Spearman Correlation Coefficients and Regression Models

Brett

January 04, 2018

Contents

Introduction:

This file compares the ability of functional parameters and *in silico* predictive tools to predict BrS1 and LQT3 penetrance. We include a test with glmer and show its limitations in our dataset. There is also a sensitivity analysis done at the end.

Define Functions

```
####### FUNCTIONS #########
calcPval=function(xName,yName,weightName,nPerms,new.mat2){
  # Pulls out variables

  x=new.mat2[,xName]
  y=new.mat2[,yName]
  w=new.mat2[,weightName]
  x2=x[!is.na(x)]
  y2=y[!is.na(x)]
  w2=w[!is.na(x)]
```

```
  # Calculate the real correlation
  realCorr=weightedCorr(x2,y2,method='spearman',weights=w2)
  # Do permutations, calculate fake correlations
  permutedCorrList=c()
  for(permNum in 1:nPerms){
    permutedX=sample(x2,length(x2),replace=FALSE)
    wCorrSim=weightedCorr(permutedX,y2,method='spearman',weights=w2)
    permutedCorrList=c(permutedCorrList,wCorrSim)
  }
  permutedCorrList2=abs(permutedCorrList)
  realCorr2=abs(realCorr)

  # Calculate pvalue
  summ=sum(realCorr2<permutedCorrList2)
  pValue=summ/nPerms
  return(list(realCorr,pValue,length(x2)))
}

calcAllPvals=function(yList,xList,nPerms,weightName,new.mat2){
  i=0
  resultTable=data.frame()
  for(yName in yList){
    for(xName in xList){
      i=i+1
      result=calcPval(xName,yName,weightName,nPerms,new.mat2)
      resultTable[i,'x']=xName
      resultTable[i,'y']=yName
      resultTable[i,'nPerms']=nPerms
      resultTable[i,'weightedCorr']=result[[1]]
      resultTable[i,'pValue']=result[[2]]
      resultTable[i,'n']=result[[3]]
      #print(resultTable[i,'pValue'])
    }
  }
  print(resultTable)
  return(resultTable)
}
```

Read in data

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

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

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

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

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

d$eaRate<-as.numeric(d$eaRate)
d$blastpssm<-as.numeric(d$blastpssm)
d$pamscore<-as.numeric(d$pamscore)

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

Summary Statistics

```
vari <- c('ipeak','ilate_norm','vhalfact','vhalfinact','recovfrominact','ilate')
t1 <- CreateTableOne(vars = vari, data=d)
print(t1, nonnormal=vari)
```

```
                              Overall
  n                            1712
  ipeak (median [IQR])        85.00 [42.93, 100.00]
  ilate_norm (median [IQR])    2.18 [1.90, 2.54]
  vhalfact (median [IQR])      0.33 [-1.50, 3.65]
  vhalfinact (median [IQR])   -0.17 [-5.10, 3.13]
  recovfrominact (median [IQR]) 2.00 [1.91, 2.12]
  ilate (median [IQR])         2.26 [2.00, 2.59]
```

Fitting data with generalized linear mixed effects model

standard deviation of error term suggest negligible influence of predictor ipeak (peak) current [or ilate (late current)]

```
e<-d[d$total_carriers>0,]
```

```
fit_partialpool_glmer <- glmer(cbind(brs1, total_carriers - brs1) ~ (1|var) + ilate, data = e, family =
```

```
print("glmer of BrS1 penetrance by late current (fixed, bad predictor) and variant (random intercept) co
```

```
[1] "glmer of BrS1 penetrance by late current (fixed, bad predictor) and variant (random intercept) cova
```

```
arm::display(fit_partialpool_glmer)
```

```
glmer(formula = cbind(brs1, total_carriers - brs1) ~ (1 | var) +
    ilate, data = e, family = binomial("logit"))
            coef.est coef.se
(Intercept) -4.06     0.00
ilate       -0.23     0.00

Error terms:
 Groups   Name        Std.Dev.
 var      (Intercept) 2.81
 Residual             1.00
---
number of obs: 96, groups: var, 96
AIC = 248.4, DIC = -205.6
deviance = 18.4
```

```
fit_partialpool_glmer <- glmer(cbind(brs1, total_carriers - brs1) ~ (1|var) + ipeak, data = e, family =
print("glmer of BrS1 penetrance by peak current (fixed, good predictor) and variant (random intercept) c
```

```
[1] "glmer of BrS1 penetrance by peak current (fixed, good predictor) and variant (random intercept) co
```

```
arm::display(fit_partialpool_glmer)
```

```
glmer(formula = cbind(brs1, total_carriers - brs1) ~ (1 | var) +
    ipeak, data = e, family = binomial("logit"), weights = weightsMilder)
            coef.est coef.se
(Intercept)  0.69     0.37
ipeak       -0.05     0.00

Error terms:
 Groups   Name        Std.Dev.
 var      (Intercept) 2.15
 Residual             1.00
---
number of obs: 231, groups: var, 231
AIC = 699.6, DIC = -544
deviance = 74.8
```

```
e<-d
```

Rank order correlations of functional parameters and predictive models to penetrance

