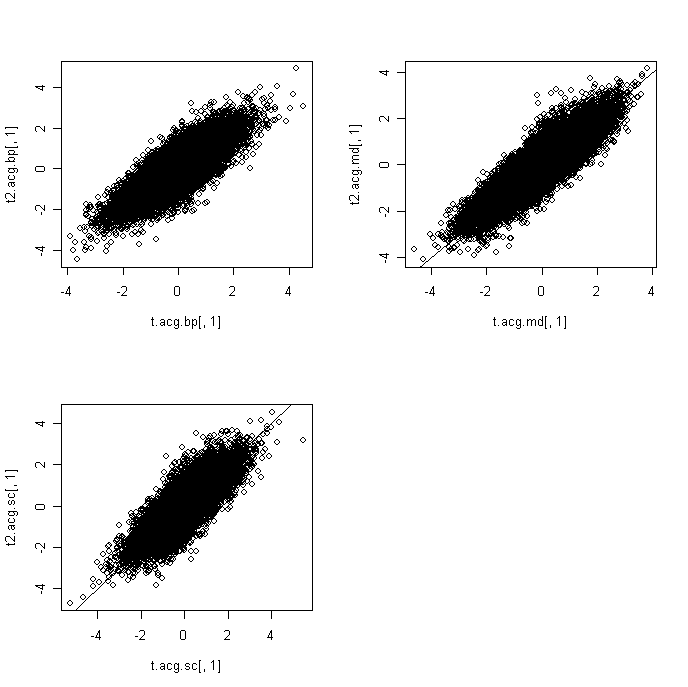
par(mfrow=c(2,2))

plot(tb.acg[,1],t2.acg.bp[,2])

plot(tb.acg[,3],t2.acg.md[,2])

plot(tb.acg[,5],t2.acg.sc[,2])

par(mfrow=c(1,1))



turn t into rawp

rawp2.acg.bp<-2\*(1-pt(abs(t2.acg.bp[,1]),59))

rawp2.acg.md<-2\*(1-pt(abs(t2.acg.md[,1]),79))

rawp2.acg.sc<-2\*(1-pt(abs(t2.acg.sc[,1]),59))

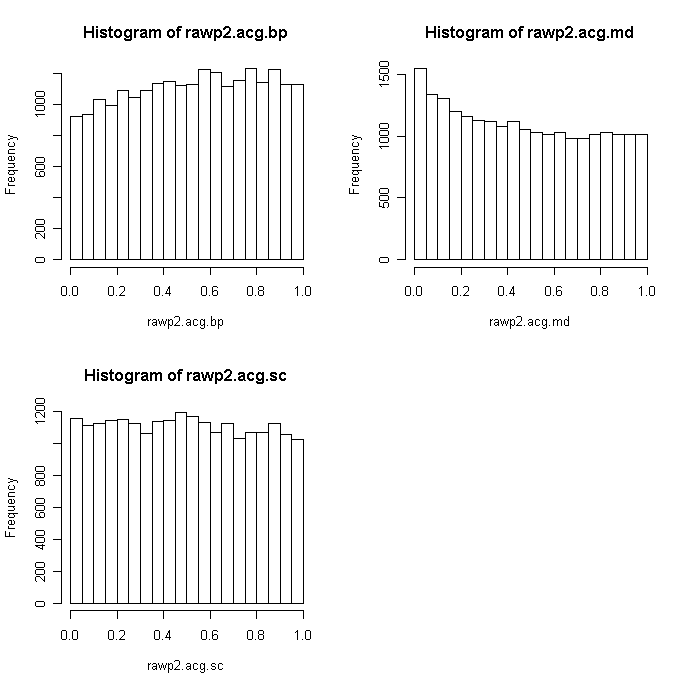
par(mfrow=c(2,2))

hist(rawp2.acg.bp)

hist(rawp2.acg.md)

hist(rawp2.acg.sc)

par(mfrow=c(1,1))



par(mfrow=c(2,2))

tmp<- rt(22177, df = 59)

qqplot(tmp,t2.acg.bp[,1],xlab="Expected t", ylab="Observed t (Acg-BP)")

abline(0,1)

tmp<- rt(22177, df = 79)

qqplot(tmp,t2.acg.md[,1] ,xlab="Expected t", ylab="Observed t (Acg-MD)")

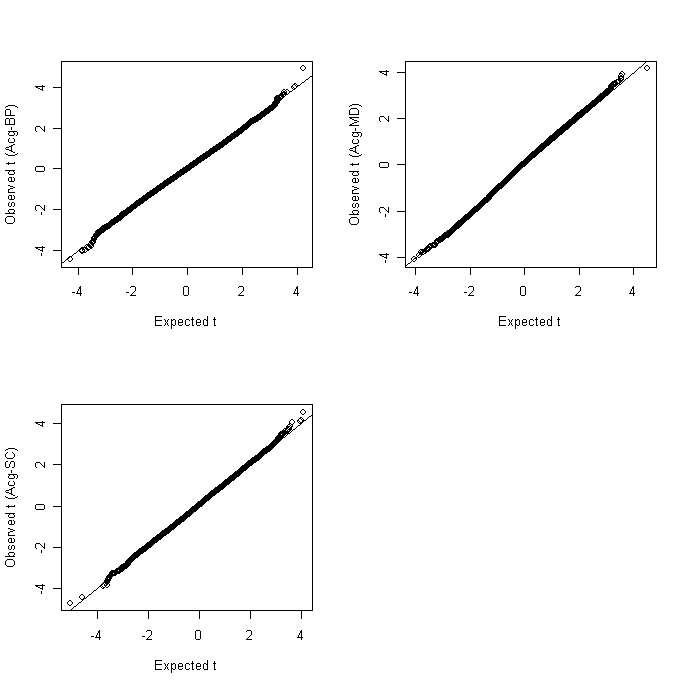
abline(0,1)

tmp<- rt(22177, df = 59)

qqplot(tmp,t2.acg.sc[,1] ,xlab="Expected t", ylab="Observed t (Acg-SC)")

abline(0,1)

par(mfrow=c(1,1))



FDR

res2.acg.bp<-mt.rawp2adjp(rawp2.acg.bp,procs)

adjp2.acg.bp<-res2.acg.bp$adjp[order(res2.acg.bp$index),]

res2.acg.md<-mt.rawp2adjp(rawp2.acg.md,procs)

adjp2.acg.md<-res2.acg.md$adjp[order(res2.acg.md$index),]

res2.acg.sc<-mt.rawp2adjp(rawp2.acg.sc,procs)

adjp2.acg.sc<-res2.acg.sc$adjp[order(res2.acg.sc$index),]

par(mfrow=c(2,2))

plot(rawp2.acg.bp,adjp2.acg.bp[,2],log="xy",main="BH(acg.bp)")

points(rawp2.acg.bp[med.p.acg<0.05],adjp2.acg.bp[med.p.acg<0.05,2],col=2)

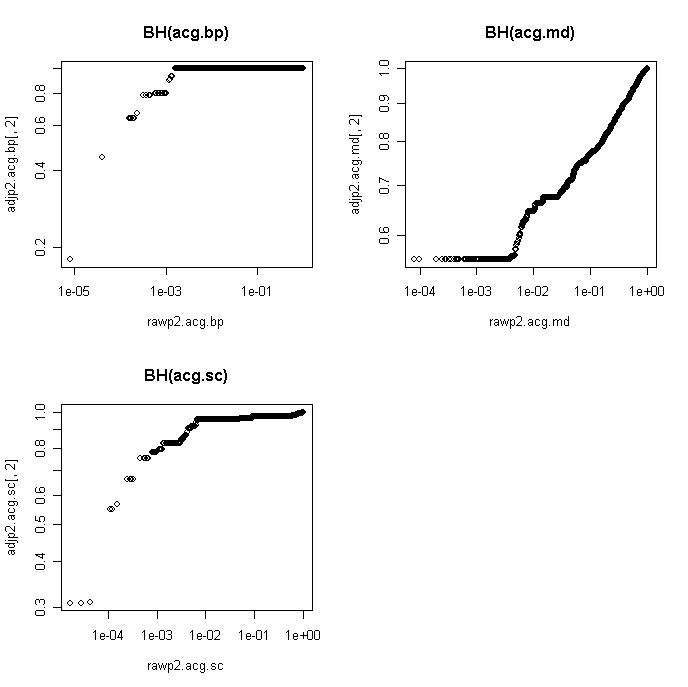
plot(rawp2.acg.md,adjp2.acg.md[,2],log="xy",main="BH(acg.md)")

points(rawp2.acg.md[med.p.acg<0.05],adjp2.acg.md[med.p.acg<0.05,2], col=2)

plot(rawp2.acg.sc,adjp2.acg.sc[,2],log="xy",main="BH(acg.sc)")

points(rawp2.acg.sc[med.p.acg<0.05],adjp2.acg.sc[med.p.acg<0.05,2], col=2)

par(mfrow=c(1,1))



export data

t

t.acg<-as.matrix(cbind(t2.acg.bp,t2.acg.md,t2.acg.sc))

colnames(t.acg)<-c("t.acg.bp","fc.acg.bp","t.acg..md","fc.acg.md","t.acg.sc","fc.acg.sc")

write.table(t.acg,"t\_fc\_acg.txt",sep="\t")

rawp

rawp.acg<-as.matrix(cbind(rawp2.acg.bp,rawp2.acg.md,rawp2.acg.sc))

colnames(rawp.acg)<-c("rawp.acg.bp","rawp.acg..md","rawp.acg.sc")

write.table(rawp.acg,"rawp\_acg.txt",sep="\t")

bh.fdr.acg<-as.matrix(cbind(adjp2.acg.bp[,2], adjp2.acg.md[,2],

adjp2.acg.sc[,2]))

rownames(bh.fdr.acg)<-rownames(norm.acg)

colnames(bh.fdr.acg)<-c("fdr.bp","fdr.md","fdr.sc")

write.table(bh.fdr.acg,"BH \_fdr\_acg.txt",sep="\t")

norm.cb<-medctr.6[,208:320]

cb.sample<-ordered.sample644[208:320,]

t.cb.bp<-ttest.disease(norm.cb,cb.sample,"BP")

t.cb.md<-ttest.disease(norm.cb,cb.sample,"MD")

t.cb.sc<-ttest.disease(norm.cb,cb.sample,"SC")

norm.cb<-medctr.6.85[,208:320]

t2.cb.bp<-ttest.disease(norm.cb,cb.sample,"BP")

t2.cb.md<-ttest.disease(norm.cb,cb.sample,"MD")

t2.cb.sc<-ttest.disease(norm.cb,cb.sample,"SC")

tb.cb <-read.delim("t2\_fc\_cb.txt",header=1,row.names=1,sep="\t")

par(mfrow=c(2,2))

plot(t.cb.bp[,1],t2.cb.bp[,1])

abline(0,1)

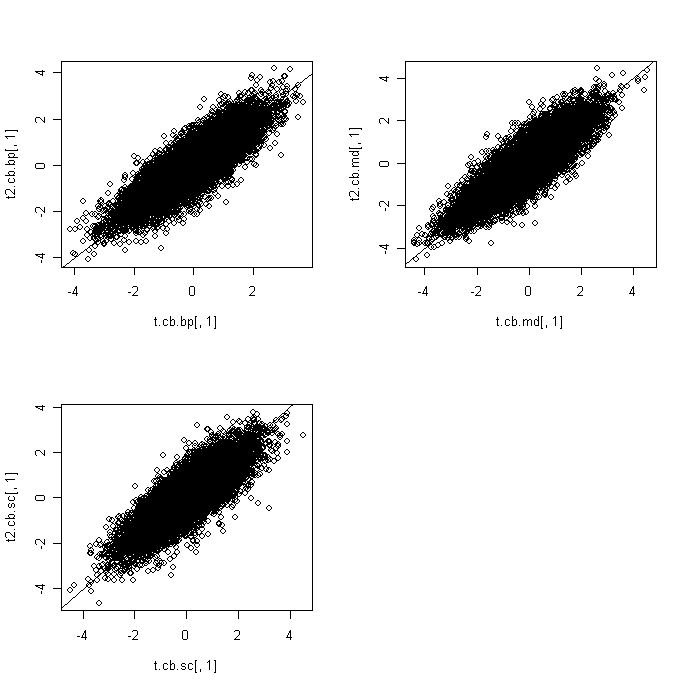
plot(t.cb.md[,1],t2.cb.md[,1])

abline(0,1)

plot(t.cb.sc[,1],t2.cb.sc[,1])

abline(0,1)

par(mfrow=c(1,1))



turn t into rawp

rawp2.cb.bp<-2\*(1-pt(abs(t2.cb.bp[,1]),61))

rawp2.cb.md<-2\*(1-pt(abs(t2.cb.md[,1]),81))

rawp2.cb.sc<-2\*(1-pt(abs(t2.cb.sc[,1]),59))

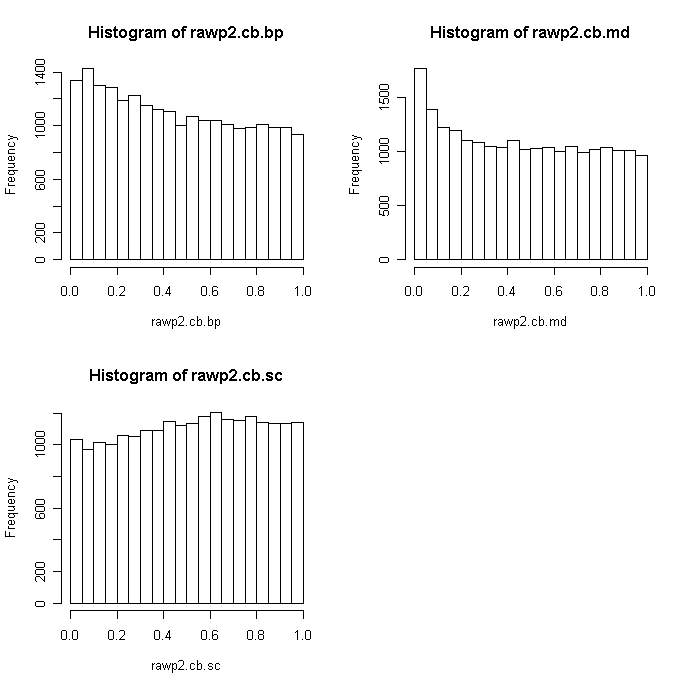
par(mfrow=c(2,2))

hist(rawp2.cb.bp)

hist(rawp2.cb.md)

hist(rawp2.cb.sc)

par(mfrow=c(1,1))



par(mfrow=c(2,2))

tmp<- rt(22177, df = 61)

qqplot(tmp,t2.cb.bp[,1],xlab="Expected t", ylab="Observed t (Cb-BP)")

abline(0,1)

tmp<- rt(22177, df = 81)

qqplot(tmp,t2.cb.md[,1] ,xlab="Expected t", ylab="Observed t (Cb-MD)")

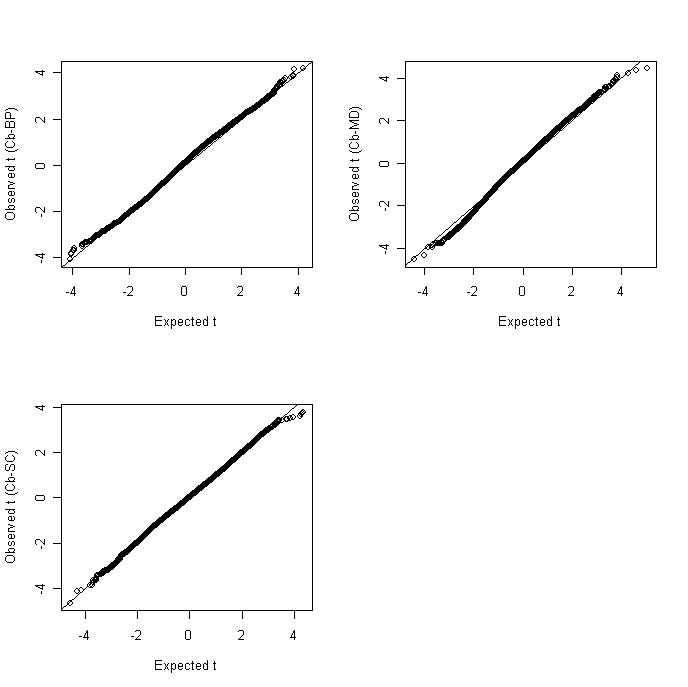
abline(0,1)

tmp<- rt(22177, df = 59)

qqplot(tmp,t2.cb.sc[,1] ,xlab="Expected t", ylab="Observed t (Cb-SC)")

abline(0,1)

par(mfrow=c(1,1))



FDR

res2.cb.bp<-mt.rawp2adjp(rawp2.cb.bp,procs)

adjp2.cb.bp<-res2.cb.bp$adjp[order(res2.cb.bp$index),]

res2.cb.md<-mt.rawp2adjp(rawp2.cb.md,procs)

adjp2.cb.md<-res2.cb.md$adjp[order(res2.cb.md$index),]

res2.cb.sc<-mt.rawp2adjp(rawp2.cb.sc,procs)

adjp2.cb.sc<-res2.cb.sc$adjp[order(res2.cb.sc$index),]

par(mfrow=c(2,2))

plot(rawp2.cb.bp,adjp2.cb.bp[,2],log="xy",main="BH(cb.bp)")

points(rawp2.cb.bp[med.p.cb<0.05],adjp2.cb.bp[med.p.cb<0.05,2],col=2)

