

```
In [1]: include("/home/nicole/Jupyter/SSBRnoJ/src/SSBR.jl")
        using SSBR
```

```
In [2]: function getPos(ped,IDs)
        posAi = Array{Int64,1}(size(IDs,1))
        for (i,id) = enumerate(IDs[:,1])
            posAi[i] = ped.idMap[id].seqID
        end
        return posAi
    end
```

```
Out[2]: getPos (generic function with 1 method)
```

```
In [3]: ; cd Data/0.5a0/G/3

/home/nicole/Jupyter/JG3/Data/0.5a0/G/3
```

```
In [4]: ;ls

Correlation.G5.G.JC.txt
G0.Genotype.ID
G0.ID
G0.noGenotype.ID
G1.Genotype.ID
G1.ID
G1.noGenotype.ID
G2.Genotype.ID
G2.ID
G2.noGenotype.ID
G3.Genotype.ID
G3.ID
G3.noGenotype.ID
G4.Genotype.ID
G4.ID
G4.noGenotype.ID
G5.Genotype.ID
G5.ID
G5.noGenotype.ID
GenNF.txt
```

```
In [5]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
```

```
In [6]: ;awk '{print $1}' GenNF.txt | sort -b > genotype.ID
```

```
In [7]: ;join -v1 all.ID genotype.ID > noGenotype.ID
```

```
In [8]: ;awk '{print $1,$2}' Phe.txt > sim.phenotype
```

```
In [9]: ;awk '{print $1,$3}' PheAll.txt > sim.bv
```

```
In [10]: ; awk 'NR >=1 && NR <=8000 {print $1}' PedAll.txt | sort -b > G0.ID
```

```
In [11]: ; awk 'NR >=8001 && NR <=16000 {print $1}' PedAll.txt | sort -b > G1.ID
```

```
In [12]: ; awk 'NR >=16001 && NR <=24000 {print $1}' PedAll.txt | sort -b > G2.ID
```

```
In [13]: ; awk 'NR >=24001 && NR <=32000 {print $1}' PedAll.txt | sort -b > G3.ID
```

```
In [14]: ; awk 'NR >=32001 && NR <=40000 {print $1}' PedAll.txt | sort -b > G4.ID
```

```
In [15]: ; awk 'NR >=40001 && NR <=48000 {print $1}' PedAll.txt | sort -b > G5.ID
```

```
In [16]: ;join G0.ID genotype.ID > G0.Genotype.ID
```

```
In [17]: ;join G1.ID genotype.ID > G1.Genotype.ID
```

```
In [18]: ;join G2.ID genotype.ID > G2.Genotype.ID
```

```
In [19]: ;join G3.ID genotype.ID > G3.Genotype.ID
```

```
In [20]: ;join G4.ID genotype.ID > G4.Genotype.ID
```

```
In [21]: ;join G5.ID genotype.ID > G5.Genotype.ID
```

```
In [22]: ;join -v1 G0.ID genotype.ID > G0.noGenotype.ID
```

```
In [23]: ;join -v1 G1.ID genotype.ID > G1.noGenotype.ID
```

```
In [24]: ;join -v1 G2.ID genotype.ID > G2.noGenotype.ID
```

```
In [25]: ;join -v1 G3.ID genotype.ID > G3.noGenotype.ID
```

```
In [26]: ;join -v1 G4.ID genotype.ID > G4.noGenotype.ID
```

```
In [27]: ;join -v1 G5.ID genotype.ID > G5.noGenotype.ID
```

```
In [28]: ;wc G0.Genotype.ID;wc G1.Genotype.ID;wc G2.Genotype.ID;wc G3.Genotype.ID;wc G4.Genotype.ID;wc G5.Genotype.ID
200 200 1200 G0.Genotype.ID
200 200 1200 G1.Genotype.ID
200 200 1200 G2.Genotype.ID
200 200 1200 G3.Genotype.ID
200 200 1200 G4.Genotype.ID
8000 8000 48000 G5.Genotype.ID
```

```
In [29]: ;wc G0.noGenotype.ID;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype
7800 7800 46800 G0.noGenotype.ID
7800 7800 46800 G1.noGenotype.ID
7800 7800 46800 G2.noGenotype.ID
7800 7800 46800 G3.noGenotype.ID
7800 7800 46800 G4.noGenotype.ID
0 0 0 G5.noGenotype.ID
```

```
In [30]: ped,A_Mats,numSSBayes = calc_Ai("PedAll.txt","genotype.ID",calculateInbreeding
nothing
df      = read_genotypes("GenNF.txt",numSSBayes)
M_Mats = make_MMats(df,A_Mats,ped,center=true);           # with
y_Vecs = make_yVecs("sim.phenotype",ped,numSSBayes);
Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
X_Mats, W_Mats = make_XWMats(Z_Mats,M_Mats,numSSBayes)     # no
nothing
```

```
In [31]: vRes      = 0.323
vG        = 0.323
nIter     = 50000
@time aHat1,alphaHat,betaHat,epsHat =
ssGibbs(M_Mats,y_Vecs,Z_Mats,X_Mats,W_Mats,A_Mats, numSSBayes,vRes,vG,nIter,
nothing

This is iteration 5000
This is iteration 10000
This is iteration 15000
This is iteration 20000
This is iteration 25000
This is iteration 30000
This is iteration 35000
This is iteration 40000
This is iteration 45000
This is iteration 50000
2564.464527 seconds (23.04 G allocations: 723.563 GB, 7.46% gc time)
```

```
In [32]: betaHat
```

```
Out[32]: 1-element Array{Float64,1}:
0.832281
```

```
In [33]: using DataFrames
```

```
