```
In [2]: include("/home/nicole/Jupyter/SSBRnoJ/src/SSBR.jl")
         using SSBR
 In [3]: function getPos(ped,IDs)
             posAi = Array(Int64, size(IDs, 1))
             for (i,id) = enumerate(IDs[:,1])
                 posAi[i] = ped.idMap[id].seqID
             end
             return posAi
         end
Out[3]: getPos (generic function with 1 method)
 In [4]: | ; cd Data/0.5/G/1
         /home/nicole/Jupyter/JG3/Data/0.5/G/1
 In [5]: ;ls
         Correlation.G5.G.JC.txt
         Correlation.G5.G.N.txt
         Correlation.G5.G.PBLUP.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
In [6]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
         ;awk '{print $1}' GenNF.txt | sort -b > genotype.ID
In [7]:
In [8]:
         ; join -v1 all.ID genotype.ID > noGenotype.ID
In [9]:
         ;awk '{print $1,$2}' Phe.txt > sim.phenotype
In [10]:
         ;awk '{print $1,$3}' PheAll.txt > sim.bv
         ; awk 'NR >=1 && NR <=8000 {print $1}' PedAll.txt | sort -b > G0.ID
In [11]:
```

```
In [12]: ; awk 'NR >=8001 && NR <=16000 {print $1}' PedAll.txt | sort -b > G1.ID
In [13]: ; awk 'NR >=16001 && NR <=24000 {print $1}' PedAll.txt | sort -b > G2.ID
         ; 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
         ; awk 'NR >=40001 && NR <=48000 {print $1}' PedAll.txt | sort -b > G5.ID
In [16]:
In [17]:
         ; join G0.ID genotype.ID > G0.Genotype.ID
         ; join G1.ID genotype.ID > G1.Genotype.ID
In [18]:
In [19]:
         ; join G2.ID genotype.ID > G2.Genotype.ID
In [20]:
         ; join G3.ID genotype.ID > G3.Genotype.ID
In [21]:
         ; join G4.ID genotype.ID > G4.Genotype.ID
In [22]:
         ; join G5.ID genotype.ID > G5.Genotype.ID
         ;join -v1 G0.ID genotype.ID > G0.noGenotype.ID
In [23]:
In [24]:
         ; join -v1 G1.ID genotype.ID > G1.noGenotype.ID
In [25]:
         ;join -v1 G2.ID genotype.ID > G2.noGenotype.ID
In [26]:
         ; join -v1 G3.ID genotype.ID > G3.noGenotype.ID
In [27]:
         ;join -v1 G4.ID genotype.ID > G4.noGenotype.ID
In [28]:
         ; join -v1 G5.ID genotype.ID > G5.noGenotype.ID
In [29]:
         ;wc G0.Genotype.ID;wc G1.Genotype.ID;wc G2.Genotype.ID;wc G3.Genotype.ID;wc (
               200 1200 GO.Genotype.ID
               200 1200 G1.Genotype.ID
          200
               200 1200 G2.Genotype.ID
               200 1200 G3.Genotype.ID
          200
          200
               200 1200 G4.Genotype.ID
          8000 8000 48000 G5.Genotype.ID
```

```
In [30]:
         ;wc G0.noGenotype.ID;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype
                7800 46800 G0.noGenotype.ID
          7800
                7800 46800 Gl.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 [31]:
         ped,A Mats,numSSBayes = calc Ai("PedAll.txt","genotype.ID",calculateInbreedia
         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
                = 0.668
In [32]: vRes
                = 0.668
         vG
         nIter = 50000
         @time aHat1,alphaHat,betaHat,epsiHat =
         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
         2556.715886 seconds (23.05 G allocations: 723.770 GB, 7.43% gc time)
In [ ]:
In [33]: | betaHat
Out[33]: 1-element Array{Float64,1}:
          11.2478
In [34]: using DataFrames
In [35]: 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 [36]: 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",
         JCAll = cor1
         SSBRJC from Gibbs - all.ID : correlation = 0.913
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.969
Out[36]: 0.9134593992136977
In [37]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[37]: 0.14383435674344183
In [38]: 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 = %6.3:
         JCAll = cor2
         SSBRJC from Gibbs - genotype.ID : correlation = 0.968
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.081
Out[38]: 0.9684372886055078
In [39]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[39]: 1.3003410754939129
In [40]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',heade
         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
         JCAll = cor3
         SSBRJC from Gibbs - noGenotype.ID : correlation = 0.876
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.952
Out[40]: 0.8760824550653251
In [41]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[41]: -0.12305180912205142
```

```
In [42]: 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 ei
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.709
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.852
Out[42]: 0.7091224741906464
In [43]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[43]: -1.2209387801401774
In [44]: 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 = 0.778
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.989
Out[44]: 0.7783040396605218
In [45]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[45]: -0.5909427228512527
In [46]: 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 ep
         @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor5
         SSBRJC from Gibbs - G2.ID : correlation = 0.769
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 0.990
Out[46]: 0.7685711190277406
In [47]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[47]: -0.04148152955399474
```

```
In [48]: 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 ei
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor6
         SSBRJC from Gibbs - G3.ID : correlation = 0.768
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 0.996
Out[48]: 0.7682499921914511
In [49]: GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[49]: 0.4603284454440368