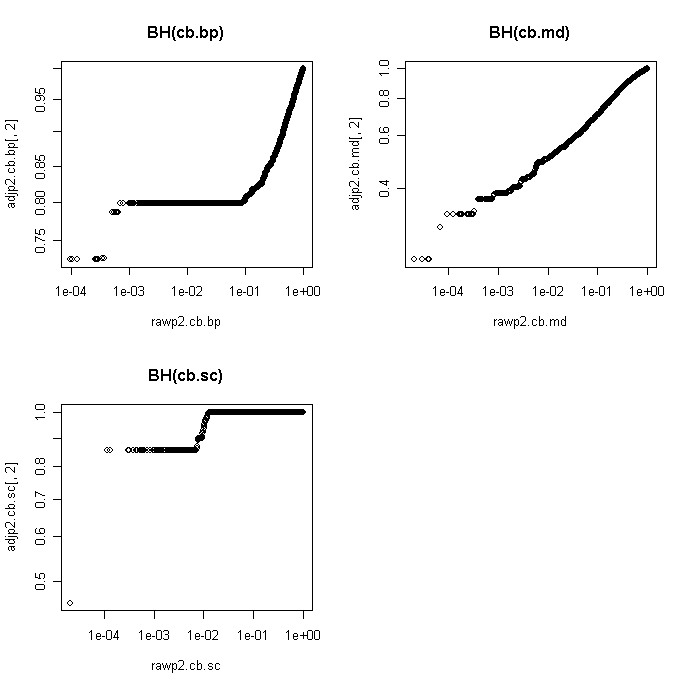
plot(rawp2.cb.md,adjp2.cb.md[,2],log="xy",main="BH(cb.md)")

points(rawp2.cb.md[med.p.cb<0.05],adjp2.cb.md[med.p.cb<0.05,2], col=2)

plot(rawp2.cb.sc,adjp2.cb.sc[,2],log="xy",main="BH(cb.sc)")

points(rawp2.cb.sc[med.p.cb<0.05],adjp2.cb.sc[med.p.cb<0.05,2], col=2)

par(mfrow=c(1,1))



export data

t

t.cb<-as.matrix(cbind(t2.cb.bp,t2.cb.md,t2.cb.sc))

colnames(t.cb)<-c("t.cb.bp","fc.cb.bp","t.cb..md","fc.cb.md","t.cb.sc","fc.cb.sc")

write.table(t.cb,"t\_fc\_cb.txt",sep="\t")

rawp

rawp.cb<-as.matrix(cbind(rawp2.cb.bp,rawp2.cb.md,rawp2.cb.sc))

colnames(rawp.cb)<-c("rawp.cb.bp","rawp.cb..md","rawp.cb.sc")

write.table(rawp.cb,"rawp\_cb.txt",sep="\t")

bh.fdr.cb<-as.matrix(cbind(adjp2.cb.bp[,2], adjp2.cb.md[,2],

adjp2.cb.sc[,2]))

rownames(bh.fdr.cb)<-rownames(norm.cb)

colnames(bh.fdr.cb)<-c("fdr.bp","fdr.md","fdr.sc")

write.table(bh.fdr.cb,"BH \_fdr\_cb.txt",sep="\t")

norm.dlpfc<-medctr.6[,321:429]

dlpfc.sample<-ordered.sample644[321:429,]

t.dlpfc.bp<-ttest.disease(norm.dlpfc,dlpfc.sample,"BP")

t.dlpfc.md<-ttest.disease(norm.dlpfc,dlpfc.sample,"MD")

t.dlpfc.sc<-ttest.disease(norm.dlpfc,dlpfc.sample,"SC")

norm.dlpfc<-medctr.6.85[,321:429]

t2.dlpfc.bp<-ttest.disease(norm.dlpfc,dlpfc.sample,"BP")

t2.dlpfc.md<-ttest.disease(norm.dlpfc,dlpfc.sample,"MD")

t2.dlpfc.sc<-ttest.disease(norm.dlpfc,dlpfc.sample,"SC")

tb.dlpfc <-read.delim("t2\_fc\_dlpfc.txt",header=1,row.names=1,sep="\t")

par(mfrow=c(2,2))

plot(t.dlpfc.bp[,1],t2.dlpfc.bp[,1])

abline(0,1)

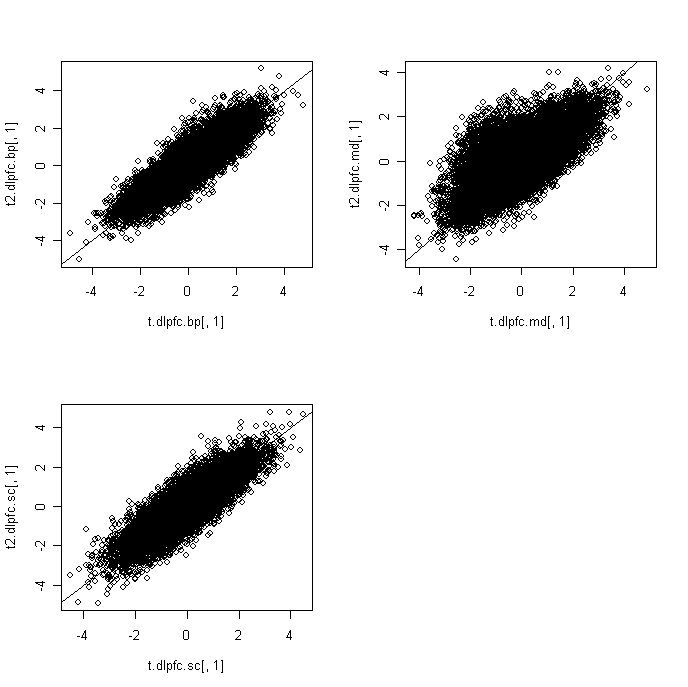
plot(t.dlpfc.md[,1],t2.dlpfc.md[,1])

abline(0,1)

plot(t.dlpfc.sc[,1],t2.dlpfc.sc[,1])

abline(0,1)

par(mfrow=c(1,1))



turn t into rawp

rawp2.dlpfc.bp<-2\*(1-pt(abs(t2.dlpfc.bp[,1]),60))

rawp2.dlpfc.md<-2\*(1-pt(abs(t2.dlpfc.md[,1]),77))

rawp2.dlpfc.sc<-2\*(1-pt(abs(t2.dlpfc.sc[,1]),58))

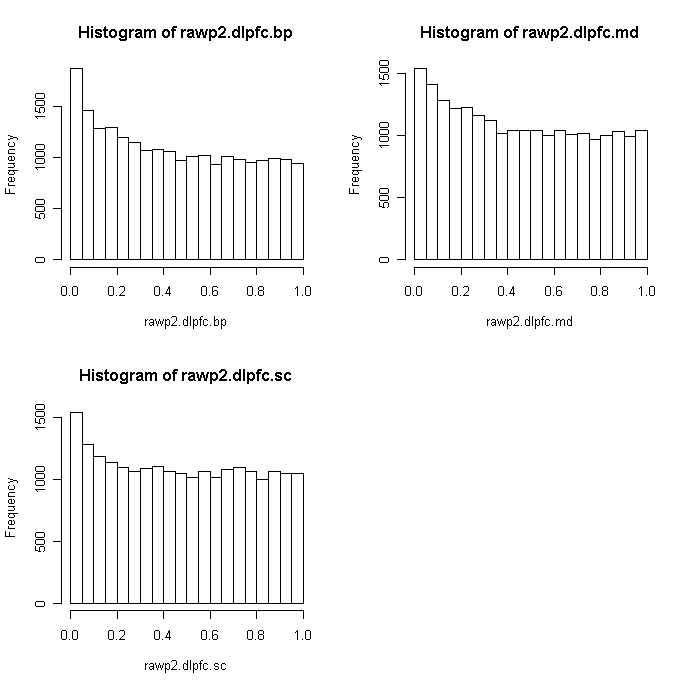
par(mfrow=c(2,2))

hist(rawp2.dlpfc.bp)

hist(rawp2.dlpfc.md)

hist(rawp2.dlpfc.sc)

par(mfrow=c(1,1))



par(mfrow=c(2,2))

tmp<- rt(22177, df = 60)

qqplot(tmp,t2.dlpfc.bp[,1],xlab="Expected t", ylab="Observed t (Dlpfc-BP)")

abline(0,1)

tmp<- rt(22177, df = 77)

qqplot(tmp,t2.dlpfc.md[,1] ,xlab="Expected t", ylab="Observed t (Dlpfc-MD)")

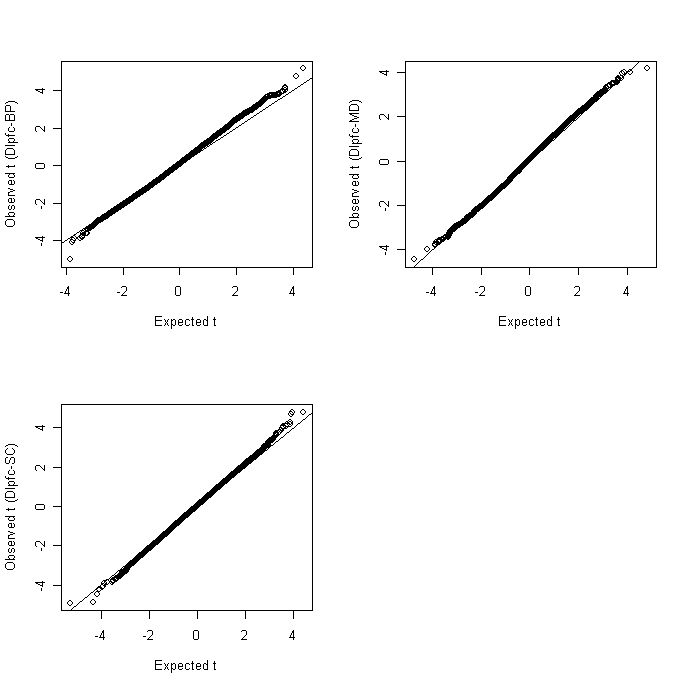
abline(0,1)

tmp<- rt(22177, df = 58)

qqplot(tmp,t2.dlpfc.sc[,1] ,xlab="Expected t", ylab="Observed t (Dlpfc-SC)")

abline(0,1)

par(mfrow=c(1,1))



FDR

res2.dlpfc.bp<-mt.rawp2adjp(rawp2.dlpfc.bp,procs)

adjp2.dlpfc.bp<-res2.dlpfc.bp$adjp[order(res2.dlpfc.bp$index),]

res2.dlpfc.md<-mt.rawp2adjp(rawp2.dlpfc.md,procs)

adjp2.dlpfc.md<-res2.dlpfc.md$adjp[order(res2.dlpfc.md$index),]

res2.dlpfc.sc<-mt.rawp2adjp(rawp2.dlpfc.sc,procs)

adjp2.dlpfc.sc<-res2.dlpfc.sc$adjp[order(res2.dlpfc.sc$index),]

par(mfrow=c(2,2))

plot(rawp2.dlpfc.bp,adjp2.dlpfc.bp[,2],log="xy",main="BH(dlpfc.bp)")

points(rawp2.dlpfc.bp[med.p.dlpfc<0.05],adjp2.dlpfc.bp[med.p.dlpfc<0.05,2],col=2)

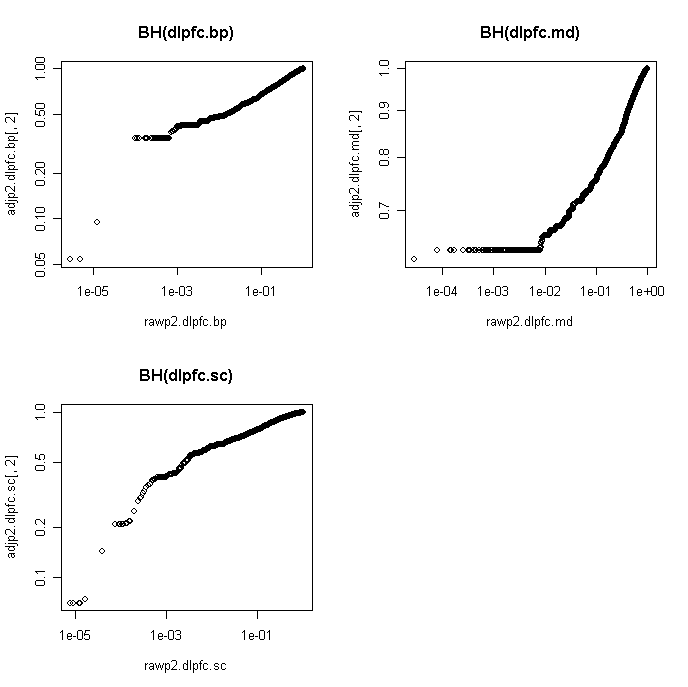
plot(rawp2.dlpfc.md,adjp2.dlpfc.md[,2],log="xy",main="BH(dlpfc.md)")

points(rawp2.dlpfc.md[med.p.dlpfc<0.05],adjp2.dlpfc.md[med.p.dlpfc<0.05,2], col=2)

plot(rawp2.dlpfc.sc,adjp2.dlpfc.sc[,2],log="xy",main="BH(dlpfc.sc)")

points(rawp2.dlpfc.sc[med.p.dlpfc<0.05],adjp2.dlpfc.sc[med.p.dlpfc<0.05,2], col=2)

par(mfrow=c(1,1))



export data

t

t.dlpfc<-as.matrix(cbind(t2.dlpfc.bp,t2.dlpfc.md,t2.dlpfc.sc))

colnames(t.dlpfc)<-c("t.dlpfc.bp","fc.dlpfc.bp","t.dlpfc..md","fc.dlpfc.md","t.dlpfc.sc","fc.dlpfc.sc")

write.table(t.dlpfc,"t\_fc\_dlpfc.txt",sep="\t")

rawp

rawp.dlpfc<-as.matrix(cbind(rawp2.dlpfc.bp,rawp2.dlpfc.md,rawp2.dlpfc.sc))

colnames(rawp.dlpfc)<-c("rawp.dlpfc.bp","rawp.dlpfc..md","rawp.dlpfc.sc")

write.table(rawp.dlpfc,"rawp\_dlpfc.txt",sep="\t")

bh.fdr.dlpfc<-as.matrix(cbind(adjp2.dlpfc.bp[,2], adjp2.dlpfc.md[,2],

adjp2.dlpfc.sc[,2]))

rownames(bh.fdr.dlpfc)<-rownames(norm.dlpfc)

colnames(bh.fdr.dlpfc)<-c("fdr.bp","fdr.md","fdr.sc")

write.table(bh.fdr.dlpfc,"BH \_fdr\_dlpfc.txt",sep="\t")

norm.hc<-medctr.6[,430:538]

hc.sample<-ordered.sample644[430:538,]

t.hc.bp<-ttest.disease(norm.hc,hc.sample,"BP")

t.hc.md<-ttest.disease(norm.hc,hc.sample,"MD")

t.hc.sc<-ttest.disease(norm.hc,hc.sample,"SC")

norm.hc<-medctr.6.85[,430:538]

t2.hc.bp<-ttest.disease(norm.hc,hc.sample,"BP")

t2.hc.md<-ttest.disease(norm.hc,hc.sample,"MD")

t2.hc.sc<-ttest.disease(norm.hc,hc.sample,"SC")

tb.hc <-read.delim("t2\_fc\_hc.txt",header=1,row.names=1,sep="\t")

par(mfrow=c(2,2))

plot(t.hc.bp[,1],t2.hc.bp[,1])

abline(0,1)

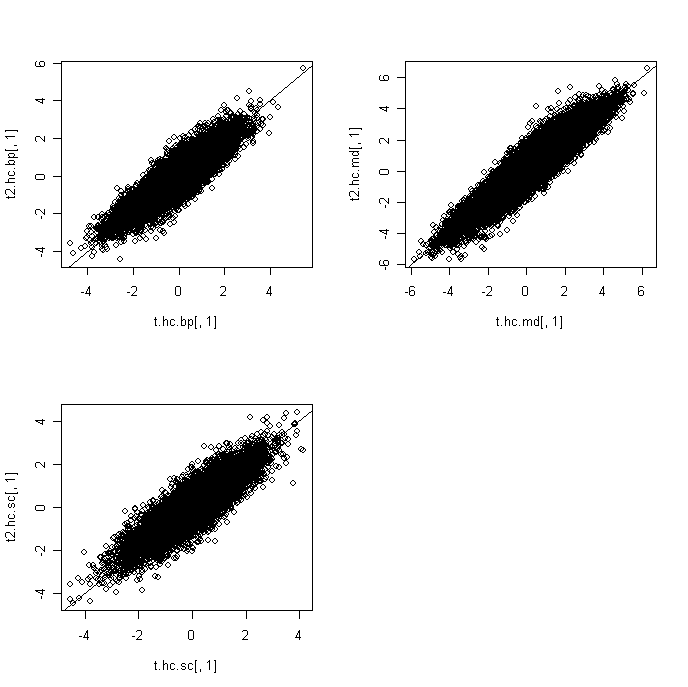
plot(t.hc.md[,1],t2.hc.md[,1])

abline(0,1)

plot(t.hc.sc[,1],t2.hc.sc[,1])

abline(0,1)

par(mfrow=c(1,1))



turn t into rawp

rawp2.hc.bp<-2\*(1-pt(abs(t2.hc.bp[,1]),59))

rawp2.hc.md<-2\*(1-pt(abs(t2.hc.md[,1]),77))

rawp2.hc.sc<-2\*(1-pt(abs(t2.hc.sc[,1]),57))

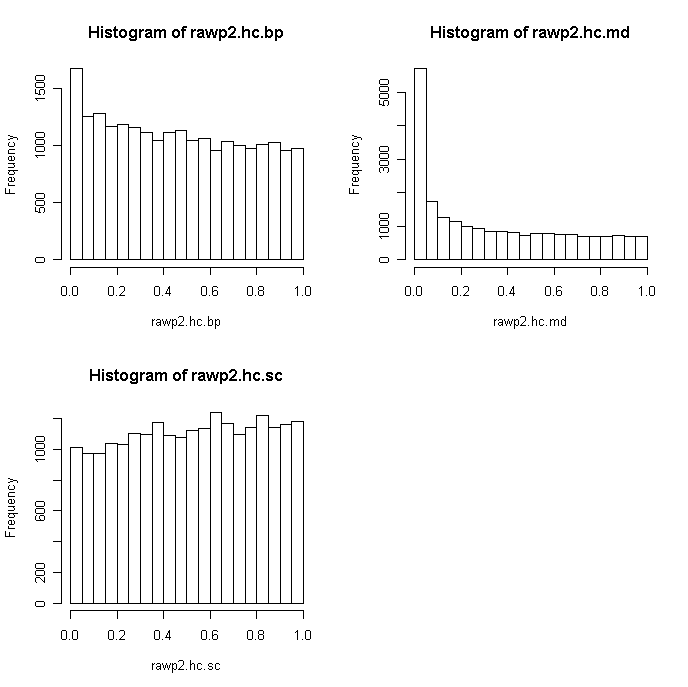
par(mfrow=c(2,2))

hist(rawp2.hc.bp)

hist(rawp2.hc.md)

hist(rawp2.hc.sc)

par(mfrow=c(1,1))



par(mfrow=c(2,2))

tmp<- rt(22177, df = 59)

qqplot(tmp,t2.hc.bp[,1],xlab="Expected t", ylab="Observed t (Hc-BP)")

abline(0,1)

tmp<- rt(22177, df = 77)

qqplot(tmp,t2.hc.md[,1] ,xlab="Expected t", ylab="Observed t (Hc-MD)")