In [34]: df = readtable("sim.bv", eltypes=[UTF8String, Float64], separator = ' ',head
a = Array{Float64,numSSBayes.num_ped}
for (i,ID) in enumerate(df[:,1])
    j = ped.idMap[ID].seqID
    a[j] = df[i,2]
end
```

```
In [35]: IDs = readtable("all.ID", eltypes =[UTF8String], separator = ' ',header=false)
posAi = getPos(ped,IDs)
cor1 = cor(a[posAi],aHat1[posAi])[1,1]
reg1 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - all.ID : correlation = %6.3f\n", cor1 ) # with
@printf("SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = %6.3f\n",
JCA11 = cor1
```

```
SSBRJC from Gibbs - all.ID : correlation = 0.912
SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.964
```

```
Out[35]: 0.9116969698292369
```

```
In [36]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[36]: 0.10575056910281516
```

```
In [37]: IDs = readtable("genotype.ID", eltypes =[UTF8String], separator = ' ',header=
posAi = getPos(ped,IDs)
cor2 = cor(a[posAi],aHat1[posAi])[1,1]
reg2 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - genotype.ID : correlation = %6.3f\n", cor2 ) #
@printf("SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = %6.3f\n",
JCA11 = cor2
```

```
SSBRJC from Gibbs - genotype.ID : correlation = 0.983
SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.119
```

```
Out[37]: 0.9834293595831247
```

```
In [38]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[38]: 1.0467255816367231
```

```
In [39]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',header=
posAi = getPos(ped,IDs)
cor3 = cor(a[posAi],aHat1[posAi])[1,1]
reg3 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - noGenotype.ID : correlation = %6.3f\n", cor3 ) ;
@printf("SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = %6.3f\n",
JCA11 = cor3
```

```
SSBRJC from Gibbs - noGenotype.ID : correlation = 0.863
SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.927
```

```
Out[39]: 0.8633448807031896
```

```
In [40]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[40]: -0.11139751071270192
```

```
In [41]: IDs = readtable("G0.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor4 = cor(a[posAi],aHat1[posAi])[1,1]
reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G0.ID : correlation = %6.3f\n", cor4 ) # with e
@printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
JCall = cor4
```

```
SSBRJC from Gibbs - G0.ID : correlation = 0.587
SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = 0.610
```

```
Out[41]: 0.586767462886083
```

```
In [42]: GEBV = aHat1[posAi]
G1GEBV=mean(GEBV)
```

```
Out[42]: -0.9118273262620235
```

```
In [43]: IDs = readtable("G1.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor4 = cor(a[posAi],aHat1[posAi])[1,1]
reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G1.ID : correlation = %6.3f\n", cor4 ) # with e
@printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", re
JCall = cor4
```

```
SSBRJC from Gibbs - G1.ID : correlation = 0.771
SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = 0.956
```

```
Out[43]: 0.7708179104196572
```

```
In [44]: GEBV = aHat1[posAi]
G1GEBV=mean(GEBV)
```

```
Out[44]: -0.5009643267359712
```

```
In [45]: IDs = readtable("G2.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor5 = cor(a[posAi],aHat1[posAi])[1,1]
reg5 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G2.ID : correlation = %6.3f\n", cor5 ) # with e
@printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
JCall = cor5
```

```
SSBRJC from Gibbs - G2.ID : correlation = 0.774
SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = 1.024
```

```
Out[45]: 0.7739109558464314
```

```
In [46]: GEBV = aHat1[posAi]
G2GEBV=mean(GEBV)
```

```
Out[46]: -0.05810484714094824
```

```
In [47]: IDs = readtable("G3.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor6 = cor(a[posAi],aHat1[posAi])[1,1]
reg6 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G3.ID : correlation = %6.3f\n", cor6 ) # with e
@printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", r
JCA11 = cor6