This part of the script compares Spearman rho's calculated using each functional parameter and predictive model against penetrance of BrS1 or LQT3 or both BrS1 and LQT3 (broad pathogenic classification)

```r
#only look at true nsSNPs
d<-d[d$mut_type == "missense", ]

# look only at ones with measured peak currents.
d<-d[!is.na(d$ipeak) | !is.na(d$ilate) | !is.na(d$vhalfact)
     | !is.na(d$vhalfinact) | !is.na(d$recovfrominact), ]
d<-d[d$total_carriers>0, ]

yList=c('BrS_penetranceBayesian','LQT_penetranceBayesian', 'all_penetranceBayesian')
xList=c('ipeak','ilate_norm','ilate','vhalfact','vhalfinact', 'tauinact',
        'recovfrominact','eaRate',
        'pph2_prob', 'SIFT.Score', 'provean_score', 'CADD_raw', 'blastpssm',
        'pamscore', 'aasimilaritymat')
resultTable<-calcAllPvals(yList, xList, 1000, 'weight', d)
```

```
                 x                      y nPerms weightedCorr pValue   n
1            ipeak BrS_penetranceBayesian   1000 -0.432395099  0.000 212
2       ilate_norm BrS_penetranceBayesian   1000 -0.052577234  0.674  86
3            ilate BrS_penetranceBayesian   1000  0.002894086  0.980  91
4         vhalfact BrS_penetranceBayesian   1000  0.314510626  0.000 170
5       vhalfinact BrS_penetranceBayesian   1000 -0.141168595  0.082 193
6         tauinact BrS_penetranceBayesian   1000 -0.488689629  0.006  38
7   recovfrominact BrS_penetranceBayesian   1000  0.170253623  0.088 127
8           eaRate BrS_penetranceBayesian   1000 -0.288977033  0.000 224
9        pph2_prob BrS_penetranceBayesian   1000  0.362811331  0.000 221
10      SIFT.Score BrS_penetranceBayesian   1000 -0.347997598  0.000 219
11   provean_score BrS_penetranceBayesian   1000 -0.340840589  0.000 221
12        CADD_raw BrS_penetranceBayesian   1000  0.367834514  0.000 189
13       blastpssm BrS_penetranceBayesian   1000 -0.201502537  0.004 224
14        pamscore BrS_penetranceBayesian   1000 -0.077109126  0.314 224
15 aasimilaritymat BrS_penetranceBayesian   1000 -0.011559075  0.884 224
16           ipeak LQT_penetranceBayesian   1000  0.153270454  0.049 212
17      ilate_norm LQT_penetranceBayesian   1000  0.331905518  0.005  86
18           ilate LQT_penetranceBayesian   1000  0.370512712  0.000  91
19        vhalfact LQT_penetranceBayesian   1000 -0.086550511  0.331 170
20      vhalfinact LQT_penetranceBayesian   1000  0.033956540  0.685 193
21        tauinact LQT_penetranceBayesian   1000  0.195381366  0.273  38
22  recovfrominact LQT_penetranceBayesian   1000 -0.251931833  0.010 127
23          eaRate LQT_penetranceBayesian   1000 -0.264892504  0.000 224
24       pph2_prob LQT_penetranceBayesian   1000  0.181694558  0.015 221
25      SIFT.Score LQT_penetranceBayesian   1000 -0.243096970  0.000 219
26   provean_score LQT_penetranceBayesian   1000 -0.228014375  0.007 221
27        CADD_raw LQT_penetranceBayesian   1000  0.116762700  0.156 189
28       blastpssm LQT_penetranceBayesian   1000 -0.077575273  0.293 224
29        pamscore LQT_penetranceBayesian   1000 -0.001605959  0.992 224
30 aasimilaritymat LQT_penetranceBayesian   1000 -0.029413594  0.707 224
31           ipeak all_penetranceBayesian   1000 -0.204633910  0.008 212
32      ilate_norm all_penetranceBayesian   1000  0.282178688  0.023  86
```

```
33           ilate all_penetranceBayesian   1000  0.325680430  0.006  91
34        vhalfact all_penetranceBayesian   1000  0.183807186  0.031 170
35      vhalfinact all_penetranceBayesian   1000 -0.092979099  0.273 193
36        tauinact all_penetranceBayesian   1000  0.052558070  0.766  38
37  recovfrominact all_penetranceBayesian   1000 -0.119352761  0.269 127
38          eaRate all_penetranceBayesian   1000 -0.462128452  0.000 224
39       pph2_prob all_penetranceBayesian   1000  0.432613792  0.000 221
40      SIFT.Score all_penetranceBayesian   1000 -0.476206196  0.000 219
41   provean_score all_penetranceBayesian   1000 -0.461417485  0.000 221
42        CADD_raw all_penetranceBayesian   1000  0.409895503  0.000 189
43       blastpssm all_penetranceBayesian   1000 -0.243124363  0.000 224
44        pamscore all_penetranceBayesian   1000 -0.045037086  0.538 224
45 aasimilaritymat all_penetranceBayesian   1000 -0.008107253  0.910 224
```