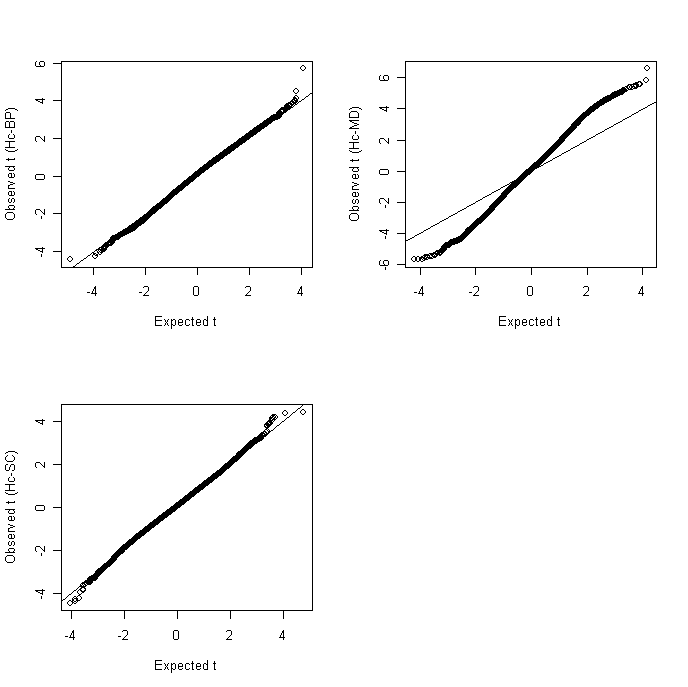
abline(0,1)

tmp<- rt(22177, df = 57)

qqplot(tmp,t2.hc.sc[,1] ,xlab="Expected t", ylab="Observed t (Hc-SC)")

abline(0,1)

par(mfrow=c(1,1))



FDR

res2.hc.bp<-mt.rawp2adjp(rawp2.hc.bp,procs)

adjp2.hc.bp<-res2.hc.bp$adjp[order(res2.hc.bp$index),]

res2.hc.md<-mt.rawp2adjp(rawp2.hc.md,procs)

adjp2.hc.md<-res2.hc.md$adjp[order(res2.hc.md$index),]

res2.hc.sc<-mt.rawp2adjp(rawp2.hc.sc,procs)

adjp2.hc.sc<-res2.hc.sc$adjp[order(res2.hc.sc$index),]

par(mfrow=c(2,2))

plot(rawp2.hc.bp,adjp2.hc.bp[,2],log="xy",main="BH(hc.bp)")

points(rawp2.hc.bp[med.p.hc<0.05],adjp2.hc.bp[med.p.hc<0.05,2],col=2)

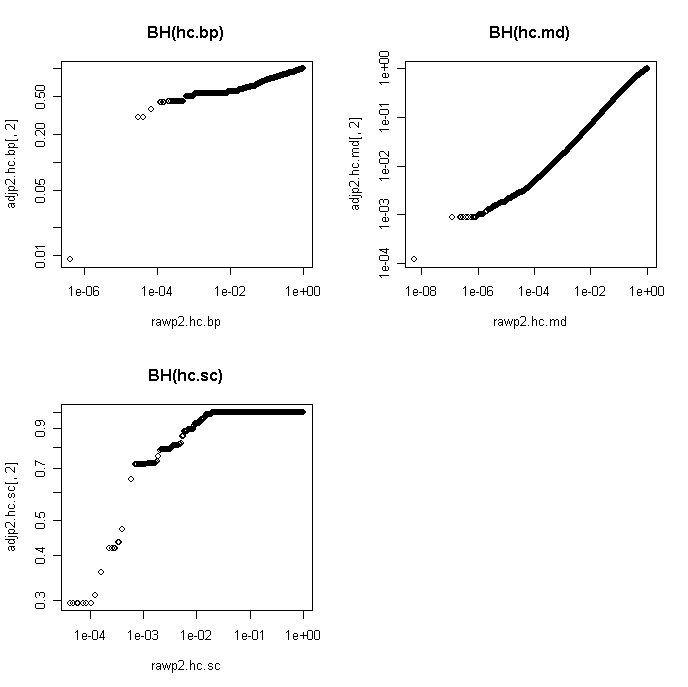
plot(rawp2.hc.md,adjp2.hc.md[,2],log="xy",main="BH(hc.md)")

points(rawp2.hc.md[med.p.hc<0.05],adjp2.hc.md[med.p.hc<0.05,2], col=2)

plot(rawp2.hc.sc,adjp2.hc.sc[,2],log="xy",main="BH(hc.sc)")

points(rawp2.hc.sc[med.p.hc<0.05],adjp2.hc.sc[med.p.hc<0.05,2], col=2)

par(mfrow=c(1,1))



export data

t

t.hc<-as.matrix(cbind(t2.hc.bp,t2.hc.md,t2.hc.sc))

colnames(t.hc)<-c("t.hc.bp","fc.hc.bp","t.hc..md","fc.hc.md","t.hc.sc","fc.hc.sc")

write.table(t.hc,"t\_fc\_hc.txt",sep="\t")

rawp

rawp.hc<-as.matrix(cbind(rawp2.hc.bp,rawp2.hc.md,rawp2.hc.sc))

colnames(rawp.hc)<-c("rawp.hc.bp","rawp.hc..md","rawp.hc.sc")

write.table(rawp.hc,"rawp\_hc.txt",sep="\t")

bh.fdr.hc<-as.matrix(cbind(adjp2.hc.bp[,2], adjp2.hc.md[,2],

adjp2.hc.sc[,2]))

rownames(bh.fdr.hc)<-rownames(norm.hc)

colnames(bh.fdr.hc)<-c("fdr.bp","fdr.md","fdr.sc")

write.table(bh.fdr.hc,"BH \_fdr\_hc.txt",sep="\t")

norm.nacc<-medctr.6[,539:644]

nacc.sample<-ordered.sample644[539:644,]

t.nacc.bp<-ttest.disease(norm.nacc,nacc.sample,"BP")

t.nacc.md<-ttest.disease(norm.nacc,nacc.sample,"MD")

t.nacc.sc<-ttest.disease(norm.nacc,nacc.sample,"SC")

norm.nacc<-medctr.6.85[,539:644]

t2.nacc.bp<-ttest.disease(norm.nacc,nacc.sample,"BP")

t2.nacc.md<-ttest.disease(norm.nacc,nacc.sample,"MD")

t2.nacc.sc<-ttest.disease(norm.nacc,nacc.sample,"SC")

tb.nacc <-read.delim("t2\_fc\_nacc.txt",header=1,row.names=1,sep="\t")

par(mfrow=c(2,2))

plot(t.nacc.bp[,1],t2.nacc.bp[,1])

abline(0,1)

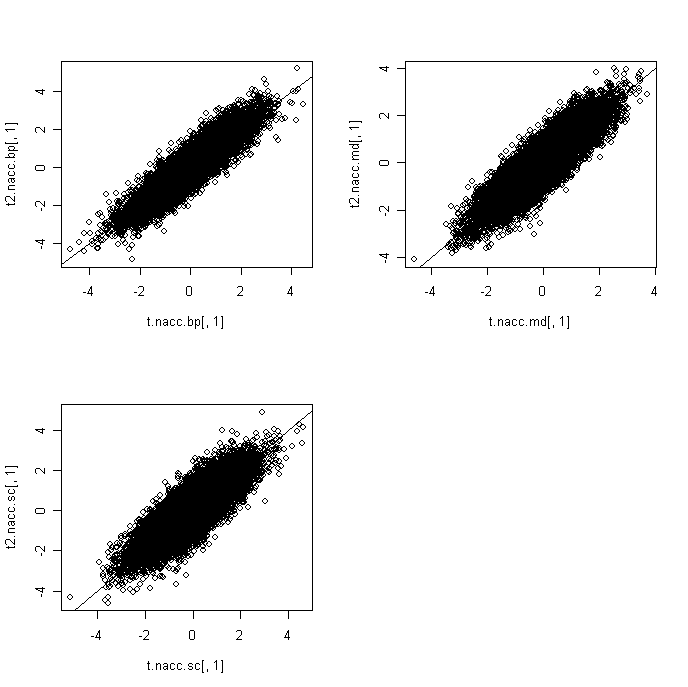
plot(t.nacc.md[,1],t2.nacc.md[,1])

abline(0,1)

plot(t.nacc.sc[,1],t2.nacc.sc[,1])

abline(0,1)

par(mfrow=c(1,1))



turn t into rawp

rawp2.nacc.bp<-2\*(1-pt(abs(t2.nacc.bp[,1]),58))

rawp2.nacc.md<-2\*(1-pt(abs(t2.nacc.md[,1]),75))

rawp2.nacc.sc<-2\*(1-pt(abs(t2.nacc.sc[,1]),57))

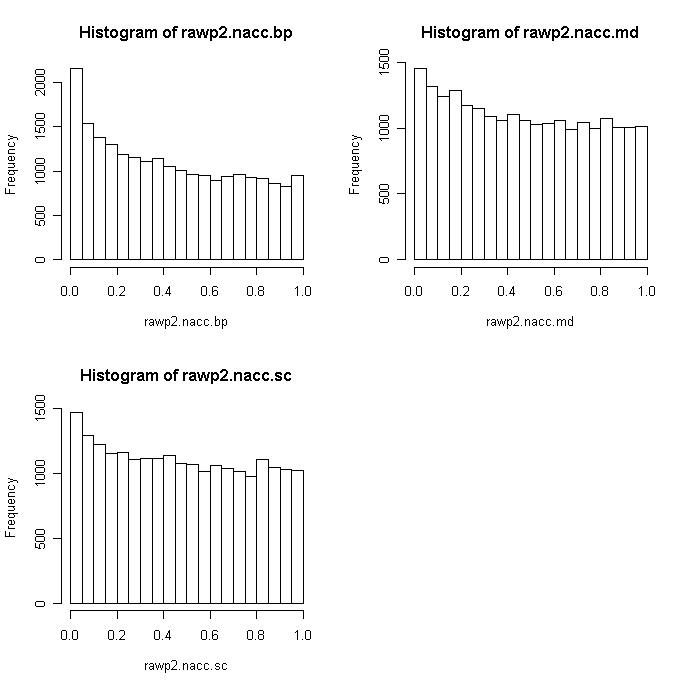
par(mfrow=c(2,2))

hist(rawp2.nacc.bp)

hist(rawp2.nacc.md)

hist(rawp2.nacc.sc)

par(mfrow=c(1,1))



par(mfrow=c(2,2))

tmp<- rt(22177, df = 58)

qqplot(tmp,t2.nacc.bp[,1],xlab="Expected t", ylab="Observed t (Nacc-BP)")

abline(0,1)

tmp<- rt(22177, df = 75)

qqplot(tmp,t2.nacc.md[,1] ,xlab="Expected t", ylab="Observed t (Nacc-MD)")

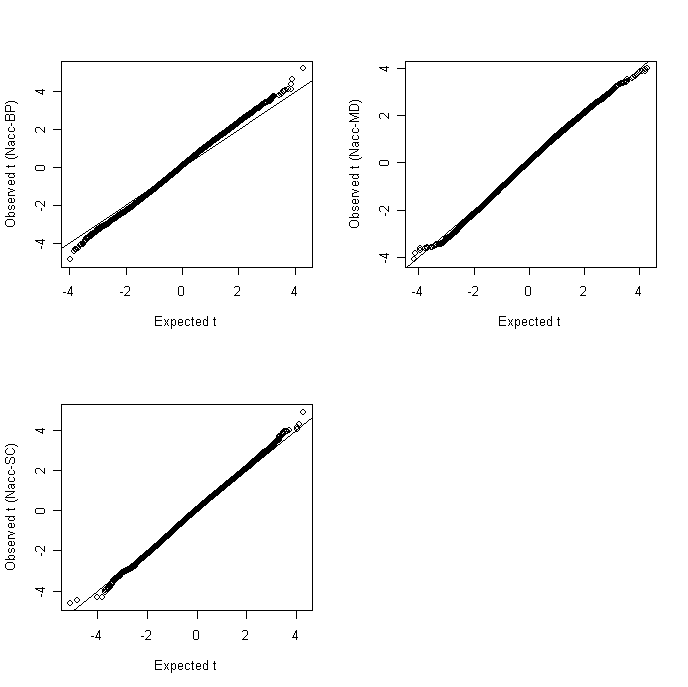
abline(0,1)

tmp<- rt(22177, df = 57)

qqplot(tmp,t2.nacc.sc[,1] ,xlab="Expected t", ylab="Observed t (Nacc-SC)")

abline(0,1)

par(mfrow=c(1,1))



FDR

res2.nacc.bp<-mt.rawp2adjp(rawp2.nacc.bp,procs)

adjp2.nacc.bp<-res2.nacc.bp$adjp[order(res2.nacc.bp$index),]

res2.nacc.md<-mt.rawp2adjp(rawp2.nacc.md,procs)

adjp2.nacc.md<-res2.nacc.md$adjp[order(res2.nacc.md$index),]

res2.nacc.sc<-mt.rawp2adjp(rawp2.nacc.sc,procs)

adjp2.nacc.sc<-res2.nacc.sc$adjp[order(res2.nacc.sc$index),]

par(mfrow=c(2,2))

plot(rawp2.nacc.bp,adjp2.nacc.bp[,2],log="xy",main="BH(nacc.bp)")

points(rawp2.nacc.bp[med.p.nacc<0.05],adjp2.nacc.bp[med.p.nacc<0.05,2],col=2)

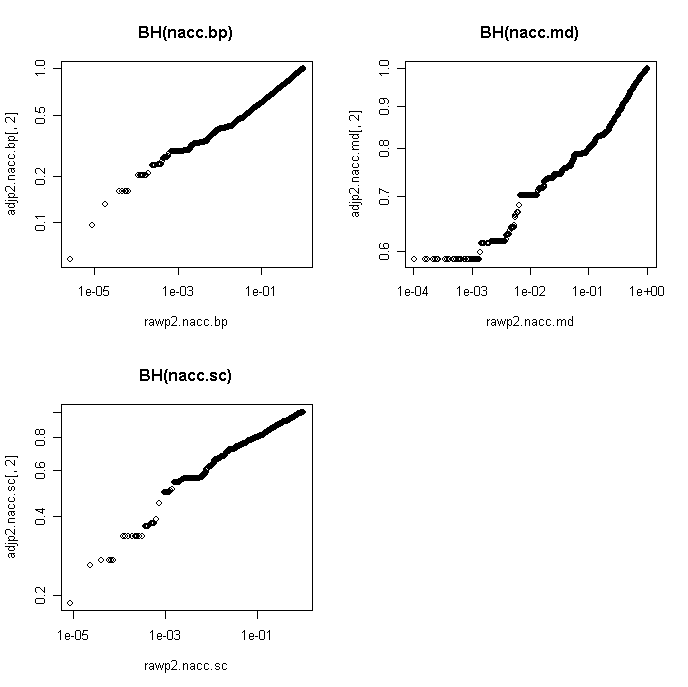
plot(rawp2.nacc.md,adjp2.nacc.md[,2],log="xy",main="BH(nacc.md)")

points(rawp2.nacc.md[med.p.nacc<0.05],adjp2.nacc.md[med.p.nacc<0.05,2], col=2)

plot(rawp2.nacc.sc,adjp2.nacc.sc[,2],log="xy",main="BH(nacc.sc)")

points(rawp2.nacc.sc[med.p.nacc<0.05],adjp2.nacc.sc[med.p.nacc<0.05,2], col=2)

par(mfrow=c(1,1))



export data

t

t.nacc<-as.matrix(cbind(t2.nacc.bp,t2.nacc.md,t2.nacc.sc))

colnames(t.nacc)<-c("t.nacc.bp","fc.nacc.bp","t.nacc..md","fc.nacc.md","t.nacc.sc","fc.nacc.sc")

write.table(t.nacc,"t\_fc\_nacc.txt",sep="\t")

rawp

rawp.nacc<-as.matrix(cbind(rawp2.nacc.bp,rawp2.nacc.md,rawp2.nacc.sc))

colnames(rawp.nacc)<-c("rawp.nacc.bp","rawp.nacc..md","rawp.nacc.sc")

write.table(rawp.nacc,"rawp\_nacc.txt",sep="\t")

bh.fdr.nacc<-as.matrix(cbind(adjp2.nacc.bp[,2], adjp2.nacc.md[,2],

adjp2.nacc.sc[,2]))

rownames(bh.fdr.nacc)<-rownames(norm.nacc)

colnames(bh.fdr.nacc)<-c("fdr.bp","fdr.md","fdr.sc")

write.table(bh.fdr.nacc,"BH \_fdr\_nacc.txt",sep="\t")

> cor(t.amy.bp[,1],t2.amy.bp[,1])

[1] 0.8185788

> cor(t.amy.md[,1],t2.amy.md[,1])

[1] 0.881987

> cor(t.amy.sc[,1],t2.amy.sc[,1])

[1] 0.8823885

cor(t.acg.bp[,1],t2.acg.bp[,1])

[1] 0.81708

> cor(t.acg.md[,1],t2.acg.md[,1])

[1] 0.8826422

> cor(t.acg.sc[,1],t2.acg.sc[,1])

[1] 0.8140165

>

> cor(t.cb.bp[,1],t2.cb.bp[,1])

[1] 0.8705292

> cor(t.cb.md[,1],t2.cb.md[,1])

[1] 0.8625025

> cor(t.cb.sc[,1],t2.cb.sc[,1])

[1] 0.8289543

>

> cor(t.dlpfc.bp[,1],t2.dlpfc.bp[,1])

[1] 0.8781031

> cor(t.dlpfc.md[,1],t2.dlpfc.md[,1])

[1] 0.729295

> cor(t.dlpfc.sc[,1],t2.dlpfc.sc[,1])

[1] 0.8657253

>

> cor(t.hc.bp[,1],t2.hc.bp[,1])

[1] 0.8948656

> cor(t.hc.md[,1],t2.hc.md[,1])

[1] 0.9318283

> cor(t.hc.sc[,1],t2.hc.sc[,1])

[1] 0.88081

>

> cor(t.nacc.bp[,1],t2.nacc.bp[,1])

[1] 0.918541

> cor(t.nacc.md[,1],t2.nacc.md[,1])

[1] 0.8854663

> cor(t.nacc.sc[,1],t2.nacc.sc[,1])

[1] 0.8374088

Comparing batch effect groups

tmp.test <- mt.teststat(medctr.6, as.numeric(filter))

write.table(tmp.test,"t\_chipEffect.txt",sep="\t")

EASE: intracellular transport at 10-8

Cohort

cohort<-matrix(NA,22177,4)

date()