```

```
SSBRJC from Gibbs - G3.ID : correlation = 0.788
SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = 1.057
```

```
Out[47]: 0.7880743871482143
```

```
In [48]: GEBV = aHat1[posAi]
G3GEBV=mean(GEBV)
```

```
Out[48]: 0.31931994858708357
```

```
In [49]: IDs = readtable("G4.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor7 = cor(a[posAi],aHat1[posAi])[1,1]
reg7 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G4.ID : correlation = %6.3f\n", cor7 ) # with e
@printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", r
JCA11 = cor7
```

```
SSBRJC from Gibbs - G4.ID : correlation = 0.798
SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 1.077
```

```
Out[49]: 0.7982198624890112
```

```
In [50]: GEBV = aHat1[posAi]
G4GEBV=mean(GEBV)
```

```
Out[50]: 0.6986825021828805
```

```
In [51]: IDs = readtable("G5.ID", eltypes=[UTF8String], separator = ' ',header=false
posAi = getPos(ped,IDs)
cor8 = cor(a[posAi],aHat1[posAi])[1,1]
reg8 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G5.ID : correlation = %6.3f\n", cor8 ) # with e
@printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", r
JCA11 = cor8
```

```
SSBRJC from Gibbs - G5.ID : correlation = 0.981
SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 1.114
```

```
Out[51]: 0.980726424790754
```

```
In [52]: GEBV = aHat1[posAi]
G5GEBV=mean(GEBV)
```

```
Out[52]: 1.0873974639858708
```

```
In [53]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor9 = cor(a[posAi],aHat1[posAi])[1,1]
reg9 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G0.Genotype.ID : correlation = %6.3f\n", cor9 )
@printf("SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg9 )
JCA11 = cor9
```

```
SSBRJC from Gibbs - G0.Genotype.ID : correlation = 0.982
SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.120
```

```
Out[53]: 0.9819741242258183
```

```
In [54]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[54]: -0.06699831064904127
```

```
In [55]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor9 = cor(a[posAi],aHat1[posAi])[1,1]
reg9 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G1.Genotype.ID : correlation = %6.3f\n", cor9 )
@printf("SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg9 )
JCA11 = cor9
```

```
SSBRJC from Gibbs - G1.Genotype.ID : correlation = 0.971
SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.101
```

```
Out[55]: 0.970683068341457
```

```
In [56]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[56]: 0.4229097327585677
```

```
In [57]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor10 = cor(a[posAi],aHat1[posAi])[1,1]
reg10 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G2.Genotype.ID : correlation = %6.3f\n", cor10 )
@printf("SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg10 )
JCA11 = cor10
```

```
SSBRJC from Gibbs - G2.Genotype.ID : correlation = 0.979
SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.105
```

```
Out[57]: 0.9794727776875481
```

```
In [58]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[58]: 0.7203471999331709
```

```
In [59]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor11 = cor(a[posAi],aHat1[posAi])[1,1]
reg11 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G3.Genotype.ID : correlation = %6.3f\n", cor11)
@printf("SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg11)
JCA11 = cor11
```

```
SSBRJC from Gibbs - G3.Genotype.ID : correlation = 0.971
SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 1.097
```

```
Out[59]: 0.9714647396857966
```

```
In [60]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[60]: 1.0827422035954295
```

```
In [61]: IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor12 = cor(a[posAi],aHat1[posAi])[1,1]
reg12 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G4.Genotype.ID : correlation = %6.3f\n", cor12)
@printf("SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg12)
JCA12 = cor12
```

```
SSBRJC from Gibbs - G4.Genotype.ID : correlation = 0.971
SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.118
```

```
Out[61]: 0.9705308765797513
```

```
In [62]: GEBV = aHat1[posAi]
mean(GEBV)
```

```
Out[62]: 1.4477517885795785
```

```
In [63]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
posAi = getPos(ped,IDs)
cor13 = cor(a[posAi],aHat1[posAi])[1,1]
reg13 = linreg(aHat1[posAi], a[posAi])[2,1]
@printf("SSBRJC from Gibbs - G5.Genotype.ID : correlation = %6.3f\n", cor13)
@printf("SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg13)
JCA13 = cor13
```

```
SSBRJC from Gibbs - G5.Genotype.ID : correlation = 0.981
SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 1.114
```

```
Out[63]: 0.980726424790754
```

```
In [64]: writedlm("Correlation.G5.G.C.txt",cor13)
```

```
In [65]: writedlm("Regression.G5.G.C.txt",reg13)
```



```
In [66]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
```

```
Out[66]: 1.0873974639858708
```

```
In [67]: IDs = readtable("G0.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat1[posAi])[1,1]
         reg14 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : correlation = %6.3f\n", cor14)
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg14)
         JCall = cor14
```

```
SSBRJC from Gibbs - G0.noGenotype.ID : correlation = 0.563
SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 0.595
```

```
Out[67]: 0.5632923875996132
```

```
In [68]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[68]: -0.9334896087136384
```

```
In [69]: IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat1[posAi])[1,1]
         reg14 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G1.noGenotype.ID : correlation = %6.3f\n", cor14)
         @printf("SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg14)
         JCall = cor14
```

```
SSBRJC from Gibbs - G1.noGenotype.ID : correlation = 0.750
SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 0.955
```

```
Out[69]: 0.7502491776155054
```

```
In [70]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[70]: -0.5246534051845493
```

```
In [71]: IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor15 = cor(a[posAi],aHat1[posAi])[1,1]
         reg15 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : correlation = %6.3f\n", cor15)
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg15)
         JCall = cor15
```

```
SSBRJC from Gibbs - G2.noGenotype.ID : correlation = 0.756
SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = 1.026
```

```
Out[71]: 0.7561190221703962
```

```
In [72]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[72]: -0.07806515604028463
```

```
In [73]: IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor16 = cor(a[posAi],aHat1[posAi])[1,1]
         reg16 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G3.noGenotype.ID : correlation = %6.3f\n", cor16)
         @printf("SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg16)
         JCall = cor16
```

```
SSBRJC from Gibbs - G3.noGenotype.ID : correlation = 0.771
SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 1.057
```

```
Out[73]: 0.7714692426610509
```

```
In [74]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[74]: 0.2997450189714849
```

```
In [75]: IDs = readtable("G4.noGenotype.ID", eltypes =[UTF8String], separator = ' ', header = 1,
         posAi = getPos(ped,IDs)
         cor17 = cor(a[posAi],aHat1[posAi])[1,1]
         reg17 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G4.noGenotype.ID : correlation = %6.3f\n", cor17)
         @printf("SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg17)
         JCall = cor17
```

```
SSBRJC from Gibbs - G4.noGenotype.ID : correlation = 0.782
SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 1.075
```

```
Out[75]: 0.7815487054389542
```

```
In [76]: GEBV = aHat1[posAi]
         mean(GEBV)
```

```
Out[76]: 0.6794755974034781
```

```
In [77]: numSSBayes
```

```
Out[77]: SSBR.NumSSBayes(54891,45891,9000,40000,39000,1000,200)
```