Predictive models

Simple linear models using peak current (ipeak), v1/2 activation, and PROVEAN (in silico predictive tool) to predict BrS1 and late current (ilate_norm), recovery from inactivation, and PROVEAN to predict LQT3 penetrance.

```
e<-e[e$total_carriers>0, ]

#only evaluate true  missense SNPs
e<-e[e$mut_type == "missense"
     & !is.na(e$provean_score), ]

# look only at ones with measured peak currents.
b<-e[!is.na(e$ipeak), ]


print("Now select only missense variants where some functional characterization exists. This is done so
```

```
[1] "Now select only missense variants where some functional characterization exists. This is done so t
```

```
print("restricted cubic spline of peak current predicting BrS1 penetrance")
```

```
[1] "restricted cubic spline of peak current predicting BrS1 penetrance"
```

```
ln_brs_peak<-lm(b$BrS_penetranceBayesian~rcs(b$ipeak,4), weights=b$weight)

print("restricted cubic spline of peak current and linear PROVEAN predicting BrS1 penetrance")
```

```
[1] "restricted cubic spline of peak current and linear PROVEAN predicting BrS1 penetrance"
```

```
ln_brs_peak_prov<-lm(b$BrS_penetranceBayesian~rcs(b$ipeak,4)+b$provean_score, weights=b$weight)

print("linear PROVEAN predicting BrS1 penetrance")
```

```
[1] "linear PROVEAN predicting BrS1 penetrance"
```

```
ln_brs_prov<-lm(b$BrS_penetranceBayesian~b$provean_score, weights=b$weight)

print("linear PROVEAN predicting LQT3 penetrance")
```

```
[1] "linear PROVEAN predicting LQT3 penetrance"
```

```r
ln_lqt_prov<-lm(b$LQT_penetranceBayesian~b$provean_score, weights=b$weight)

print("restricted cubic spline of peak current predicting BrS1 penetrance")

## [1] "restricted cubic spline of peak current predicting BrS1 penetrance"

ln_lqt_peak<-lm(b$LQT_penetranceBayesian~rcs(b$ipeak,4), weights=b$weight)

print("restricted cubic spline of peak current and linear PROVEAN predicting LQT penetrance")

## [1] "restricted cubic spline of peak current and linear PROVEAN predicting LQT penetrance"

ln_lqt_peak_prov<-lm(b$LQT_penetranceBayesian~rcs(b$ipeak,4)+b$provean_score, weights=b$weight)

summary(ln_lqt_peak_prov)
```

```
Call:
lm(formula = b$LQT_penetranceBayesian ~ rcs(b$ipeak, 4) + b$provean_score,
    weights = b$weight)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-0.29218 -0.13045 -0.04416  0.05859  0.66664

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)        -8.232e-02  6.056e-02  -1.359   0.1756
rcs(b$ipeak, 4)b   -1.392e-05  1.606e-03  -0.009   0.9931
rcs(b$ipeak, 4)b'   4.983e-03  3.054e-03   1.632   0.1043
rcs(b$ipeak, 4)b'' -3.228e-02  1.770e-02  -1.824   0.0696 .
b$provean_score    -3.103e-02  6.825e-03  -4.547 9.31e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2005 on 204 degrees of freedom
Multiple R-squared:  0.1523,    Adjusted R-squared:  0.1357
F-statistic: 9.164 on 4 and 204 DF,  p-value: 7.915e-07
```

```r
summary(ln_brs_peak_prov)
```

```
Call:
lm(formula = b$BrS_penetranceBayesian ~ rcs(b$ipeak, 4) + b$provean_score,
    weights = b$weight)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-0.42889 -0.08864 -0.01031  0.08161  0.55074

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         0.504525   0.050334  10.024  < 2e-16 ***
rcs(b$ipeak, 4)b   -0.008023   0.001335  -6.011 8.39e-09 ***
rcs(b$ipeak, 4)b'   0.007466   0.002538   2.941 0.003644 **
rcs(b$ipeak, 4)b'' -0.026489   0.014711  -1.801 0.073233 .
b$provean_score    -0.021590   0.005672  -3.806 0.000187 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1666 on 204 degrees of freedom
Multiple R-squared:  0.413, Adjusted R-squared:  0.4015
F-statistic: 35.89 on 4 and 204 DF,  p-value: < 2.2e-16
```

```
summary(ln_brs_prov)
```

```
Call:
lm(formula = b$BrS_penetranceBayesian ~ b$provean_score, weights = b$weight)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-0.36119 -0.13444 -0.03966  0.10680  0.58310

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.07789    0.02833    2.75  0.00649 **
b$provean_score -0.03700    0.00656   -5.64 5.54e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.201 on 207 degrees of freedom
Multiple R-squared:  0.1332,    Adjusted R-squared:  0.129
F-statistic: 31.81 on 1 and 207 DF,  p-value: 5.541e-08
```

```
summary(ln_lqt_prov)
```

```
Call:
lm(formula = b$LQT_penetranceBayesian ~ b$provean_score, weights = b$weight)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-0.24705 -0.12393 -0.05755  0.05844  0.76386

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.084281   0.029727   2.835  0.00503 **
b$provean_score -0.022160   0.006884  -3.219  0.00149 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2109 on 207 degrees of freedom
Multiple R-squared:  0.04767,   Adjusted R-squared:  0.04307
F-statistic: 10.36 on 1 and 207 DF,  p-value: 0.001493
```