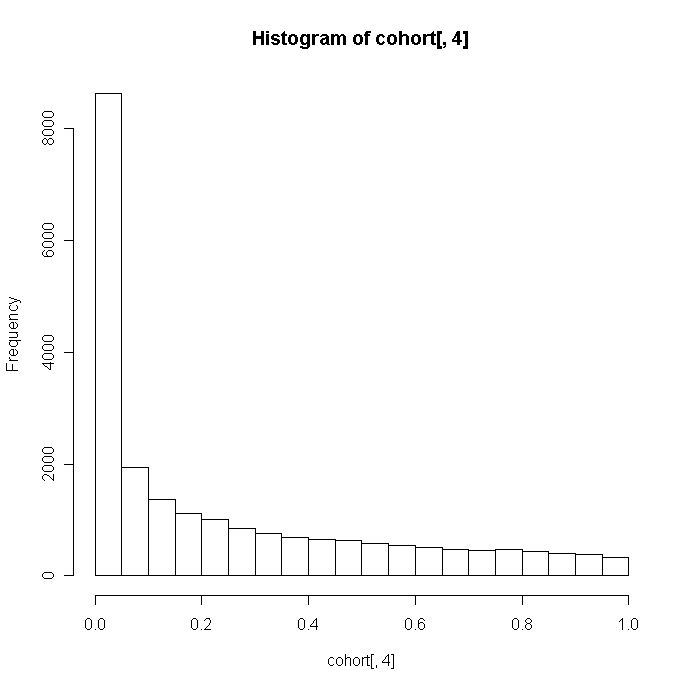
for (i in 1:22177) {

tmp<-anova (lm(medctr.6[i,]~as.factor(ordered.sample644$cohort)))

cohort[i,]<-as.numeric(tmp[1,2:5])

}

date()



cohort2<-matrix(NA,22177,4)

date()

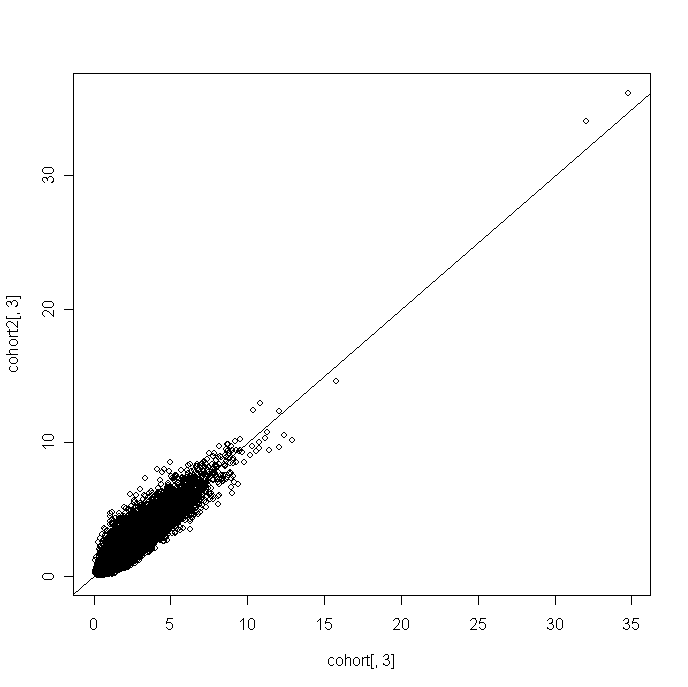
for (i in 1:22177) {

tmp<-anova (lm(medctr.6.85[i,]~as.factor(ordered.sample644$cohort)))

cohort2[i,]<-as.numeric(tmp[1,2:5])

}

date()



PCA for cohort

plot(pca2[,1],pca2[,2])

points(pca2[ordered.sample644$cohort=="dep1",1],pca2[ordered.sample644$cohort=="dep1",2],col=2,pch=19)

points(pca2[ordered.sample644$cohort=="dep2",1],pca2[ordered.sample644$cohort=="dep2",2],col=3,pch=19)

points(pca2[ordered.sample644$cohort=="dep3",1],pca2[ordered.sample644$cohort=="dep3",2],col=4,pch=19)

points(pca2[ordered.sample644$cohort=="dep4",1],pca2[ordered.sample644$cohort=="dep4",2],col=5,pch=19)

points(pca2[ordered.sample644$cohort=="dep5",1],pca2[ordered.sample644$cohort=="dep5",2],col=6,pch=19)

points(pca2[ordered.sample644$cohort=="dep6",1],pca2[ordered.sample644$cohort=="dep6",2],col=2,pch=19)

points(pca2[ordered.sample644$cohort=="dep7",1],pca2[ordered.sample644$cohort=="dep7",2],col=3,pch=19)

points(pca2[ordered.sample644$cohort=="scz1",1],pca2[ordered.sample644$cohort=="scz1",2],col=4,pch=19)

points(pca2[ordered.sample644$cohort=="scz2",1],pca2[ordered.sample644$cohort=="scz2",2],col=5,pch=19)

not seeing anything

anova(lm(pca2[,1]~as.factor(ordered.sample644$cohort)))

Df Sum Sq Mean Sq F value Pr(>F)

as.factor(ordered.sample644$cohort) 9 5298 589 7.6627 9.976e-11 \*\*\*

plot(density(pca2[ordered.sample644$cohort=="dep1",1]))

> lines(density(pca2[ordered.sample644$cohort=="dep2",1]),col=2)

> lines(density(pca2[ordered.sample644$cohort=="dep3",1]),col=3)

> lines(density(pca2[ordered.sample644$cohort=="dep4",1]),col=4)

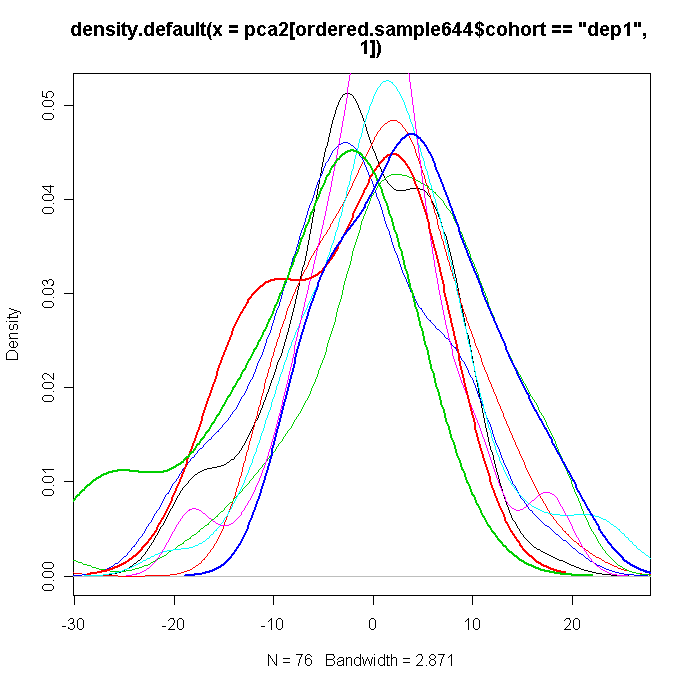
> lines(density(pca2[ordered.sample644$cohort=="dep5",1]),col=5)

> lines(density(pca2[ordered.sample644$cohort=="dep6",1]),col=6)

> lines(density(pca2[ordered.sample644$cohort=="dep7",1]),col=2,lwd=2)

> lines(density(pca2[ordered.sample644$cohort=="scz1",1]),col=3,lwd=2)

> lines(density(pca2[ordered.sample644$cohort=="scz2",1]),col=4,lwd=2)



plot(pca2[order(ordered.sample644$cohort),1])

#not seeing much

anova(lm(pca2[,1]~as.factor(ordered.sample644$chip)))

label,hyb and position

label<-hyb<-pos<-matrix(NA,22177,4)

date()

for (i in 1:22177) {

tmp<-anova (lm(medctr.6[i,]~as.factor(ordered.sample644$label)))

label[i,]<-as.numeric(tmp[1,2:5])

tmp<-anova (lm(medctr.6[i,]~as.factor(ordered.sample644$hyb)))

hyb[i,]<-as.numeric(tmp[1,2:5])

tmp<-anova (lm(medctr.6[i,]~as.factor(ordered.sample644$pos)))

pos[i,]<-as.numeric(tmp[1,2:5])

}

date()

label2<-hyb2<-pos2<-matrix(NA,22177,4)

date()

for (i in 1:22177) {

tmp<-anova (lm(medctr.6.85[i,]~as.factor(ordered.sample644$label)))

label2[i,]<-as.numeric(tmp[1,2:5])

tmp<-anova (lm(medctr.6.85[i,]~as.factor(ordered.sample644$hyb)))

hyb2[i,]<-as.numeric(tmp[1,2:5])

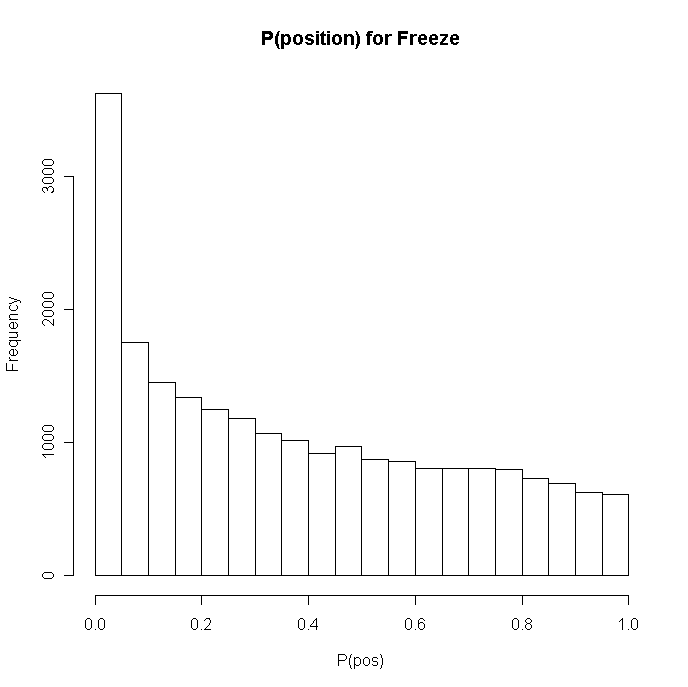
tmp<-anova (lm(medctr.6.85[i,]~as.factor(ordered.sample644$pos)))

pos2[i,]<-as.numeric(tmp[1,2:5])

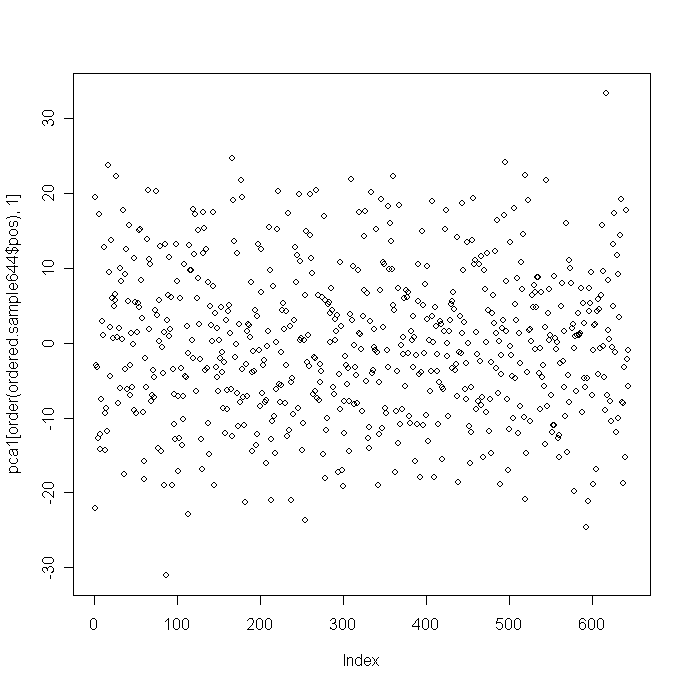
}

date()

hist(pos2)

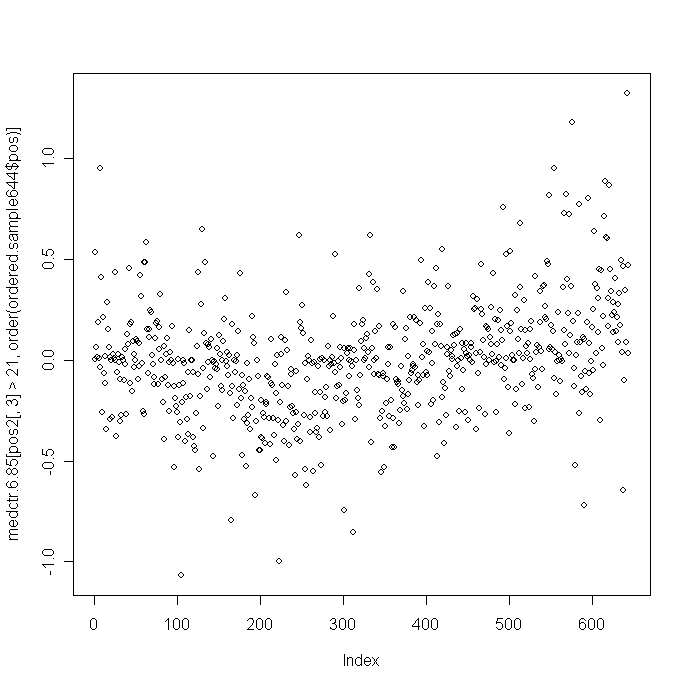


plot(pca2[order(ordered.sample644$pos),1])



anova(lm(pca2[,1]~as.factor(ordered.sample644$pos)))

plot(medctr.6.85[pos2[,3]>21,order(ordered.sample644$pos)])



par(mfrow=c(2,2))

plot(medctr.6.85[rank(pos2[,3])==22177,order(ordered.sample644$pos)],main="gene with the top position effect",pch=19)

lines(runmed(medctr.6.85[rank(pos2[,3])==22177,order(ordered.sample644$pos)],85),col=2,lwd=2)

plot(medctr.6.85[rank(pos2[,3])==22167,order(ordered.sample644$pos)],main="gene with the 10th largest position effect",pch=19)

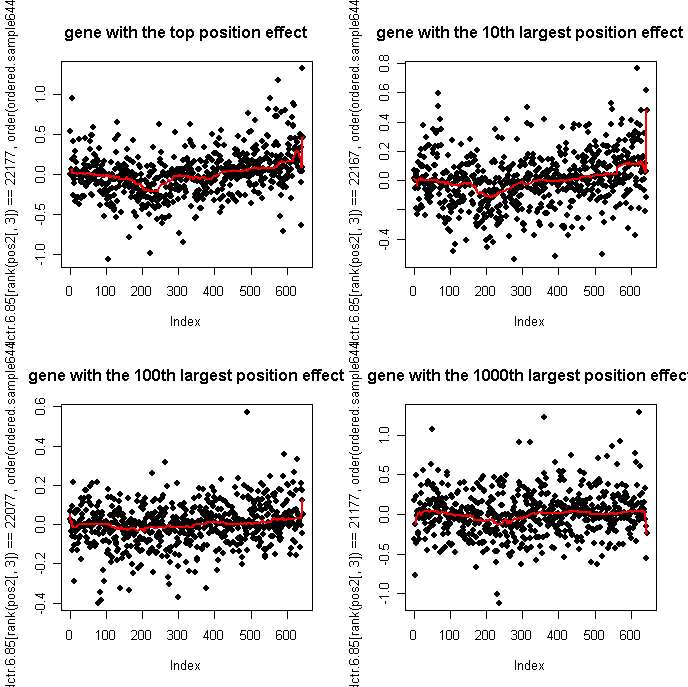
lines(runmed(medctr.6.85[rank(pos2[,3])==22167,order(ordered.sample644$pos)],85),col=2,lwd=2)

plot(medctr.6.85[rank(pos2[,3])==22077,order(ordered.sample644$pos)],main="gene with the 100th largest position effect",pch=19)

lines(runmed(medctr.6.85[rank(pos2[,3])==22077,order(ordered.sample644$pos)],85),col=2,lwd=2)

plot(medctr.6.85[rank(pos2[,3])==21177,order(ordered.sample644$pos)],main="gene with the 1000th largest position effect",pch=19)

lines(runmed(medctr.6.85[rank(pos2[,3])==21177,order(ordered.sample644$pos)],85),col=2,lwd=2)



chip.pos<-chip.b<-pos.b<-matrix(NA,22177,4)

date()

for (i in 1:22177) {

tmp<-anova (lm(medctr.6[i,]~as.factor(ordered.sample644$chip)+ as.factor(ordered.sample644$pos)+ as.factor(ordered.sample644$chip)\* as.factor(ordered.sample644$pos)))

chip.b[i,]<-as.numeric(tmp[1,2:5])

tmp<-anova (lm(medctr.6[i,]~as.factor(ordered.sample644$hyb)))

pos.b[i,]<-as.numeric(tmp[2,2:5])

tmp<-anova (lm(medctr.6[i,]~as.factor(ordered.sample644$pos)))

chip.pos[i,]<-as.numeric(tmp[3,2:5])

}

date()

save.image()

set up per-position median centering

pos.levels<-levels(factor(ordered.sample644$pos))

medctr.6.85.pos<-matrix(NA,22177,644)

for (i in 1:8) {

tmp<-medctr.6.85[,ordered.sample644$pos==pos.levels[i]]

med.chip<-apply(tmp,1,median)

medctr.6.85.pos[,ordered.sample644$pos ==pos.levels[i]]<-medctr.6.85[,ordered.sample644$pos ==pos.levels[i]]-med.chip

}

norm.amy<-medctr.6.85.pos[,1:98]

amy.sample<-ordered.sample644[1:98,]

t3.amy.bp<-ttest.disease(norm.amy,amy.sample,"BP")

t3.amy.md<-ttest.disease(norm.amy,amy.sample,"MD")

t3.amy.sc<-ttest.disease(norm.amy,amy.sample,"SC")

cor(cbind(t2.amy.bp[,1],t2.amy.md[,1],t2.amy.sc[,1],t3.amy.bp[,1],t3.amy.md[,1],t3.amy.sc[,1]))

[,1] [,2] [,3] [,4] [,5] [,6]

