```
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/0.3/G/2
        /home/nicole/Jupyter/JG3/Data/0.3/G/2
In [4]:
        ;ls
        Correlation.G5.G.J.txt
        Correlation.G5.G.JC.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 [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 GO.ID genotype.ID > GO.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
         ; join -v1 G3.ID genotype.ID > G3.noGenotype.ID
In [25]:
         ; join -v1 G4.ID genotype.ID > G4.noGenotype.ID
In [26]:
         ;join -v1 G5.ID genotype.ID > G5.noGenotype.ID
In [27]:
         ;wc G0.Genotype.ID;wc G1.Genotype.ID;wc G2.Genotype.ID;wc G3.Genotype.ID;wc G4.Genotype.ID;wc G5.Genotype
In [28]:
          200 200 1200 GO.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
         ;wc G0.noGenotype.ID;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype.ID;wc G4.noGenotype.ID;wc G
In [29]:
          7800 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", calculateInbreeding=false)
         nothing
         df
                = read genotypes("GenNF.txt", numSSBayes)
                                                                                 # with centering
         M Mats = make MMats(df, A Mats, ped, center=true);
         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 J
         nothing
```

```
In [31]:
         vG
                = 0.731
                = 1.706
         vRes
         nIter = 50000
         @time aHat1,alphaHat,betaHat,epsiHat =
         ssGibbs(M Mats,y Vecs, Z Mats, X Mats, W Mats, A Mats, numSSBayes, vRes, vG, nIter, outFreq=5000);
         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
         2479.954453 seconds (23.06 G allocations: 724.155 GB, 7.43% gc time)
In [32]: betaHat
Out[32]: 1-element Array{Float64,1}:
          12.2173
In [33]: using DataFrames
         df = readtable("sim.bv", eltypes =[UTF8String, Float64], separator = ' ',header=false)
In [34]:
         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 epsilon
         @printf("SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = %6.3f\n", reg1)
         JCAll = cor1
         SSBRJC from Gibbs - all.ID : correlation = 0.857
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.989
Out[35]: 0.8568567436843337
In [36]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[36]: 0.13224472801675563
In [37]: IDs = readtable("genotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @printf("SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = %6.3f\n", reg2)
         JCAll = cor2
         SSBRJC from Gibbs - genotype.ID : correlation = 0.938
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.117
Out[37]: 0.9379404473458274
In [38]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[38]: 1.0281220851792365
```

```
In [39]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @printf("SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = %6.3f\n", reg3)
         JCAll = cor3
         SSBRJC from Gibbs - noGenotype.ID : correlation = 0.808
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.969
Out[39]: 0.8083279429324747
In [40]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[40]: -0.07449620055920134
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 epsilon
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", req4)
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.690
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 1.148
Out[41]: 0.6896548692711091
In [42]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[42]: -1.0780967809541548
```

```
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 epsilon
         @printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", req4)
         JCAll = cor4
         SSBRJC from Gibbs - G1.ID : correlation = 0.685
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 1.007
Out[43]: 0.6850293930851707
In [44]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[44]: -0.3926723173021774
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 epsilon
         @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", reg5)
         JCAll = cor5
         SSBRJC from Gibbs - G2.ID : correlation = 0.676
         SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = 0.992
Out[45]: 0.6761584978914771
In [46]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[46]: 0.0577406766323755
```

```
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 epsilon
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", reg6)
         JCAll = cor6
         SSBRJC from Gibbs - G3.ID : correlation = 0.666
         SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = 0.993
Out[47]: 0.6655591835880451
In [48]: GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[48]: 0.4184059604918812
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 epsilon
         @printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", req7)
         JCAll = cor7
         SSBRJC from Gibbs - G4.ID : correlation = 0.695
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 1.015
Out[49]: 0.6948461573841903
In [50]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[50]: 0.7327401884851374
```

```
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 epsilon
         @printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", reg8)
         JCAll = cor8
         SSBRJC from Gibbs - G5.ID : correlation = 0.934
         SSBRJC from Gibbs - G5.ID: regression of TBV on GEBV = 1.115
Out[51]: 0.9338038731346935
In [52]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[52]: 1.0553506407474724
In [53]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @printf("SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg9)
         JCAll = cor9
         SSBRJC from Gibbs - G0.Genotype.ID : correlation = 0.940
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.047
Out[53]: 0.9400389009510088
In [54]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[54]: 0.3188080607127937
```

```
In [55]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @printf("SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg9)
         JCAll = cor9
         SSBRJC from Gibbs - G1.Genotype.ID : correlation = 0.936
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.080
Out[55]: 0.9361629048435645
In [56]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[56]: 0.5428044712134167
In [57]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @printf("SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = %6.3f\n", req10)
         JCAll = cor10
         SSBRJC from Gibbs - G2.Genotype.ID : correlation = 0.938
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.106
Out[57]: 0.93847550301213
In [58]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[58]: 0.8078041946110062
```

```
In [59]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @printf("SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = %6.3f\n", req11)
         JCAll = cor11
         SSBRJC from Gibbs - G3.Genotype.ID : correlation = 0.933
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 1.087
Out[59]: 0.9330499740929016
In [60]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[60]: 1.0416498715354607
In [61]: IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @printf("SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = %6.3f\n", reg12)
         JCAll = cor12
         SSBRJC from Gibbs - G4.Genotype.ID : correlation = 0.926
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.118
Out[61]: 0.9263542844376513
In [62]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[62]: 1.3404016050940757
```

```
In [63]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @printf("SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = %6.3f\n", req13)
         JCAll = cor13
         SSBRJC from Gibbs - G5.Genotype.ID : correlation = 0.934
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 1.115
Out[63]: 0.9338038731346935
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]: 1.0553506407474724
In [67]: IDs = readtable("G0.noGenotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @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 = 0.672
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 1.227
Out[67]: 0.6722861347366013
In [68]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[68]: -1.1139148538174102
```

```
In [69]: IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @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 = 0.668
         SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 1.016
Out[69]: 0.6684062898059506
In [70]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[70]: -0.41665890162309005
In [71]: IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @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 = 0.659
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 0.988
Out[71]: 0.6594354149627717
In [72]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[72]: 0.038508278735487546
```

```
In [73]: IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @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 = 0.651
         SSBRJC from Gibbs - G3.noGenotype.ID: regression of TBV on GEBV = 0.987
Out[73]: 0.6505018803071381
In [74]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[74]: 0.40242534738819963
In [75]: IDs = readtable("G4.noGenotype.ID", eltypes =[UTF8String], separator = ' ',header=false)
         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 ) # with epsilon
         @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 = 0.681
         SSBRJC from Gibbs - G4.noGenotype.ID: regression of TBV on GEBV = 1.007
Out[75]: 0.6812382556291628
In [76]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[76]: 0.7171591265208057
In [77]: numSSBayes
Out[77]: SSBR.NumSSBayes(54934,45934,9000,40000,39000,1000,200)
```