```
anova(ln_brs_peak_prov)
```

```
Analysis of Variance Table

Response: b$BrS_penetranceBayesian
                 Df Sum Sq Mean Sq F value    Pr(>F)
rcs(b$ipeak, 4)   3 3.5829 1.19430  43.018 < 2.2e-16 ***
b$provean_score   1 0.4022 0.40222  14.488 0.0001865 ***
```

```
Residuals       204 5.6636 0.02776
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(ln_lqt_peak_prov)
```

```
Analysis of Variance Table

Response: b$LQT_penetranceBayesian
                 Df Sum Sq Mean Sq F value    Pr(>F)
rcs(b$ipeak, 4)   3 0.6421 0.21404  5.3257  0.001491 **
b$provean_score   1 0.8311 0.83105 20.6778 9.311e-06 ***
Residuals       204 8.1989 0.04019
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
AIC(ln_brs_prov)
```

```
[1] 52.47538
```

```
AIC(ln_brs_peak_prov)
```

```
[1] -22.99824
```

```
AIC(ln_brs_peak)
```

```
[1] -10.65898
```

```
AIC(ln_lqt_peak_prov)
```

```
[1] 54.31715
```

```
AIC(ln_lqt_peak)
```

```
[1] 72.49557
```

```
AIC(ln_lqt_prov)
```

```
[1] 72.64372
```

```
#only include variants with measured late current
l<-e[!is.na(e$ilate), ]
```

```
print("linear PROVEAN predicting LQT3 penetrance")
```

```
[1] "linear PROVEAN predicting LQT3 penetrance"
```

```
ln_lqt_prov<-lm(l$LQT_penetranceBayesian~l$provean_score, weights=l$weight)
```

```
print("linear non-normalized late current predicting LQT3 penetrance")
```

```
[1] "linear non-normalized late current predicting LQT3 penetrance"
```

```
ln_lqt_late <-lm(l$LQT_penetranceBayesian~l$ilate, weights=l$weight)
```

```
print("linear non-normalized late current and PROVEAN predicting LQT3 penetrance")
```

```
[1] "linear non-normalized late current and PROVEAN predicting LQT3 penetrance"
```

```
ln_lqt_late_prov <-lm(l$LQT_penetranceBayesian~l$ilate+l$provean_score, weights=l$weight)
```

```
print("linear PROVEAN predicting BrS3 penetrance")
```

```
[1] "linear PROVEAN predicting BrS3 penetrance"
```

```r
ln_brs_prov<-lm(l$BrS_penetranceBayesian~l$provean_score, weights=l$weight)
```

```r
print("linear non-normalized late current predicting BrS3 penetrance")
```

```
[1] "linear non-normalized late current predicting BrS3 penetrance"
```

```r
ln_brs_late <-lm(l$BrS_penetranceBayesian~l$ilate, weights=l$weight)
```

```r
print("linear non-normalized late current and PROVEAN predicting BrS3 penetrance")
```

```
[1] "linear non-normalized late current and PROVEAN predicting BrS3 penetrance"
```

```r
ln_brs_late_prov <-lm(l$BrS_penetranceBayesian~l$ilate+l$provean_score, weights=l$weight)
```

```r
summary(ln_lqt_late_prov)
```

```
Call:
lm(formula = l$LQT_penetranceBayesian ~ l$ilate + l$provean_score,
    weights = l$weight)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-0.54425 -0.13894  0.00116  0.12835  0.64589

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     -0.37637    0.16316  -2.307 0.023469 *
l$ilate          0.20804    0.07430   2.800 0.006305 **
l$provean_score -0.04768    0.01253  -3.806 0.000264 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2271 on 86 degrees of freedom
Multiple R-squared:  0.2826,    Adjusted R-squared:  0.266
F-statistic: 16.94 on 2 and 86 DF,  p-value: 6.259e-07
```

```r
summary(ln_brs_late_prov)
```

```
Call:
lm(formula = l$BrS_penetranceBayesian ~ l$ilate + l$provean_score,
    weights = l$weight)

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-0.15739 -0.05541 -0.02371  0.02739  0.67822

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.075872   0.093294   0.813   0.4183
l$ilate         -0.019445   0.042483  -0.458   0.6483
l$provean_score -0.018070   0.007163  -2.523   0.0135 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1298 on 86 degrees of freedom
```

```
Multiple R-squared:  0.07095,   Adjusted R-squared:  0.04934
F-statistic: 3.284 on 2 and 86 DF,  p-value: 0.04224
```

```r
anova(ln_brs_late_prov)
```

```
Analysis of Variance Table

Response: l$BrS_penetranceBayesian
                Df  Sum Sq  Mean Sq F value  Pr(>F)
l$ilate          1 0.00342 0.003424  0.2031 0.65339
l$provean_score  1 0.10730 0.107295  6.3642 0.01349 *
Residuals       86 1.44989 0.016859
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(ln_lqt_late_prov)
```

```
Analysis of Variance Table

Response: l$LQT_penetranceBayesian
                Df Sum Sq Mean Sq F value    Pr(>F)
l$ilate          1 1.0003 1.00033  19.400 3.038e-05 ***
l$provean_score  1 0.7469 0.74689  14.485  0.000264 ***
Residuals       86 4.4344 0.05156
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
AIC(ln_brs_prov)
```

```
[1] -53.36812
```

```r
AIC(ln_brs_late_prov)
```

```
[1] -51.58466
```

```r
AIC(ln_brs_late)
```

```
[1] -47.23078
```

```r
AIC(ln_lqt_late_prov)
```

```
[1] 47.90906
```

```r
AIC(ln_lqt_late)
```

```
[1] 59.7629
```

```r
AIC(ln_lqt_prov)
```

```
[1] 53.67492
```