[1,] 1.0000000 0.4433583 0.2533433 0.9963874 0.4454304 0.2537575

[2,] 0.4433583 1.0000000 0.3035303 0.4350470 0.9974521 0.2995935

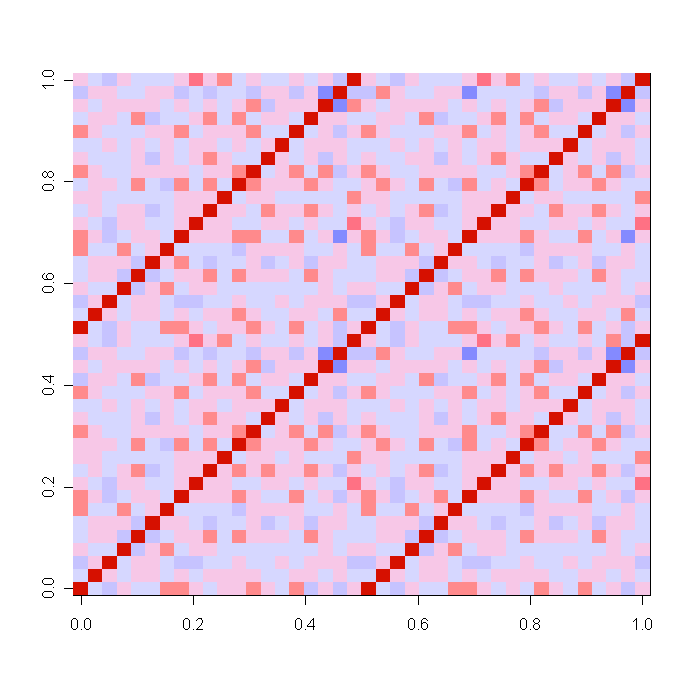
[3,] 0.2533433 0.3035303 1.0000000 0.2455263 0.2979534 0.9953327

[4,] 0.9963874 0.4350470 0.2455263 1.0000000 0.4411381 0.2512495

[5,] 0.4454304 0.9974521 0.2979534 0.4411381 1.0000000 0.2967187

[6,] 0.2537575 0.2995935 0.9953327 0.2512495 0.2967187 1.0000000

image(cor(cbind(medctr.6.85[,1:20],medctr.6.85.pos[,1:20])),col=redblue,zlim=c(-1,1))

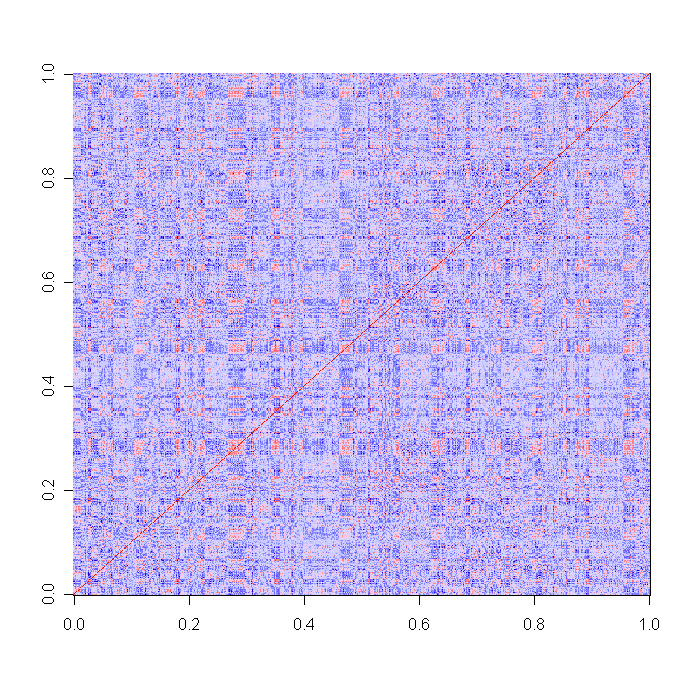


multi-region

tmp <-medctr.6[order(ordered.sample644$Region, ordered.sample644$Individual),]

tmp.cor<-cor(tmp)

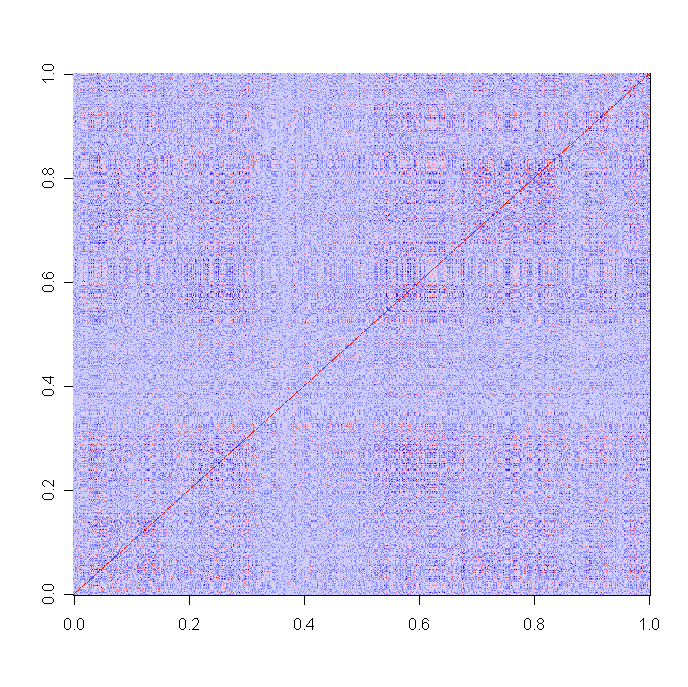
image(tmp.cor,col=redblue)



tmp <-medctr.6.85[order(ordered.sample644$Region, ordered.sample644$Individual),]

tmp.cor<-cor(tmp)

image(tmp.cor,col=redblue)



plot(pca2[,1],pca2[,2] ,pch=19, xlab="PC1-freeze3",ylab="PC2-freeze3")

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

points(pca2[filter,1],pca2[filter,2],pch=19,col=2,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

points(pca2[filter,1],pca2[filter,2],pch=19,col=4,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

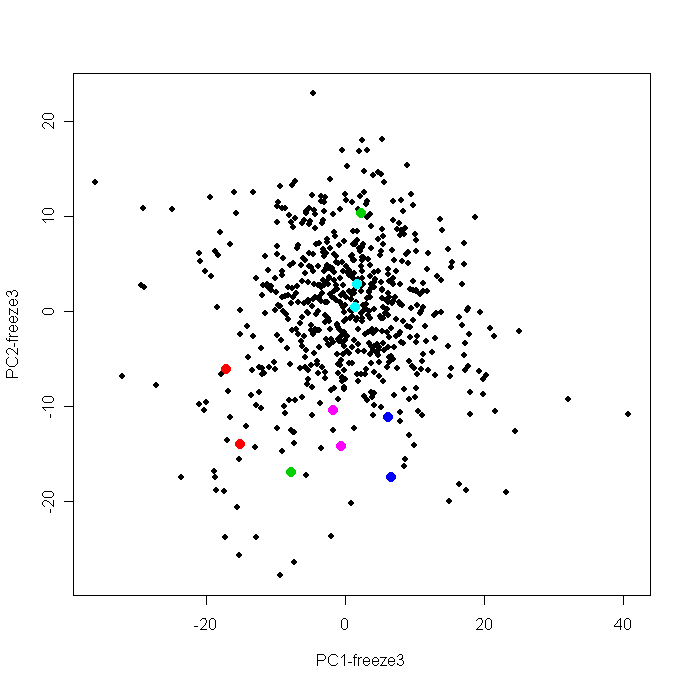
points(pca2[filter,1],pca2[filter,2],pch=19,col=5,cex=1.5)

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

points(pca2[filter,1],pca2[filter,2],pch=19,col=6,cex=1.5)

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

points(pca2[filter,1],pca2[filter,2],pch=19,col=3,cex=1.5)



median(cor(medctr.6.85))

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

cor(medctr.6.85[,filter])

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

cor(medctr.6.85[,filter])

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

cor(medctr.6.85[,filter])

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

cor(medctr.6.85[,filter])

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

cor(medctr.6.85[,filter])

median(cor(medctr.6.85))

[1] 0.000747149

> filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

> cor(medctr.6.85[,filter])

X1692264045\_A X1743113224\_H

X1692264045\_A 1.0000000 0.8234895

X1743113224\_H 0.8234895 1.0000000

> filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

> cor(medctr.6.85[,filter])

X1692264012\_G X1714478082\_H

X1692264012\_G 1.0000000 0.6735202

X1714478082\_H 0.6735202 1.0000000

> filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

> cor(medctr.6.85[,filter])

X1692264012\_E X1743113224\_G

X1692264012\_E 1.0000000 0.4718889

X1743113224\_G 0.4718889 1.0000000

> filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

> cor(medctr.6.85[,filter])

X1692264012\_F X1692264035\_H

X1692264012\_F 1.0000000 0.5571956

X1692264035\_H 0.5571956 1.0000000

> filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

> cor(medctr.6.85[,filter])

X1692264012\_D X1743113008\_H

X1692264012\_D 1.00000000 0.08232898

X1743113008\_H 0.08232898 1.00000000

plot(pca1[,1],pca1[,2],pch=19,xlab="PC1-medctr-region",ylab="PC2-medctr-region")

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

points(pca1[filter,1],pca1[filter,2],pch=19,col=2,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

points(pca1[filter,1],pca1[filter,2],pch=19,col=4,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

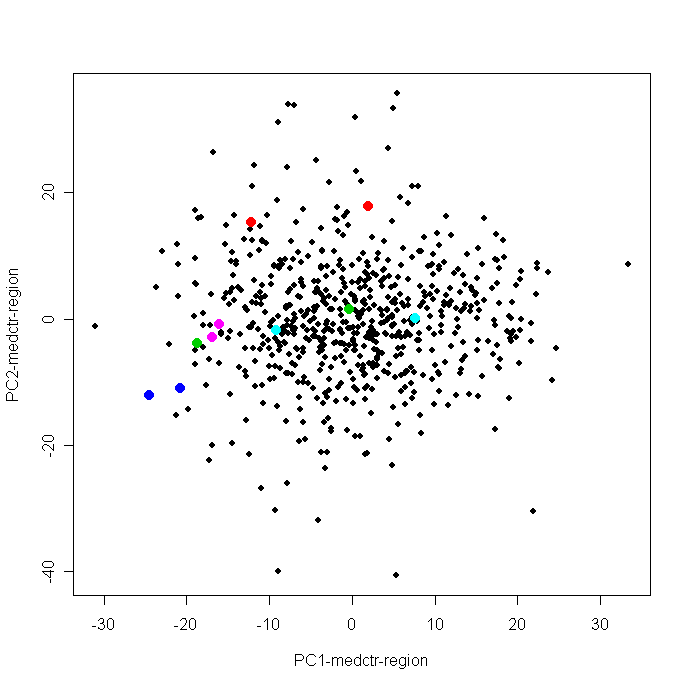
points(pca1[filter,1],pca1[filter,2],pch=19,col=5,cex=1.5)

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

points(pca1[filter,1],pca1[filter,2],pch=19,col=6,cex=1.5)

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

points(pca1[filter,1],pca1[filter,2],pch=19,col=3,cex=1.5)



Add HC comparison with old Illu

Add med.amy, etc back to medctr.6.85 and re-run pca

medctr.6.85.addregion<-matrix(NA,22177,644)

medctr.6.85.addregion[,1:98]<-medctr.6.85[,1:98]+med.amy

medctr.6.85.addregion[,99:207]<-medctr.6.85[,99:207]+med.acg

medctr.6.85.addregion[,208:320]<-medctr.6.85[,208:320]+med.cb

medctr.6.85.addregion[,321:429]<-medctr.6.85[,321:429]+med.dlpfc

medctr.6.85.addregion[,430:538]<-medctr.6.85[,430:538]+med.hc

medctr.6.85.addregion[,539:644]<-medctr.6.85[,539:644]+med.nacc

pca4<-prcomp(t(medctr.6.85.addregion))

pca4<-pca4$x[,1:6]

plot(pca4[,1],pca4[,2],pch=19)

points(pca4[ordered.sample644$Region=="ANCG",1],pca4[ordered.sample644$Region=="ANCG",2],col=2,pch=19)

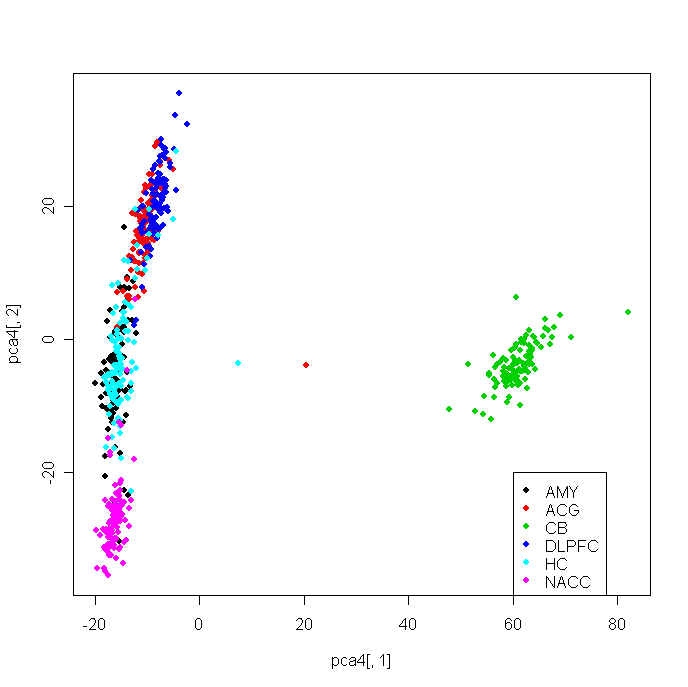
points(pca4[ordered.sample644$Region=="CB",1],pca4[ordered.sample644$Region=="CB",2],col=3,pch=19)

points(pca4[ordered.sample644$Region=="DLPFC",1],pca4[ordered.sample644$Region=="DLPFC",2],col=4,pch=19)

points(pca4[ordered.sample644$Region=="HC",1],pca4[ordered.sample644$Region=="HC",2],col=5,pch=19)

points(pca4[ordered.sample644$Region=="NACC",1],pca4[ordered.sample644$Region=="NACC",2],col=6,pch=19)

legend(60,-20,c("AMY","ACG","CB","DLPFC","HC","NACC"),pch=rep(19,6),col=c(1:6))



plot(pca4[,3],pca4[,2],pch=19)

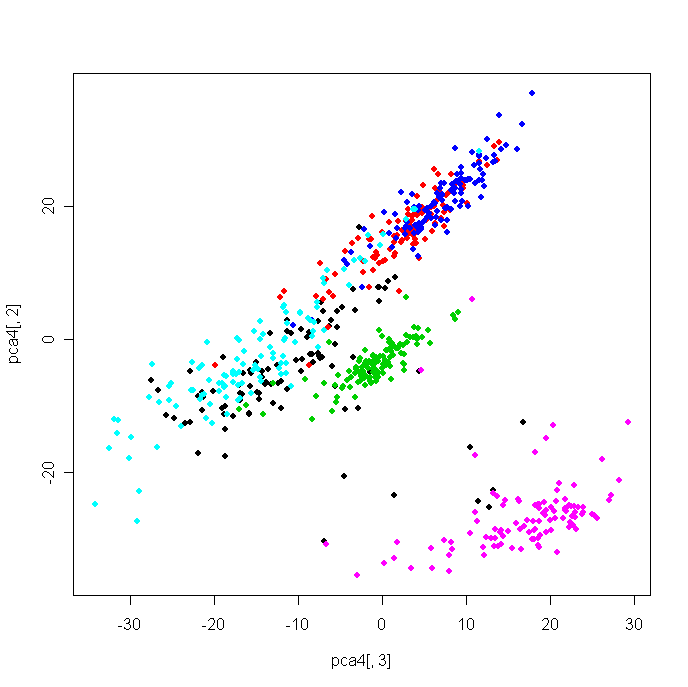
points(pca4[ordered.sample644$Region=="ANCG",3],pca4[ordered.sample644$Region=="ANCG",2],col=2,pch=19)

points(pca4[ordered.sample644$Region=="CB",3],pca4[ordered.sample644$Region=="CB",2],col=3,pch=19)

points(pca4[ordered.sample644$Region=="DLPFC",3],pca4[ordered.sample644$Region=="DLPFC",2],col=4,pch=19)

points(pca4[ordered.sample644$Region=="HC",3],pca4[ordered.sample644$Region=="HC",2],col=5,pch=19)

points(pca4[ordered.sample644$Region=="NACC",3],pca4[ordered.sample644$Region=="NACC",2],col=6,pch=19)



plot(pca4[,1],pca4[,2],pch=19)

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

points(pca4[filter,1],pca4[filter,2],pch=19,col=2,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

points(pca4[filter,1],pca4[filter,2],pch=19,col=4,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

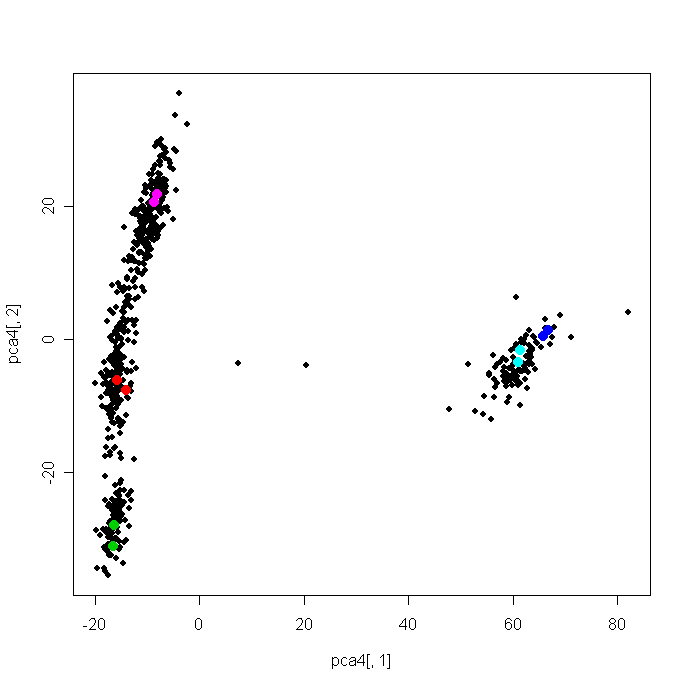
points(pca4[filter,1],pca4[filter,2],pch=19,col=5,cex=1.5)

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

points(pca4[filter,1],pca4[filter,2],pch=19,col=6,cex=1.5)

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

points(pca4[filter,1],pca4[filter,2],pch=19,col=3,cex=1.5)



median(cor(medctr.6.85.addregion))

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

cor(medctr.6.85.addregion[,filter])

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

cor(medctr.6.85.addregion[,filter])

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

cor(medctr.6.85.addregion[,filter])

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

cor(medctr.6.85.addregion[,filter])

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

cor(medctr.6.85.addregion[,filter])

median(cor(medctr.6.85.addregion))

[1] 0.965389

> filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

> cor(medctr.6.85.addregion[,filter])