In [50]: 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", re
         JCAll = cor7
         SSBRJC from Gibbs - G4.ID : correlation = 0.787
         SSBRJC from Gibbs - G4.ID: regression of TBV on GEBV = 1.008
Out[50]: 0.7865395878502194
In [51]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[51]: 0.9123590323739923
In [52]: 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 ep
         @printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor8
         SSBRJC from Gibbs - G5.ID : correlation = 0.964
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 1.076
Out[52]: 0.9642432087368493
In [53]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[53]: 1.3436816951880466
```

```
In [54]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         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
         JCAll = cor9
         SSBRJC from Gibbs - G0.Genotype.ID : correlation = 0.962
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.090
Out[54]: 0.9620764789616452
In [55]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[55]: 0.08377111241222515
In [56]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ', head
         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
         JCAll = cor9
         SSBRJC from Gibbs - G1.Genotype.ID : correlation = 0.968
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.039
Out[56]: 0.9680278017929095
In [57]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[57]: 0.5626403688933109
In [58]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ', head
         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
         JCAll = cor10
         SSBRJC from Gibbs - G2.Genotype.ID : correlation =
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.042
Out[58]: 0.9575604672093873
In [59]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[59]: 1.0135147760412275
```

```
In [60]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         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
         JCAll = cor11
         SSBRJC from Gibbs - G3.Genotype.ID : correlation = 0.966
         SSBRJC from Gibbs - G3.Genotype.ID: regression of TBV on GEBV = 1.056
Out[60]: 0.9657278358373766
In [61]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[61]: 1.3754584112463057
In [62]: IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         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
         JCAll = cor12
         SSBRJC from Gibbs - G4.Genotype.ID : correlation = 0.945
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.045
Out[62]: 0.9445836318545584
In [63]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[63]: 1.732695921111143
In [64]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ', head
         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
         JCAll = cor13
         SSBRJC from Gibbs - G5.Genotype.ID : correlation =
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 1.076
Out[64]: 0.9642432087368493
In [65]: | writedlm("Correlation.G5.G.C.txt",cor13)
In [66]: writedlm("Regression.G5.G.C.txt",reg13)
```

```
In [67]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
Out[67]: 1.3436816951880466
In [68]: IDs = readtable("G0.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         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", corl
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor14
         SSBRJC from Gibbs - G0.noGenotype.ID : correlation = 0.690
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 0.855
Out[68]: 0.6899870737584175
In [69]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[69]: -1.2543928799492132
In [70]: IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         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", corl
         @printf("SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor14
         SSBRJC from Gibbs - G1.noGenotype.ID : correlation = 0.761
         SSBRJC from Gibbs - G1.noGenotype.ID: regression of TBV on GEBV = 0.993
Out[70]: 0.7612832235273792
In [71]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[71]: -0.6205217764857286
In [72]: IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         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", cor1!
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor15
         SSBRJC from Gibbs - G2.noGenotype.ID : correlation = 0.751
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 0.990
Out[72]: 0.7508661790101163
```

```
In [73]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[73]: -0.06853271687694912
In [74]: IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         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", cor1
         @printf("SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor16
         SSBRJC from Gibbs - G3.noGenotype.ID : correlation = 0.752
         SSBRJC from Gibbs - G3.noGenotype.ID: regression of TBV on GEBV = 0.992
Out[74]: 0.7519443835163266
In [75]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[75]: 0.4368635745260299
In [76]: IDs = readtable("G4.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         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", cor1
         @printf("SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor17
         SSBRJC from Gibbs - G4.noGenotype.ID : correlation = 0.775
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 1.005
Out[76]: 0.7748802485852807
In [77]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[77]: 0.8913247531756039
In [78]: numSSBayes
Out[78]: SSBR.NumSSBayes(54906,45906,9000,40000,39000,1000,200)
```