Correlation between functional parameters and *in silico* predictive tools

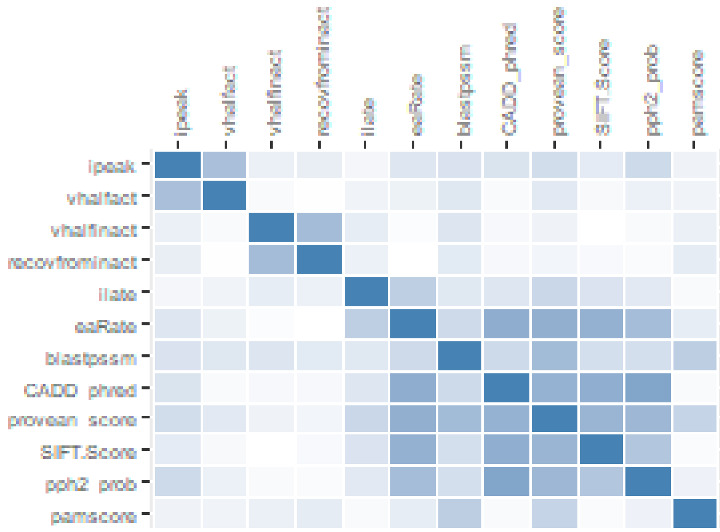
```r
rcor<-corr.test(d[c(15:17,19:22,40,47,50,53,23)], method = "pearson")
rcor$r<-abs(rcor$r)

p<-ggplot(melt(rcor$r), aes(Var1,ordered(Var2, levels = rev(sort(unique(Var2))))))+ geom_tile(aes(fill =

#ordered(Var2, levels = rev(sort(unique(Var2))))
```

```
base_size<-9

p+ theme_grey(base_size = base_size) + labs(x = "", y = "") +
  scale_x_discrete(expand = c(0, 0)) +
  scale_y_discrete(expand = c(0, 0))  + theme(legend.position = "none", axis.text.x.top = element_text(
```



Sensitivity Analysis

This part of the script compares Spearman rho's calculated using each functional parameter and predictive model against penetrance of BrS1 or LQT3 or both BrS1 and LQT3 (more like the broad pathogenic classification) using different priors (empirical Bayes, optimistic, uninformative, and pessimistic)

Empirical Bayes (used in manuscript)

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

#only look at true nsSNPs
d<-d[d$mut_type == "missense", ]

# look only at ones with measured peak currents.
d<-d[!is.na(d$ipeak) | !is.na(d$ilate) | !is.na(d$vhalfact)
     | !is.na(d$vhalfinact) | !is.na(d$recovfrominact), ]
d<-d[d$total_carriers>0, ]

yList=c('BrS_penetranceBayesian','LQT_penetranceBayesian', 'all_penetranceBayesian')
xList=c('ipeak','ilate_norm','ilate','vhalfact','vhalfinact', 'tauinact',
        'recovfrominact','eaRate',
```

```
        'pph2_prob', 'SIFT.Score', 'provean_score', 'CADD_raw', 'blastpssm',
        'pamscore', 'aasimilaritymat')
resultTable<-calcAllPvals(yList, xList, 1000, 'weight', d)