[,1] [,2]

[1,] 1.0000000 0.9926283

[2,] 0.9926283 1.0000000

> filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

> cor(medctr.6.85.addregion[,filter])

[,1] [,2]

[1,] 1.0000000 0.9882316

[2,] 0.9882316 1.0000000

> filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

> cor(medctr.6.85.addregion[,filter])

[,1] [,2]

[1,] 1.000000 0.994699

[2,] 0.994699 1.000000

> filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

> cor(medctr.6.85.addregion[,filter])

[,1] [,2]

[1,] 1.0000000 0.9923794

[2,] 0.9923794 1.0000000

> filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

> cor(medctr.6.85.addregion[,filter])

[,1] [,2]

[1,] 1.0000000 0.9651652

[2,] 0.9651652 1.0000000

rm(medctr.6.85.addregion)

Try Irvine approach

I1. median center every sample

irvine1<-matrix(NA,22177,644)

for (i in 1:644){

irvine1[,i]<-total[,i]-q10[i,5]

}

pca.irvine1<-prcomp(t(irvine1))

pca.irvine1<-pca.irvine1$x[,1:6]

plot(pca.irvine1[,1],pca.irvine1[,2],pch=19,xlab="PC1-Irvine\_medctr",ylab="PC2-Irvine\_medctr ")

points(pca.irvine1[ordered.sample644$Region=="ANCG",1],pca.irvine1[ordered.sample644$Region=="ANCG",2],col=2,pch=19)

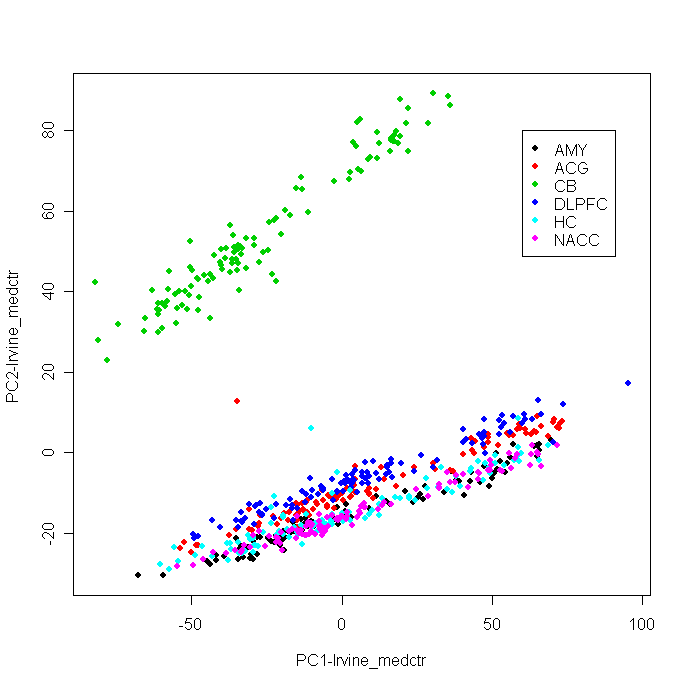
points(pca.irvine1[ordered.sample644$Region=="CB",1],pca.irvine1[ordered.sample644$Region=="CB",2],col=3,pch=19)

points(pca.irvine1[ordered.sample644$Region=="DLPFC",1],pca.irvine1[ordered.sample644$Region=="DLPFC",2],col=4,pch=19)

points(pca.irvine1[ordered.sample644$Region=="HC",1],pca.irvine1[ordered.sample644$Region=="HC",2],col=5,pch=19)

points(pca.irvine1[ordered.sample644$Region=="NACC",1],pca.irvine1[ordered.sample644$Region=="NACC",2],col=6,pch=19)

legend(60,80,c("AMY","ACG","CB","DLPFC","HC","NACC"),pch=rep(19,6),col=c(1:6))



plot(rank(ordered.sample644$Chip), pca.irvine1[,1])

plot(pca.irvine1[,1],pca.irvine1[,2],pch=19)

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

points(pca.irvine1[filter,1],pca.irvine1[filter,2],pch=19,col=2,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

points(pca.irvine1[filter,1],pca.irvine1[filter,2],pch=19,col=4,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

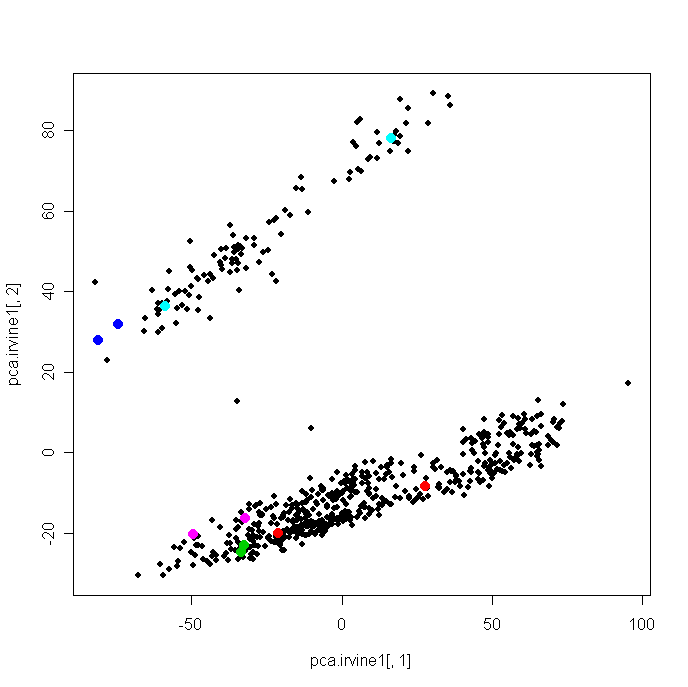
points(pca.irvine1[filter,1],pca.irvine1[filter,2],pch=19,col=5,cex=1.5)

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

points(pca.irvine1[filter,1],pca.irvine1[filter,2],pch=19,col=6,cex=1.5)

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

points(pca.irvine1[filter,1],pca.irvine1[filter,2],pch=19,col=3,cex=1.5)



plot(ordered.sample644$chip.order,pca.irvine1[,1], main="pca.irvine1\_1, by chip.order",pch=19,cex=0.7)

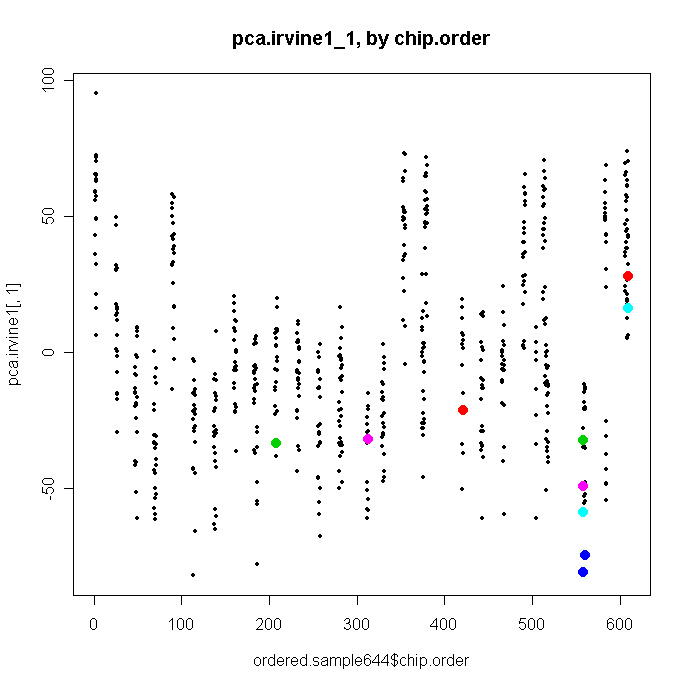
points(ordered.sample644$chip.order[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")],pca.irvine1[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445"),1],pch=19,cex=1.5,col=2)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")],pca.irvine1[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667"),1],pch=19,cex=1.5,col=4)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")],pca.irvine1[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196"),1],pch=19,cex=1.5,col=5)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")],pca.irvine1[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273"),1],pch=19,cex=1.5,col=6)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")],pca.irvine1[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228"),1],pch=19,cex=1.5,col=3)



I2. take residue of chip

irvine2<- matrix(NA,22177,644)

for (i in 1:22177) {

irvine2[i,]<-lm(irvine1[i,]~as.factor(ordered.sample644$chip))$resid

}

irvine2.chip<-matrix(NA,22177,4)

irvine2.region<-matrix(NA,22177,4)

date()

for (i in 222:22177) {

tmp<-anova (lm(irvine2[i,]~as.factor(ordered.sample644$chip.order)+ as.factor(ordered.sample644$Region)))

irvine2.chip[i,]<-as.numeric(tmp[1,2:5])

irvine2.region[i,]<-as.numeric(tmp[2,2:5])

}

date()

chip5<-matrix(NA,22177,4)

region5<-matrix(NA,22177,4)

date()

for (i in 1:22177) {

tmp<-anova (lm(medctr.6.85.addregion[i,]~as.factor(ordered.sample644$chip.order)+ as.factor(ordered.sample644$Region)))

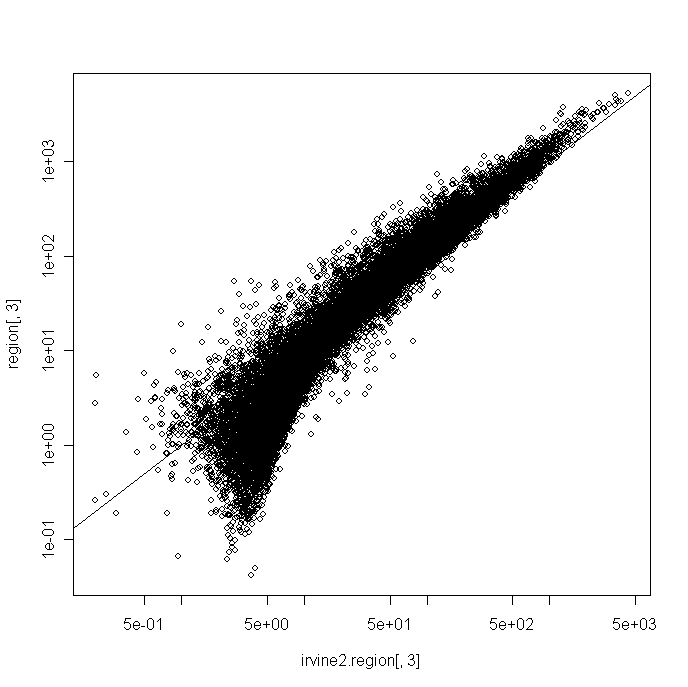
chip5[i,]<-as.numeric(tmp[1,2:5])

region5[i,]<-as.numeric(tmp[2,2:5])

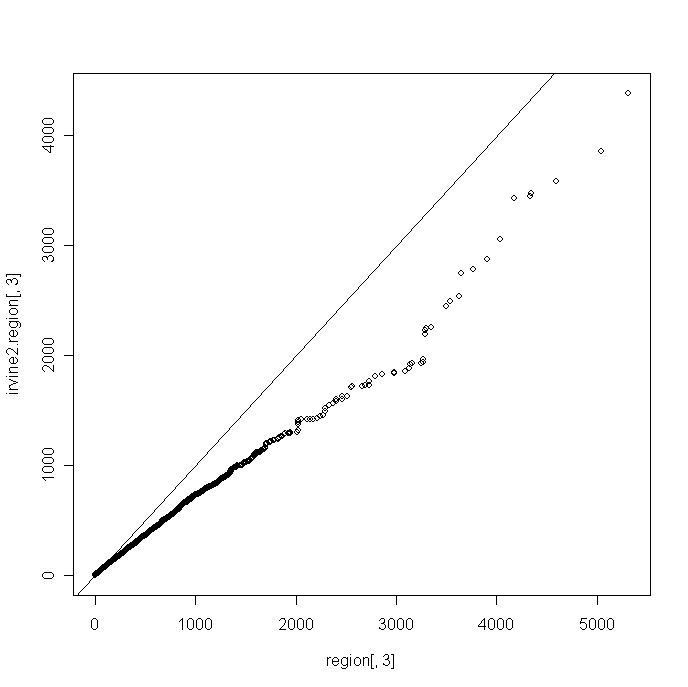
}

date()

plot(irvine2.region[,3],region[,3],log="xy")



qqplot(region[,3],irvine2.region[,3])



plot(ordered.sample644$chip.order,pca.irvine2[,1], main="pca.irvine2\_1, by chip.order",pch=19,cex=0.7)

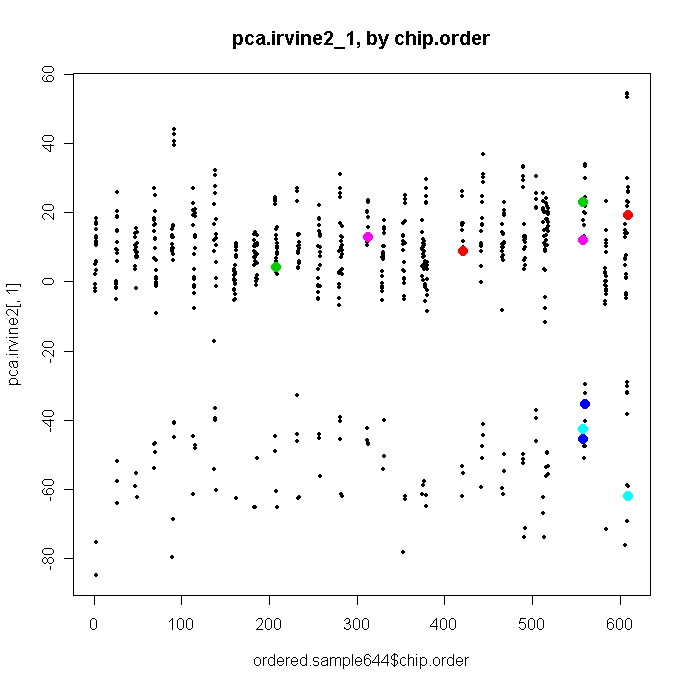
points(ordered.sample644$chip.order[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")],pca.irvine2[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445"),1],pch=19,cex=1.5,col=2)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")],pca.irvine2[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667"),1],pch=19,cex=1.5,col=4)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")],pca.irvine2[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196"),1],pch=19,cex=1.5,col=5)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")],pca.irvine2[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273"),1],pch=19,cex=1.5,col=6)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")],pca.irvine2[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228"),1],pch=19,cex=1.5,col=3)



plot(ordered.sample644$chip.order,pca.irvine2[,2], main="pca.irvine2\_1, by chip.order",pch=19,cex=0.7)

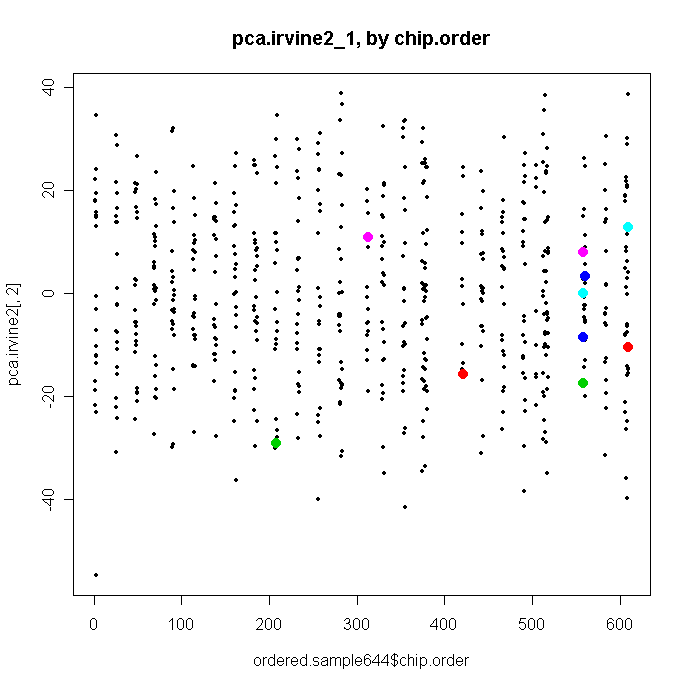
points(ordered.sample644$chip.order[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")],pca.irvine2[(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445"),2],pch=19,cex=1.5,col=2)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")],pca.irvine2[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667"),2],pch=19,cex=1.5,col=4)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")],pca.irvine2[(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196"),2],pch=19,cex=1.5,col=5)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")],pca.irvine2[(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273"),2],pch=19,cex=1.5,col=6)

points(ordered.sample644$chip.order[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")],pca.irvine2[(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228"),2],pch=19,cex=1.5,col=3)



I3. medctr by chip

irvine3<-matrix(NA,22177,644)

for (i in 1:85) {

tmp<-irvine1[,ordered.sample644$chip.order==chip.levels[i]]

med.chip<-apply(tmp,1,median)

irvine3[,ordered.sample644$chip.order==chip.levels[i]]<-irvine1[,ordered.sample644$chip.order==chip.levels[i]]-med.chip

}

colnames(irvine3)<-colnames(total)

rownames(irvine3)<-rownames(total)

colnames(irvine2)<-colnames(total)

rownames(irvine2)<-rownames(total)

write.table(irvine2,"Irvine\_resid\_chip.txt",sep="\t")

write.table(irvine3,"Irvine\_medctr\_chip.txt",sep="\t")

I4. take residue of chip+hyb

irvine4<- matrix(NA,22177,644)

for (i in 1:22177) {

irvine4[i,]<-lm(irvine1[i,]~as.factor(ordered.sample644$chip)+ as.factor(ordered.sample644$hyb))$resid

}

sum(irvine4==irvine2)/644

[1] 22177

rm(irvine4)

I5. meanctr by chip

irvine5<-matrix(NA,22177,644)

for (i in 1:85) {

tmp<-irvine1[,ordered.sample644$chip.order==chip.levels[i]]

med.chip<-apply(tmp,1,mean)

irvine5[,ordered.sample644$chip.order==chip.levels[i]]<-irvine1[,ordered.sample644$chip.order==chip.levels[i]]-med.chip

}

check rep chips

pca.irvine2<-prcomp(t(irvine2))

pca.irvine2<-pca.irvine2$x[,1:6]

plot(pca.irvine2[,1],pca.irvine2[,2],pch=19)

points(pca.irvine2[ordered.sample644$Region=="ANCG",1],pca.irvine2[ordered.sample644$Region=="ANCG",2],col=2,pch=19)

