```
In [1]: include("/home/nicole/Jupyter/SSBRnoJ/src/SSBR.jl")
         using SSBR
In [2]: 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[2]: getPos (generic function with 1 method)
In [3]: | ; cd Data/0.5a0/G/8
         /home/nicole/Jupyter/JG3/Data/0.5a0/G/8
In [4]:
         ;ls
         GEIINT . LAL
         PedAll.txt
         Phe.txt
         PheAll.txt
         Regression.G5.G.J.txt
         Regression.G5.G.JC.txt
         Regression.G5.G.PBLUP.txt
         all.ID
         alphaEstimates
         genotype.ID
         meanOfSNPGAll
         meanOfSNPGG0
         meanOfSNPGG1
         meanOfSNPGG2
         meanOfSNPGG3
         meanOfSNPGG4
         meanOfSNPGG5
         noGenotype.ID
         sim.bv
         sim.phenotype
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
         ;awk '{print $1,$2}' Phe.txt > sim.phenotype
In [8]:
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
         ; awk 'NR >=24001 && NR <=32000 {print $1}' PedAll.txt | sort -b > G3.ID
In [13]:
         ; 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 [15]:
In [16]:
         ; join G0.ID genotype.ID > G0.Genotype.ID
         ; join G1.ID genotype.ID > G1.Genotype.ID
In [17]:
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
         ;join -v1 G0.ID genotype.ID > G0.noGenotype.ID
In [22]:
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 (
               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 [29]:
         ;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 [30]:
         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
                = 0.741
In [31]: vRes
                = 0.741
         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
         2507.247410 seconds (23.06 G allocations: 723.991 GB, 7.71% gc time)
In [32]: betaHat
Out[32]: 1-element Array{Float64,1}:
          4.04464
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",
         JCAll = cor1
         SSBRJC from Gibbs - all.ID : correlation = 0.920
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.970
Out[35]: 0.920412614858272
In [36]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[36]: 0.16088101392236945
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 = %6.3:
         JCAll = cor2
         SSBRJC from Gibbs - genotype.ID : correlation = 0.983
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.048
Out[37]: 0.9829657287414549
In [38]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[38]: 1.435761799317526
In [39]: 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 = 0.885
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.963
Out[39]: 0.8846041567377392
In [40]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[40]: -0.13332224424574357
```

```
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 ei
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.715
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.887
Out[41]: 0.7152119574321202
In [42]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[42]: -1.3341171975784853
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 = 0.782
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.989
Out[43]: 0.7819902876570167
In [44]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[44]: -0.6538016555711365
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 ep
         @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor5
         SSBRJC from Gibbs - G2.ID : correlation = 0.785
         SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = 1.024
Out[45]: 0.785365863206522
In [46]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[46]: -0.05341644145965353
```

```
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 ei
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor6
         SSBRJC from Gibbs - G3.ID : correlation = 0.765
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 0.999
Out[47]: 0.7654499096010323
In [48]: GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[48]: 0.5109822924009058
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", re
         JCAll = cor7
         SSBRJC from Gibbs - G4.ID : correlation = 0.779
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 0.993
Out[49]: 0.778816689522527
In [50]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[50]: 1.009874803049669
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 ep
         @printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor8
         SSBRJC from Gibbs - G5.ID : correlation = 0.981
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 1.046
Out[51]: 0.9805888425781926
In [52]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[52]: 1.4857642826929174
```

```
In [53]: 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 = 0.980
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.093
Out[53]: 0.9799959609096395
In [54]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[54]: 0.05807750389209795
In [55]: 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 = 0.981
         SSBRJC from Gibbs - G1.Genotype.ID: regression of TBV on GEBV = 1.058
Out[55]: 0.980761250254465
In [56]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[56]: 0.5791174925627445
In [57]: 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 = 1.037
Out[57]: 0.9752572472687668
In [58]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[58]: 1.0954269707997402
```

```
In [59]: 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 = 0.975
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 1.019
Out[59]: 0.9745093096413263
In [60]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[60]: 1.5188489347649523
In [61]: 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 = 0.970
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.028
Out[61]: 0.9701779214702356
In [62]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[62]: 1.9272387595524447
In [63]: 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 = 1.046
Out[63]: 0.9805888425781926
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]: 1.4857642826929174
In [67]: 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 = 0.693
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 0.889
Out[67]: 0.6933583675289976
In [68]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[68]: -1.3698144976161926
In [69]: 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 = 0.763
         SSBRJC from Gibbs - G1.noGenotype.ID: regression of TBV on GEBV = 0.992
Out[69]: 0.7634341479921789
In [70]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[70]: -0.6854149670617489
In [71]: 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 = 0.770
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 1.030
Out[71]: 0.7699220278039617
```

```
In [72]: | GEBV = aHat1[posAi]
         mean(GEBV)
Out[72]: -0.08287396485092002
In [73]: 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 = 0.750
         SSBRJC from Gibbs - G3.noGenotype.ID: regression of TBV on GEBV = 1.000
Out[73]: 0.7496198618191624
In [74]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[74]: 0.4851395579813149
In [75]: 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 = 0.765
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 0.993
Out[75]: 0.7646473887429909
In [76]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[76]: 0.9863526503188287
In [77]: numSSBayes
Out[77]: SSBR.NumSSBayes(54922,45922,9000,40000,39000,1000,200)
```