                 x                       y nPerms weightedCorr pValue   n
1            ipeak BrS_penetranceBayesian   1000 -0.432395099  0.000 212
2        ilate_norm BrS_penetranceBayesian   1000 -0.052577234  0.700  86
3            ilate BrS_penetranceBayesian   1000  0.002894086  0.984  91
4          vhalfact BrS_penetranceBayesian   1000  0.314510626  0.000 170
5        vhalfinact BrS_penetranceBayesian   1000 -0.141168595  0.083 193
6          tauinact BrS_penetranceBayesian   1000 -0.488689629  0.001  38
7     recovfrominact BrS_penetranceBayesian   1000  0.170253623  0.096 127
8            eaRate BrS_penetranceBayesian   1000 -0.288977033  0.000 224
9         pph2_prob BrS_penetranceBayesian   1000  0.362811331  0.000 221
10       SIFT.Score BrS_penetranceBayesian   1000 -0.347997598  0.000 219
11    provean_score BrS_penetranceBayesian   1000 -0.340840589  0.000 221
12         CADD_raw BrS_penetranceBayesian   1000  0.367834514  0.000 189
13        blastpssm BrS_penetranceBayesian   1000 -0.201502537  0.005 224
14         pamscore BrS_penetranceBayesian   1000 -0.077109126  0.306 224
15 aasimilaritymat BrS_penetranceBayesian   1000 -0.011559075  0.905 224
16            ipeak LQT_penetranceBayesian   1000  0.153270454  0.042 212
17       ilate_norm LQT_penetranceBayesian   1000  0.331905518  0.009  86
18            ilate LQT_penetranceBayesian   1000  0.370512712  0.003  91
19         vhalfact LQT_penetranceBayesian   1000 -0.086550511  0.290 170
20       vhalfinact LQT_penetranceBayesian   1000  0.033956540  0.667 193
21         tauinact LQT_penetranceBayesian   1000  0.195381366  0.306  38
22    recovfrominact LQT_penetranceBayesian   1000 -0.251931833  0.011 127
23           eaRate LQT_penetranceBayesian   1000 -0.264892504  0.000 224
24        pph2_prob LQT_penetranceBayesian   1000  0.181694558  0.008 221
25       SIFT.Score LQT_penetranceBayesian   1000 -0.243096970  0.004 219
26    provean_score LQT_penetranceBayesian   1000 -0.228014375  0.006 221
27         CADD_raw LQT_penetranceBayesian   1000  0.116762700  0.153 189
28        blastpssm LQT_penetranceBayesian   1000 -0.077575273  0.305 224
29         pamscore LQT_penetranceBayesian   1000 -0.001605959  0.990 224
30 aasimilaritymat LQT_penetranceBayesian   1000 -0.029413594  0.701 224
31            ipeak all_penetranceBayesian   1000 -0.204633910  0.006 212
32       ilate_norm all_penetranceBayesian   1000  0.282178688  0.023  86
33            ilate all_penetranceBayesian   1000  0.325680430  0.007  91
34         vhalfact all_penetranceBayesian   1000  0.183807186  0.041 170
35       vhalfinact all_penetranceBayesian   1000 -0.092979099  0.237 193
36         tauinact all_penetranceBayesian   1000  0.052558070  0.781  38
37    recovfrominact all_penetranceBayesian   1000 -0.119352761  0.227 127
38           eaRate all_penetranceBayesian   1000 -0.462128452  0.000 224
39        pph2_prob all_penetranceBayesian   1000  0.432613792  0.000 221
40       SIFT.Score all_penetranceBayesian   1000 -0.476206196  0.000 219
41    provean_score all_penetranceBayesian   1000 -0.461417485  0.000 221
42         CADD_raw all_penetranceBayesian   1000  0.409895503  0.000 189
43        blastpssm all_penetranceBayesian   1000 -0.243124363  0.004 224
44         pamscore all_penetranceBayesian   1000 -0.045037086  0.518 224
45 aasimilaritymat all_penetranceBayesian   1000 -0.008107253  0.907 224
```

Uninformative Prior

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

#only look at true nsSNPs
d<-d[d$mut_type == "missense", ]

# look only at ones with measured peak currents.
d<-d[!is.na(d$ipeak) | !is.na(d$ilate) | !is.na(d$vhalfact)
     | !is.na(d$vhalfinact) | !is.na(d$recovfrominact), ]
d<-d[d$total_carriers>0, ]

yList=c('BrS_penetranceBayesian','LQT_penetranceBayesian', 'all_penetranceBayesian')
xList=c('ipeak','ilate_norm','ilate','vhalfact','vhalfinact', 'tauinact',
        'recovfrominact','eaRate',
        'pph2_prob', 'SIFT.Score', 'provean_score', 'CADD_raw', 'blastpssm',
        'pamscore', 'aasimilaritymat')
resultTable<-calcAllPvals(yList, xList, 1000, 'weight', d)
```

```
                 x                      y nPerms  weightedCorr pValue   n
1            ipeak BrS_penetranceBayesian   1000 -0.4097305972  0.000 212
2       ilate_norm BrS_penetranceBayesian   1000 -0.0416252589  0.707  86
3            ilate BrS_penetranceBayesian   1000  0.0154447825  0.895  91
4         vhalfact BrS_penetranceBayesian   1000  0.3035426908  0.002 170
5       vhalfinact BrS_penetranceBayesian   1000 -0.1347798510  0.099 193
6         tauinact BrS_penetranceBayesian   1000 -0.4163036936  0.024  38
7   recovfrominact BrS_penetranceBayesian   1000  0.1644912311  0.098 127
8           eaRate BrS_penetranceBayesian   1000 -0.3016164840  0.000 224
9        pph2_prob BrS_penetranceBayesian   1000  0.3554203824  0.000 221
10      SIFT.Score BrS_penetranceBayesian   1000 -0.3703975551  0.000 219
11   provean_score BrS_penetranceBayesian   1000 -0.3602875344  0.000 221
12        CADD_raw BrS_penetranceBayesian   1000  0.3658062748  0.000 189
13       blastpssm BrS_penetranceBayesian   1000 -0.2192574218  0.001 224
14        pamscore BrS_penetranceBayesian   1000 -0.0748580781  0.337 224
15 aasimilaritymat BrS_penetranceBayesian   1000 -0.0182305646  0.808 224
16           ipeak LQT_penetranceBayesian   1000  0.0954991259  0.217 212
17      ilate_norm LQT_penetranceBayesian   1000  0.2797235135  0.019  86
18           ilate LQT_penetranceBayesian   1000  0.3143608957  0.011  91
19        vhalfact LQT_penetranceBayesian   1000  0.0147623150  0.871 170
20      vhalfinact LQT_penetranceBayesian   1000 -0.0005762637  0.995 193
21        tauinact LQT_penetranceBayesian   1000  0.1813224725  0.345  38
22  recovfrominact LQT_penetranceBayesian   1000 -0.1959343754  0.053 127
23          eaRate LQT_penetranceBayesian   1000 -0.3090196712  0.001 224
24       pph2_prob LQT_penetranceBayesian   1000  0.2146542516  0.005 221
25      SIFT.Score LQT_penetranceBayesian   1000 -0.3114920429  0.000 219
26   provean_score LQT_penetranceBayesian   1000 -0.2543135244  0.002 221
27        CADD_raw LQT_penetranceBayesian   1000  0.1727755179  0.036 189
28       blastpssm LQT_penetranceBayesian   1000 -0.0994261011  0.172 224
```

```
29        pamscore LQT_penetranceBayesian   1000  0.0307270496  0.668 224
30 aasimilaritymat LQT_penetranceBayesian   1000  0.0122922162  0.866 224
31           ipeak all_penetranceBayesian   1000 -0.1809800695  0.019 212
32      ilate_norm all_penetranceBayesian   1000  0.2744720167  0.021  86
33           ilate all_penetranceBayesian   1000  0.2935753727  0.011  91
34        vhalfact all_penetranceBayesian   1000  0.1881686095  0.025 170
35      vhalfinact all_penetranceBayesian   1000 -0.0880001370  0.299 193
36        tauinact all_penetranceBayesian   1000  0.0797407455  0.665  38
37  recovfrominact all_penetranceBayesian   1000 -0.1011948276  0.341 127
38          eaRate all_penetranceBayesian   1000 -0.4523216627  0.000 224
39       pph2_prob all_penetranceBayesian   1000  0.4008390299  0.000 221
40      SIFT.Score all_penetranceBayesian   1000 -0.4807688741  0.000 219
41   provean_score all_penetranceBayesian   1000 -0.4497257122  0.000 221
42        CADD_raw all_penetranceBayesian   1000  0.3884429761  0.000 189
43       blastpssm all_penetranceBayesian   1000 -0.2397772775  0.002 224
44        pamscore all_penetranceBayesian   1000 -0.0338604256  0.653 224
45 aasimilaritymat all_penetranceBayesian   1000  0.0052081714  0.952 224
```