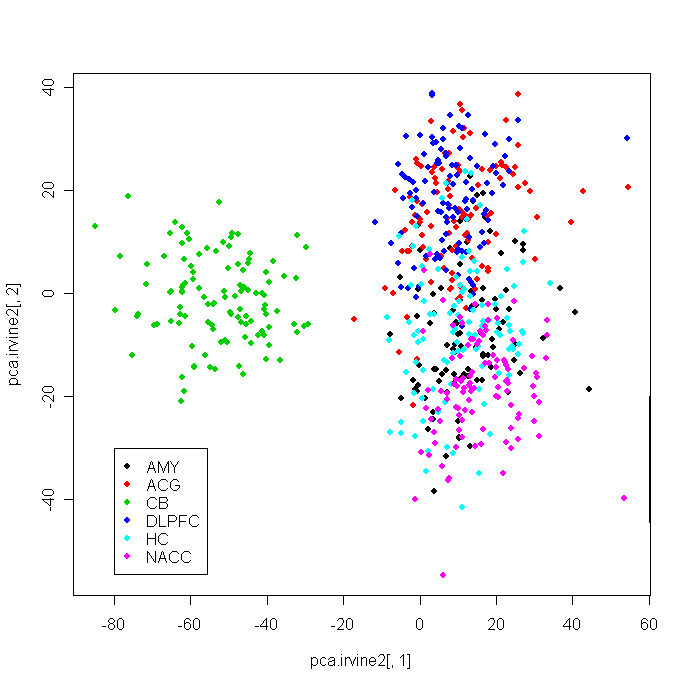
points(pca.irvine2[ordered.sample644$Region=="CB",1],pca.irvine2[ordered.sample644$Region=="CB",2],col=3,pch=19)

points(pca.irvine2[ordered.sample644$Region=="DLPFC",1],pca.irvine2[ordered.sample644$Region=="DLPFC",2],col=4,pch=19)

points(pca.irvine2[ordered.sample644$Region=="HC",1],pca.irvine2[ordered.sample644$Region=="HC",2],col=5,pch=19)

points(pca.irvine2[ordered.sample644$Region=="NACC",1],pca.irvine2[ordered.sample644$Region=="NACC",2],col=6,pch=19)

legend(-80,-30,c("AMY","ACG","CB","DLPFC","HC","NACC"),pch=rep(19,6),col=c(1:6))



plot(pca.irvine2[,1],pca.irvine2[,2] ,pch=19)

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

points(pca.irvine2[filter,1],pca.irvine2[filter,2],pch=19,col=2,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

points(pca.irvine2[filter,1],pca.irvine2[filter,2],pch=19,col=4,cex=1.5)

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

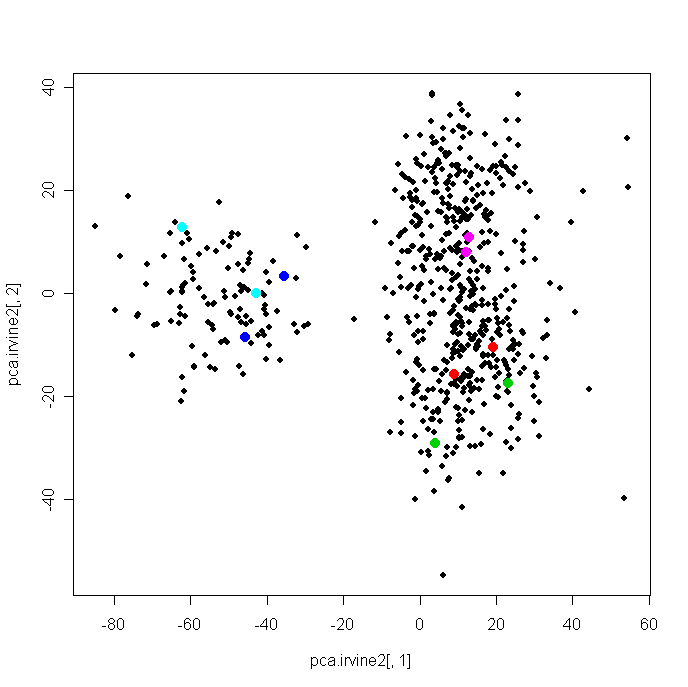
points(pca.irvine2[filter,1],pca.irvine2[filter,2],pch=19,col=5,cex=1.5)

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

points(pca.irvine2[filter,1],pca.irvine2[filter,2],pch=19,col=6,cex=1.5)

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

points(pca.irvine2[filter,1],pca.irvine2[filter,2],pch=19,col=3,cex=1.5)



q10.i2<-matrix(NA,644,10)

q10.i2[,1]<-apply(irvine2,2,min)

for (i in 2:9) {

q10.i2[,i]<-apply(irvine2,2,function(x) quantile(x,0.1\*i))

}

q10.i2[,10]<-apply(irvine2,2,max)

plot(q10.i2[,1],ylim=c(-6,1),type="l",main="By region")

lines(q10.i2[,3],col=2)

lines(q10.i2[,5],col=3)

lines(q10.i2[,7],col=4)

lines(q10.i2[,9],col=5)

median(cor(irvine2))

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

cor(irvine2[,filter])

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

cor(irvine2[,filter])

filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

cor(irvine2[,filter])

filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

cor(irvine2[,filter])

filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

cor(irvine2[,filter])

> median(cor(irvine2))

[1] -0.01816336

> filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

> cor(irvine2[,filter])

[,1] [,2]

[1,] 1.0000000 0.8715566

[2,] 0.8715566 1.0000000

> filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3667")

> cor(irvine2[,filter])

[,1] [,2]

[1,] 1.000000 0.861041

[2,] 0.861041 1.000000

> filter<-(ordered.sample644$Region=="CB")&(ordered.sample644$Individual=="3196")

> cor(irvine2[,filter])

[,1] [,2]

[1,] 1.000000 0.890993

[2,] 0.890993 1.000000

> filter<-(ordered.sample644$Region=="DLPFC")&(ordered.sample644$Individual=="3273")

> cor(irvine2[,filter])

[,1] [,2]

[1,] 1.0000000 0.8197524

[2,] 0.8197524 1.0000000

> filter<-(ordered.sample644$Region=="NACC")&(ordered.sample644$Individual=="3228")

> cor(irvine2[,filter])

[,1] [,2]

[1,] 1.0000000 0.2607082

[2,] 0.2607082 1.0000000

compare t scores and with pca4

irvine2.amy<-irvine2[,1:98]

t.i2.amy.bp<-ttest.disease(irvine2.amy,amy.sample,"BP")

t.i2.amy.md<-ttest.disease(irvine2.amy,amy.sample,"MD")

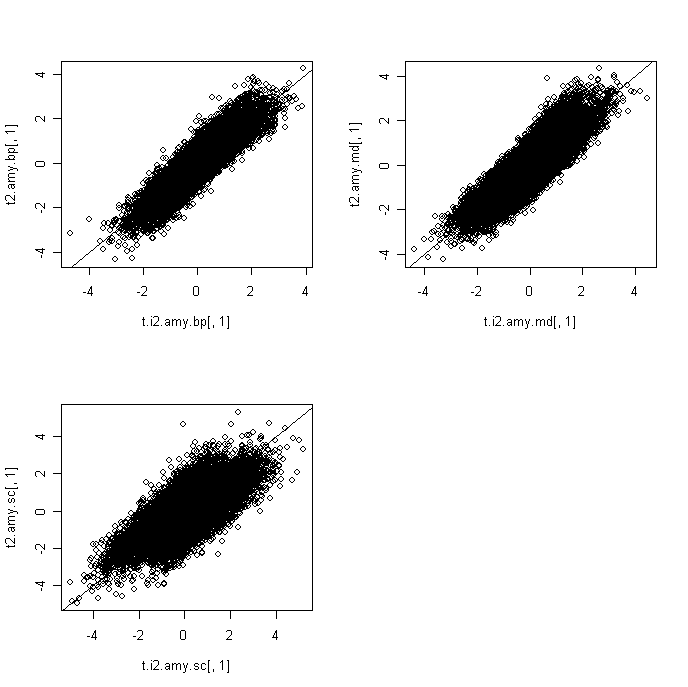
t.i2.amy.sc<-ttest.disease(irvine2.amy,amy.sample,"SC")

irvine3.amy<-irvine3[,1:98]

t.i3.amy.bp<-ttest.disease(irvine3.amy,amy.sample,"BP")

t.i3.amy.md<-ttest.disease(irvine3.amy,amy.sample,"MD")

t.i3.amy.sc<-ttest.disease(irvine3.amy,amy.sample,"SC")



compare 5 replicate samples with their regions across total, norm.total,medctr.6.85, and irvine2

tmp<-total[,ordered.sample644$Region=="AMY"]

n<-dim(tmp)[2]

tmp.dist<-matrix(0,n,n)

for (i in 1:(n-1)) {

for (j in (i+1):n) {

tmp.dist[i,j]<-median(abs(tmp[,i]-tmp[,j]))

tmp.dist[j,i]<-tmp.dist[i,j]

}

}

median(tmp.cor)

[1] 0.2561109

filter<-(ordered.sample644$Region=="AMY")&(ordered.sample644$Individual=="3445")

tmp<-total[,filter]

median(abs(tmp[,1]-tmp[,2]))

0.3390926

get.dist<-function(data,region,sample){

tmp<-data[,ordered.sample644$Region==region]

n<-dim(tmp)[2]

tmp.dist<-matrix(0,n,n)

for (i in 1:(n-1)) {

for (j in (i+1):n) {

tmp.dist[i,j]<-median(abs(tmp[,i]-tmp[,j]))

tmp.dist[j,i]<-tmp.dist[i,j]

}

}

dist1<- median(tmp.dist)

filter<-(ordered.sample644$Region==region)&(ordered.sample644$Individual==sample)

tmp<-data[,filter]

dist2<-median(abs(tmp[,1]-tmp[,2]))

return(c(dist1,dist2))

}

get.dist(total,"AMY","3445")

get.dist(total,"CB","3667")

get.dist(total,"CB","3196")

get.dist(total,"DLPFC","3273")

get.dist(total,"NACC","3228")

get.dist(norm.total,"AMY","3445")

get.dist(norm.total,"CB","3667")

get.dist(norm.total,"CB","3196")

get.dist(norm.total,"DLPFC","3273")

get.dist(norm.total,"NACC","3228")

get.dist(medctr.6.85,"AMY","3445")

get.dist(medctr.6.85,"CB","3667")

get.dist(medctr.6.85,"CB","3196")

get.dist(medctr.6.85,"DLPFC","3273")

get.dist(medctr.6.85,"NACC","3228")

get.dist(irvine2,"AMY","3445")

get.dist(irvine2,"CB","3667")

get.dist(irvine2,"CB","3196")

get.dist(irvine2,"DLPFC","3273")

get.dist(irvine2,"NACC","3228")

get.dist(total,"AMY","3445")

[1] 0.2561109 0.3390926

> get.dist(total,"CB","3667")

[1] 0.2458246 0.3461287

> get.dist(total,"CB","3196")

[1] 0.2458246 0.6271463

> get.dist(total,"DLPFC","3273")

[1] 0.2799377 0.2994483

> get.dist(total,"NACC","3228")

[1] 0.2596760 0.1964952

>

> get.dist(norm.total,"AMY","3445")

[1] 0.1552657 0.1079565

> get.dist(norm.total,"CB","3667")

[1] 0.1400469 0.1302381

> get.dist(norm.total,"CB","3196")

[1] 0.1400469 0.1260309

> get.dist(norm.total,"DLPFC","3273")

[1] 0.1385340 0.1262399

> get.dist(norm.total,"NACC","3228")

[1] 0.1407873 0.1577954

>

> get.dist(medctr.6.85,"AMY","3445")

[1] 0.12470835 0.09981434

> get.dist(medctr.6.85,"CB","3667")

[1] 0.1116107 0.1262418

> get.dist(medctr.6.85,"CB","3196")

[1] 0.11161069 0.08655029

> get.dist(medctr.6.85,"DLPFC","3273")

[1] 0.1115445 0.1061594

> get.dist(medctr.6.85,"NACC","3228")

[1] 0.1135113 0.1657555

>

> get.dist(irvine2,"AMY","3445")

[1] 0.1424723 0.1091387

> get.dist(irvine2,"CB","3667")

[1] 0.13193799 0.09824162

> get.dist(irvine2,"CB","3196")

[1] 0.1319380 0.1174413

> get.dist(irvine2,"DLPFC","3273")

[1] 0.1309272 0.0796930

> get.dist(irvine2,"NACC","3228")

[1] 0.1348751 0.1687627

>

get.dist(medctr.6,"AMY","3445")

get.dist(medctr.6,"CB","3667")

get.dist(medctr.6,"CB","3196")

get.dist(medctr.6,"DLPFC","3273")

get.dist(medctr.6,"NACC","3228")

same as medctr.6.85

get.dist(irvine1,"AMY","3445")

get.dist(irvine1,"CB","3667")

get.dist(irvine1,"CB","3196")

get.dist(irvine1,"DLPFC","3273")

get.dist(irvine1,"NACC","3228")

get.dist(irvine1,"AMY","3445")

[1] 0.1987064 0.2654112

> get.dist(irvine1,"CB","3667")

[1] 0.1795306 0.1043743

> get.dist(irvine1,"CB","3196")

[1] 0.1795306 0.4474110

> get.dist(irvine1,"DLPFC","3273")

[1] 0.1963169 0.1198038

> get.dist(irvine1,"NACC","3228")

[1] 0.1931045 0.1216504

Between irvine1 and irvine2 check the correlation between CB fraction and amount of correction across chips

for (i in 4:85) {

tmp<-medctr.6[,ordered.sample644$chip.order==chip.levels[i]]

med.chip<-apply(tmp,1,median)

medctr.6.85[,ordered.sample644$chip.order==chip.levels[i]]<-medctr.6[,ordered.sample644$chip.order==chip.levels[i]]-med.chip

}

CB.count<-rep(NA,85)

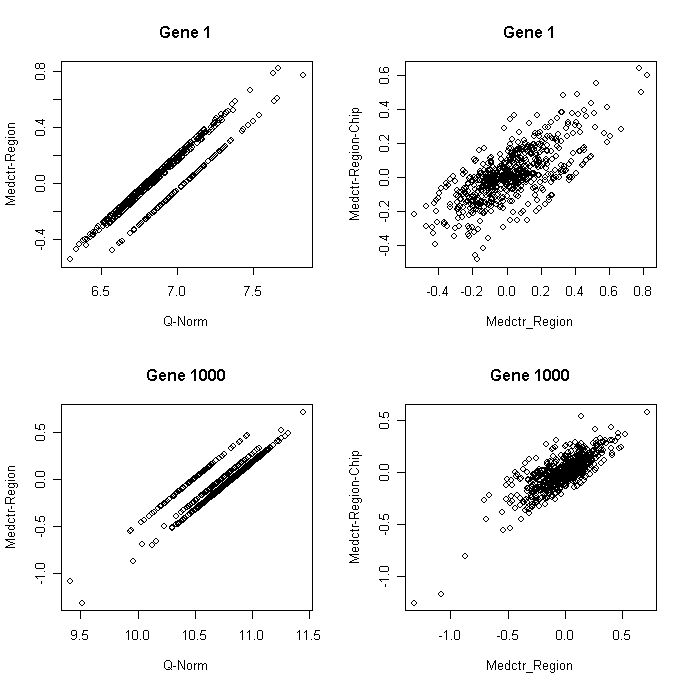
Make some plots:

plot(norm.total[1,],medctr.6[1,],main="Gene 1",xlab="Q-Norm",ylab="Medctr-Region")

plot(medctr.6[1,],medctr.6.85[1,],main="Gene 1",xlab="Medctr\_Region",ylab="Medctr-Region-Chip")

plot(norm.total[1000,],medctr.6[1000,],main="Gene 1000",xlab="Q-Norm",ylab="Medctr-Region")

plot(medctr.6[1000,],medctr.6.85[1000,],main="Gene 1000",xlab="Medctr\_Region",ylab="Medctr-Region-Chip")



test pca1's case-control impact

tmp<-t(pca1[1:98,])

t.pca1.amy.bp<-ttest.disease(tmp,amy.sample,"BP")

t.pca1.amy.md<-ttest.disease(tmp,amy.sample,"MD")

t.pca1.amy.sc<-ttest.disease(tmp,amy.sample,"SC")

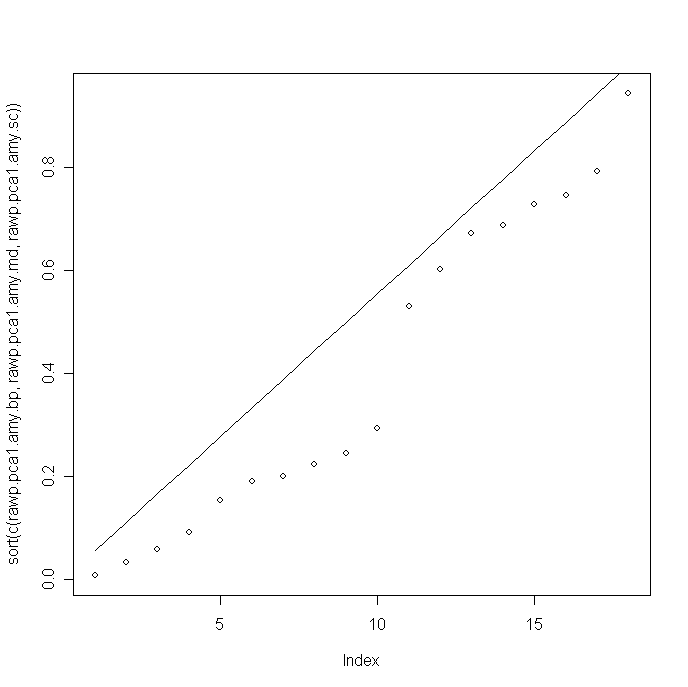
rawp.pca1.amy.bp<-2\*(1-pt(abs(t.pca1.amy.bp[,1]),53))

rawp.pca1.amy.md<-2\*(1-pt(abs(t.pca1.amy.md[,1]),74))

rawp.pca1.amy.sc<-2\*(1-pt(abs(t.pca1.amy.sc[,1]),51))

plot(sort(c( rawp.pca1.amy.bp,rawp.pca1.amy.md,rawp.pca1.amy.sc)))

lines(c(1:18)/18)



tmp<-t(pca1[99:207,])

t.pca1.acg.bp<-ttest.disease(tmp,acg.sample,"BP")

t.pca1.acg.md<-ttest.disease(tmp,acg.sample,"MD")

t.pca1.acg.sc<-ttest.disease(tmp,acg.sample,"SC")

rawp.pca1.acg.bp<-2\*(1-pt(abs(t.pca1.acg.bp[,1]),59))

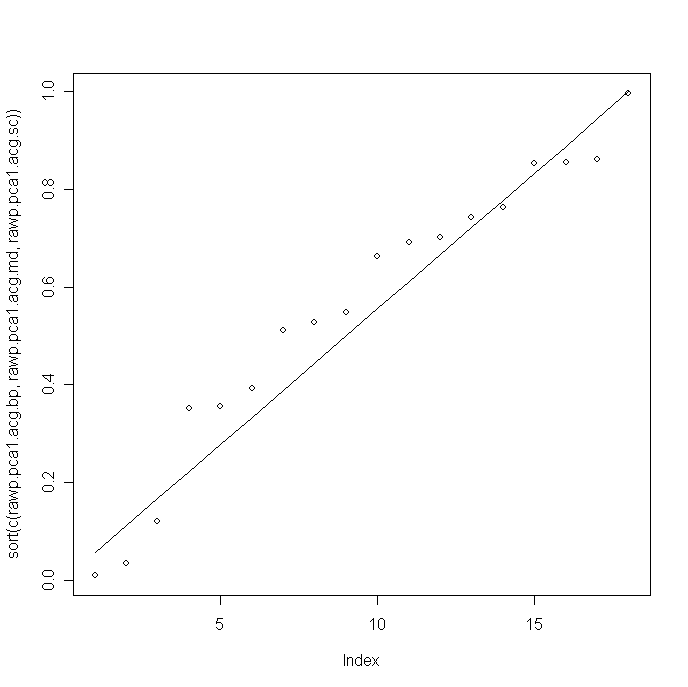
rawp.pca1.acg.md<-2\*(1-pt(abs(t.pca1.acg.md[,1]),79))

rawp.pca1.acg.sc<-2\*(1-pt(abs(t.pca1.acg.sc[,1]),59))

plot(c( rawp.pca1.acg.bp,rawp.pca1.acg.md,rawp.pca1.acg.sc),log="y")

plot(sort(c( rawp.pca1.acg.bp,rawp.pca1.acg.md,rawp.pca1.acg.sc)))

lines(c(1:18)/18)



tmp<-t(pca1[208:320,])

t.pca1.cb.bp<-ttest.disease(tmp,cb.sample,"BP")

t.pca1.cb.md<-ttest.disease(tmp,cb.sample,"MD")

t.pca1.cb.sc<-ttest.disease(tmp,cb.sample,"SC")

rawp.pca1.cb.bp<-2\*(1-pt(abs(t.pca1.cb.bp[,1]),59))

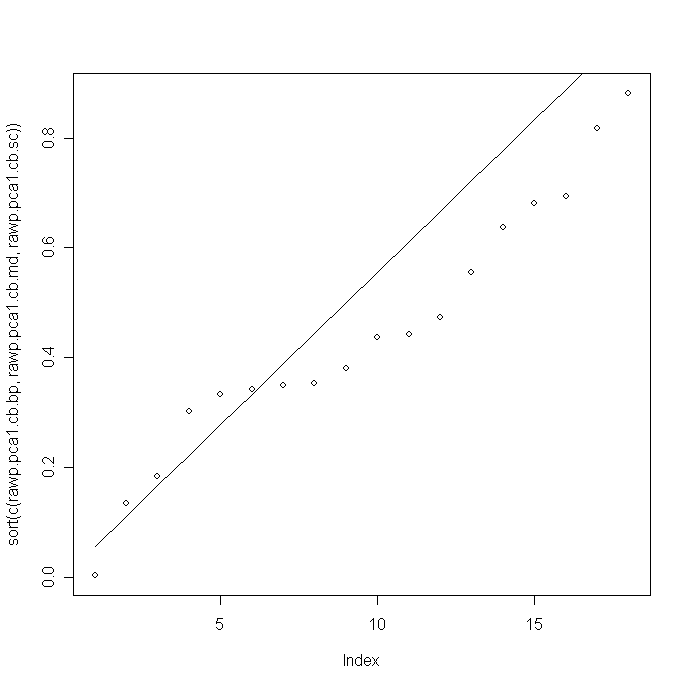
rawp.pca1.cb.md<-2\*(1-pt(abs(t.pca1.cb.md[,1]),79))

rawp.pca1.cb.sc<-2\*(1-pt(abs(t.pca1.cb.sc[,1]),59))

plot(c( rawp.pca1.cb.bp,rawp.pca1.cb.md,rawp.pca1.cb.sc),log="y")

plot(sort(c( rawp.pca1.cb.bp,rawp.pca1.cb.md,rawp.pca1.cb.sc)))

lines(c(1:18)/18)



tmp<-t(pca1[321:429,])

t.pca1.dlpfc.bp<-ttest.disease(tmp,dlpfc.sample,"BP")

t.pca1.dlpfc.md<-ttest.disease(tmp,dlpfc.sample,"MD")

t.pca1.dlpfc.sc<-ttest.disease(tmp,dlpfc.sample,"SC")

rawp.pca1.dlpfc.bp<-2\*(1-pt(abs(t.pca1.dlpfc.bp[,1]),59))

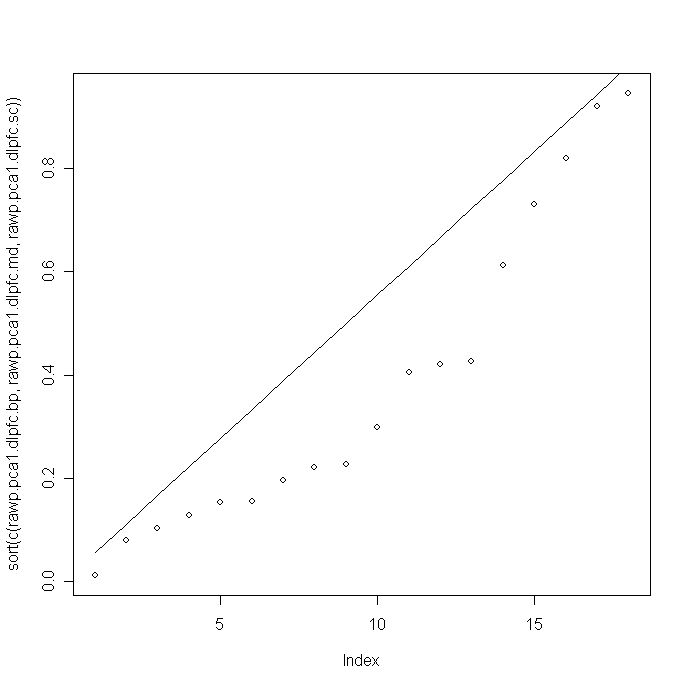
rawp.pca1.dlpfc.md<-2\*(1-pt(abs(t.pca1.dlpfc.md[,1]),79))

rawp.pca1.dlpfc.sc<-2\*(1-pt(abs(t.pca1.dlpfc.sc[,1]),59))

plot(c( rawp.pca1.dlpfc.bp,rawp.pca1.dlpfc.md,rawp.pca1.dlpfc.sc),log="y")

plot(sort(c( rawp.pca1.dlpfc.bp,rawp.pca1.dlpfc.md,rawp.pca1.dlpfc.sc)))

lines(c(1:18)/18)



tmp<-t(pca1[430:538,])

t.pca1.hc.bp<-ttest.disease(tmp,hc.sample,"BP")

t.pca1.hc.md<-ttest.disease(tmp,hc.sample,"MD")

t.pca1.hc.sc<-ttest.disease(tmp,hc.sample,"SC")

rawp.pca1.hc.bp<-2\*(1-pt(abs(t.pca1.hc.bp[,1]),59))

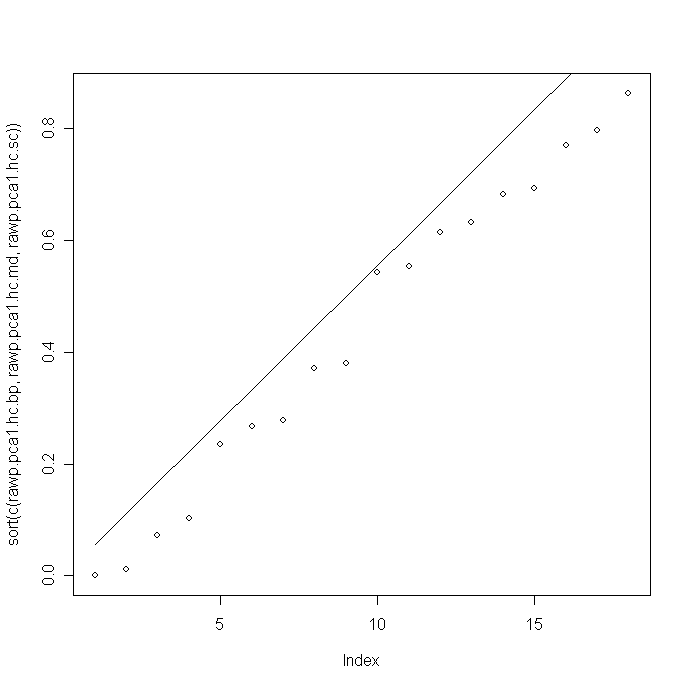
rawp.pca1.hc.md<-2\*(1-pt(abs(t.pca1.hc.md[,1]),79))

rawp.pca1.hc.sc<-2\*(1-pt(abs(t.pca1.hc.sc[,1]),59))

plot(c( rawp.pca1.hc.bp,rawp.pca1.hc.md,rawp.pca1.hc.sc),log="y")

plot(sort(c( rawp.pca1.hc.bp,rawp.pca1.hc.md,rawp.pca1.hc.sc)))

lines(c(1:18)/18)



tmp<-t(pca1[539:644,])

t.pca1.nacc.bp<-ttest.disease(tmp,nacc.sample,"BP")

t.pca1.nacc.md<-ttest.disease(tmp,nacc.sample,"MD")

t.pca1.nacc.sc<-ttest.disease(tmp,nacc.sample,"SC")

rawp.pca1.nacc.bp<-2\*(1-pt(abs(t.pca1.nacc.bp[,1]),59))

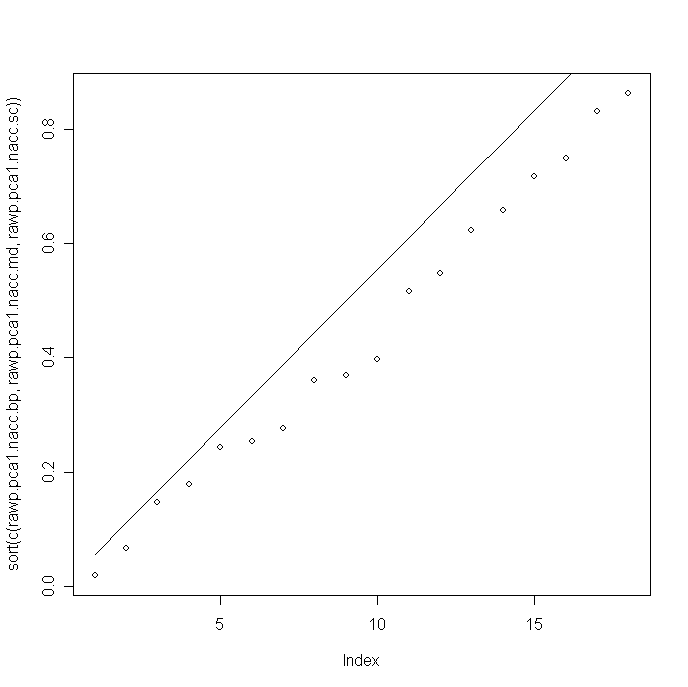
rawp.pca1.nacc.md<-2\*(1-pt(abs(t.pca1.nacc.md[,1]),79))

rawp.pca1.nacc.sc<-2\*(1-pt(abs(t.pca1.nacc.sc[,1]),59))

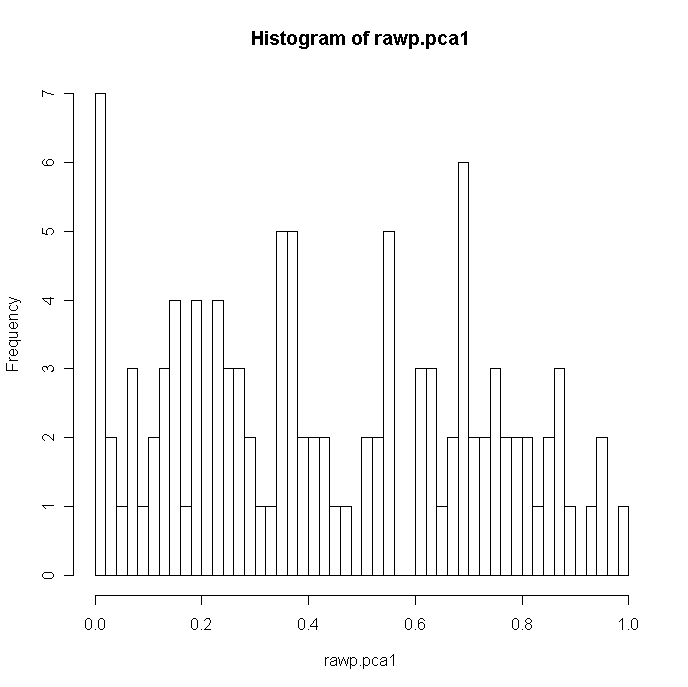
plot(c( rawp.pca1.nacc.bp,rawp.pca1.nacc.md,rawp.pca1.nacc.sc),log="y")

plot(sort(c( rawp.pca1.nacc.bp,rawp.pca1.nacc.md,rawp.pca1.nacc.sc)))

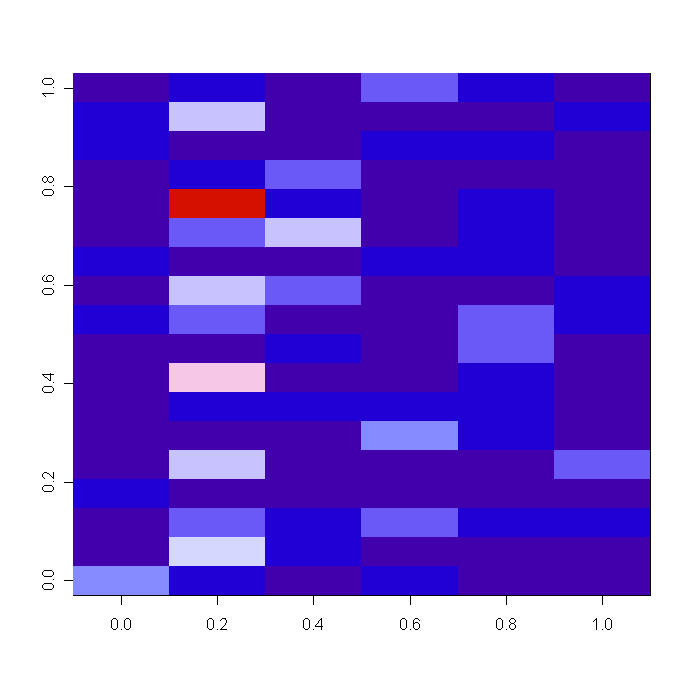
lines(c(1:18)/18)



rawp.pca1<-cbind(rawp.pca1.amy.bp, rawp.pca1.amy.md, rawp.pca1.amy.sc, rawp.pca1.acg.bp, rawp.pca1.acg.md, rawp.pca1.acg.sc, rawp.pca1.cb.bp, rawp.pca1.cb.md, rawp.pca1.cb.sc, rawp.pca1.dlpfc.bp, rawp.pca1.dlpfc.md, rawp.pca1.dlpfc.sc, rawp.pca1.hc.bp, rawp.pca1.hc.md, rawp.pca1.hc.sc, rawp.pca1.nacc.bp, rawp.pca1.nacc.md, rawp.pca1.nacc.sc)



image((-1)\*log(rawp.pca1),col=redblue)



New analysis (6/10/2010) in repose to David Lyon's request.

Need to ru-run the analysis as the old one was on Stanford server

norm.hc<-medctr.6[,430:538]

hc.sample<-ordered.sample644[430:538,]

t.hc.bp<-ttest.disease(norm.hc,hc.sample,"BP")

t.hc.md<-ttest.disease(norm.hc,hc.sample,"MD")

t.hc.sc<-ttest.disease(norm.hc,hc.sample,"SC")

norm.hc<-medctr.6.85[,430:538]

t2.hc.bp<-ttest.disease(norm.hc,hc.sample,"BP")

t2.hc.md<-ttest.disease(norm.hc,hc.sample,"MD")

t2.hc.sc<-ttest.disease(norm.hc,hc.sample,"SC")

turn t into rawp

rawp2.hc.bp<-2\*(1-pt(abs(t2.hc.bp[,1]),59))

rawp2.hc.md<-2\*(1-pt(abs(t2.hc.md[,1]),77))

rawp2.hc.sc<-2\*(1-pt(abs(t2.hc.sc[,1]),57))

Run male/female separately

added gender to sample\_ordered3.txt

ordered.sample3<-read.delim("sample\_ordered3.txt",header=T,row.names=1,sep="\t")

ordered.sample644<- ordered.sample3[ordered.sample3$remove=="keep",1:16]

hc.sample<-ordered.sample644[430:538,]

norm.hc.m<-norm.hc[,hc.sample[,16]=="M"]

hc.sample.m<- hc.sample[hc.sample[,16]=="M",]

table(hc.sample.m$Disease)

BP C MD SC

11 36 27 10

t2.hc.md.m<-ttest.disease(norm.hc.m,hc.sample.m,"MD")

rawp2.hc.md.m<-2\*(1-pt(abs(t2.hc.md.m[,1]),61))

norm.hc.f<-norm.hc[,hc.sample[,16]=="F"]

hc.sample.f<- hc.sample[hc.sample[,16]=="F",]

table(hc.sample.f$Disease)

BP C MD SC

5 9 7 4

t2.hc.md.f<-ttest.disease(norm.hc.f,hc.sample.f,"MD")

rawp2.hc.md.f<-2\*(1-pt(abs(t2.hc.md.f[,1]),14))

Export data

hc.m<-as.matrix(cbind(t2.hc.md.m,rawp2.hc.md.m))

rownames(hc.m)<-rownames(acg)

colnames(hc.m)<-c("t.hc.md.m","fc.hc.md.m","rawp.hc.md.m")

write.table(hc.m,"hc\_male.txt",sep="\t")

hc.f<-as.matrix(cbind(t2.hc.md.f,rawp2.hc.md.f))

rownames(hc.f)<-rownames(acg)

colnames(hc.f)<-c("t.hc.md.f","fc.hc.md.f","rawp.hc.md.f")

write.table(hc.f,"hc\_female.txt",sep="\t")