Optimistic Prior

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

#only look at true nsSNPs
d<-d[d$mut_type == "missense", ]

# look only at ones with measured peak currents.
d<-d[!is.na(d$ipeak) | !is.na(d$ilate) | !is.na(d$vhalfact)
     | !is.na(d$vhalfinact) | !is.na(d$recovfrominact), ]
d<-d[d$total_carriers>0, ]

yList=c('BrS_penetranceBayesian','LQT_penetranceBayesian', 'all_penetranceBayesian')
xList=c('ipeak','ilate_norm','ilate','vhalfact','vhalfinact', 'tauinact',
        'recovfrominact','eaRate',
        'pph2_prob', 'SIFT.Score', 'provean_score', 'CADD_raw', 'blastpssm',
        'pamscore', 'aasimilaritymat')
resultTable<-calcAllPvals(yList, xList, 1000, 'weight', d)
```

```
                 x                      y nPerms weightedCorr pValue   n
1            ipeak BrS_penetranceBayesian   1000 -0.419994011  0.000 212
2       ilate_norm BrS_penetranceBayesian   1000 -0.006503513  0.959  86
3            ilate BrS_penetranceBayesian   1000  0.009142985  0.943  91
4         vhalfact BrS_penetranceBayesian   1000  0.319816932  0.000 170
5       vhalfinact BrS_penetranceBayesian   1000 -0.127847792  0.113 193
6         tauinact BrS_penetranceBayesian   1000 -0.569624540  0.000  38
7   recovfrominact BrS_penetranceBayesian   1000  0.165529102  0.089 127
8           eaRate BrS_penetranceBayesian   1000 -0.240158906  0.001 224
9        pph2_prob BrS_penetranceBayesian   1000  0.326179625  0.000 221
10      SIFT.Score BrS_penetranceBayesian   1000 -0.264293366  0.000 219
```

```
11   provean_score BrS_penetranceBayesian   1000 -0.303458966  0.000 221
12        CADD_raw BrS_penetranceBayesian   1000  0.314565365  0.000 189
13       blastpssm BrS_penetranceBayesian   1000 -0.185196101  0.011 224
14        pamscore BrS_penetranceBayesian   1000 -0.112860705  0.120 224
15 aasimilaritymat BrS_penetranceBayesian   1000 -0.036179647  0.628 224
16           ipeak LQT_penetranceBayesian   1000  0.254186711  0.000 212
17      ilate_norm LQT_penetranceBayesian   1000  0.376446534  0.000  86
18           ilate LQT_penetranceBayesian   1000  0.394177659  0.000  91
19        vhalfact LQT_penetranceBayesian   1000 -0.135896627  0.108 170
20      vhalfinact LQT_penetranceBayesian   1000  0.054139314  0.494 193
21        tauinact LQT_penetranceBayesian   1000  0.202430571  0.250  38
22  recovfrominact LQT_penetranceBayesian   1000 -0.267312294  0.009 127
23          eaRate LQT_penetranceBayesian   1000 -0.209830055  0.002 224
24       pph2_prob LQT_penetranceBayesian   1000  0.103158604  0.166 221
25      SIFT.Score LQT_penetranceBayesian   1000 -0.182333638  0.011 219
26   provean_score LQT_penetranceBayesian   1000 -0.175679426  0.015 221
27        CADD_raw LQT_penetranceBayesian   1000  0.081572462  0.320 189
28       blastpssm LQT_penetranceBayesian   1000 -0.058755370  0.452 224
29        pamscore LQT_penetranceBayesian   1000 -0.010162547  0.894 224
30 aasimilaritymat LQT_penetranceBayesian   1000 -0.022839129  0.765 224
31           ipeak all_penetranceBayesian   1000 -0.207937773  0.008 212
32      ilate_norm all_penetranceBayesian   1000  0.297517339  0.026  86
33           ilate all_penetranceBayesian   1000  0.339702059  0.003  91
34        vhalfact all_penetranceBayesian   1000  0.161217627  0.065 170
35      vhalfinact all_penetranceBayesian   1000 -0.089402399  0.256 193
36        tauinact all_penetranceBayesian   1000  0.062202418  0.750  38
37  recovfrominact all_penetranceBayesian   1000 -0.135405438  0.189 127
38          eaRate all_penetranceBayesian   1000 -0.452095884  0.000 224
39       pph2_prob all_penetranceBayesian   1000  0.439746872  0.000 221
40      SIFT.Score all_penetranceBayesian   1000 -0.458422446  0.000 219
41   provean_score all_penetranceBayesian   1000 -0.463392235  0.000 221
42        CADD_raw all_penetranceBayesian   1000  0.408471214  0.000 189
43       blastpssm all_penetranceBayesian   1000 -0.238526349  0.000 224
44        pamscore all_penetranceBayesian   1000 -0.052582463  0.469 224
45 aasimilaritymat all_penetranceBayesian   1000 -0.032960709  0.667 224
```

Pessimistic Prior

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

#only look at true nsSNPs
d<-d[d$mut_type == "missense", ]

# look only at ones with measured peak currents.
d<-d[!is.na(d$ipeak) | !is.na(d$ilate) | !is.na(d$vhalfact)
     | !is.na(d$vhalfinact) | !is.na(d$recovfrominact), ]
d<-d[d$total_carriers>0, ]
```

```r
yList=c('BrS_penetranceBayesian','LQT_penetranceBayesian', 'all_penetranceBayesian')
xList=c('ipeak','ilate_norm','ilate','vhalfact','vhalfinact', 'tauinact',
        'recovfrominact','eaRate',
        'pph2_prob', 'SIFT.Score', 'provean_score', 'CADD_raw', 'blastpssm',
        'pamscore', 'aasimilaritymat')
resultTable<-calcAllPvals(yList, xList, 1000, 'weight', d)
```

```
                x                      y nPerms weightedCorr pValue   n
1           ipeak BrS_penetranceBayesian   1000 -0.409730597  0.000 212
2      ilate_norm BrS_penetranceBayesian   1000 -0.041625259  0.749  86
3           ilate BrS_penetranceBayesian   1000  0.015444783  0.890  91
4        vhalfact BrS_penetranceBayesian   1000  0.303542691  0.000 170
5      vhalfinact BrS_penetranceBayesian   1000 -0.134779851  0.093 193
6        tauinact BrS_penetranceBayesian   1000 -0.416303694  0.024  38
7  recovfrominact BrS_penetranceBayesian   1000  0.164491231  0.100 127
8          eaRate BrS_penetranceBayesian   1000 -0.301616484  0.000 224
9       pph2_prob BrS_penetranceBayesian   1000  0.355420382  0.000 221
10     SIFT.Score BrS_penetranceBayesian   1000 -0.370397555  0.000 219
11  provean_score BrS_penetranceBayesian   1000 -0.360287534  0.000 221
12       CADD_raw BrS_penetranceBayesian   1000  0.365806275  0.000 189
13      blastpssm BrS_penetranceBayesian   1000 -0.219257422  0.001 224
14       pamscore BrS_penetranceBayesian   1000 -0.074858078  0.303 224
15 aasimilaritymat BrS_penetranceBayesian  1000 -0.018230565  0.823 224
16          ipeak LQT_penetranceBayesian   1000  0.088774099  0.277 212
17     ilate_norm LQT_penetranceBayesian   1000  0.279730301  0.018  86
18          ilate LQT_penetranceBayesian   1000  0.316531219  0.007  91
19       vhalfact LQT_penetranceBayesian   1000  0.020158523  0.819 170
20     vhalfinact LQT_penetranceBayesian   1000  0.009316738  0.911 193
21       tauinact LQT_penetranceBayesian   1000  0.193803435  0.274  38
22 recovfrominact LQT_penetranceBayesian   1000 -0.192793322  0.061 127
23         eaRate LQT_penetranceBayesian   1000 -0.304808567  0.000 224
24      pph2_prob LQT_penetranceBayesian   1000  0.210548835  0.006 221
25     SIFT.Score LQT_penetranceBayesian   1000 -0.314254374  0.000 219
26  provean_score LQT_penetranceBayesian   1000 -0.258894241  0.001 221
27       CADD_raw LQT_penetranceBayesian   1000  0.168082642  0.032 189
28      blastpssm LQT_penetranceBayesian   1000 -0.099298123  0.191 224
29       pamscore LQT_penetranceBayesian   1000  0.026535198  0.724 224
30 aasimilaritymat LQT_penetranceBayesian  1000  0.011233180  0.859 224
31          ipeak all_penetranceBayesian   1000 -0.182971376  0.018 212
32     ilate_norm all_penetranceBayesian   1000  0.264096455  0.034  86
33          ilate all_penetranceBayesian   1000  0.285111042  0.018  91
34       vhalfact all_penetranceBayesian   1000  0.177173412  0.033 170
35     vhalfinact all_penetranceBayesian   1000 -0.066731962  0.417 193
36       tauinact all_penetranceBayesian   1000  0.080552702  0.652  38
37 recovfrominact all_penetranceBayesian   1000 -0.085437817  0.400 127
38         eaRate all_penetranceBayesian   1000 -0.437839401  0.000 224
39      pph2_prob all_penetranceBayesian   1000  0.389447747  0.000 221
40     SIFT.Score all_penetranceBayesian   1000 -0.469993561  0.000 219
41  provean_score all_penetranceBayesian   1000 -0.449315304  0.000 221
42       CADD_raw all_penetranceBayesian   1000  0.376383173  0.000 189
43      blastpssm all_penetranceBayesian   1000 -0.236116858  0.000 224
44       pamscore all_penetranceBayesian   1000 -0.035031715  0.633 224
45 aasimilaritymat all_penetranceBayesian  1000 -0.001015510  0.990 224
```