```
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/7
        /home/nicole/Jupyter/JG3/Data/0.3/G/7
        ;ls
In [4]:
        سد، دی
        G5.noGenotype.ID
        GenNF.txt
        PedAll.txt
        Phe.txt
        PheAll.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
 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
In [29]: ;wc G0.noGenotype.ID;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype.ID;wc G4.noGenotype.ID;wc G
          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
                = 0.647
                = 1.510
         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
         5248.121520 seconds (23.07 G allocations: 724.266 GB, 6.55% gc time)
In [32]: betaHat
Out[32]: 1-element Array{Float64,1}:
          12.0993
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 = 0.867
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.983
Out[35]: 0.8666973592046295
In [36]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[36]: 0.13502370038814945
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 = 0.951
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.077
Out[37]: 0.9506296207185001
In [38]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[38]: 1.1003586439628106
```

```
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 = 0.815
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.972
Out[39]: 0.8148837792477959
In [40]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[40]: -0.08774590197523391
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 = 0.677
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 1.097
Out[41]: 0.6765508945057549
In [42]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[42]: -1.0742792288210725
```

```
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 = 0.689
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 1.031
Out[43]: 0.6894978673017784
In [44]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[44]: -0.448093680670392
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 = 0.679
         SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = 1.021
Out[45]: 0.6785336408850646
In [46]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[46]: 0.004858721297516091
```

```
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 = 0.672
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 1.008
Out[47]: 0.6719292550666583
In [48]: | GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[48]: 0.4084561085506018
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 = 0.713
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 1.043
Out[49]: 0.713110581086378
In [50]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[50]: 0.7837988071747191
```

```
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 = 0.947
         SSBRJC from Gibbs - G5.ID: regression of TBV on GEBV = 1.075
Out[51]: 0.9469930984799638
In [52]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[52]: 1.1354014747975243
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 = 0.947
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.057
Out[53]: 0.946617743820394
In [54]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[54]: 0.2058692925878045
```

```
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 = 0.944
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.020
Out[55]: 0.9437507857281404
In [56]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[56]: 0.47368436185251817
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 = 0.954
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.057
Out[57]: 0.9536363895869129
In [58]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[58]: 0.8114293575182088
```

```
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 = 0.944
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 1.070
Out[59]: 0.9442552800594832
In [60]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[60]: 1.1539735567103606
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 = 0.929
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.037
Out[61]: 0.9290824343921623
In [62]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[62]: 1.455123417756605
```

```
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 = 0.947
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 1.075
Out[63]: 0.9469930984799638
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]: 1.1354014747975243
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 = 0.659
         SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 1.164
Out[67]: 0.6586817096404074
In [68]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[68]: -1.1071035498828388
```

```
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 = 0.672
         SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 1.041
Out[69]: 0.6724869483406303
In [70]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[70]: -0.47172901509405646
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 = 0.660
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 1.021
Out[71]: 0.6599542202750854
In [72]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[72]: -0.01582257706711703
```

```
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 = 0.654
         SSBRJC from Gibbs - G3.noGenotype.ID: regression of TBV on GEBV = 1.004
Out[73]: 0.6538246248682226
In [74]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[74]: 0.38934027654650544
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 = 0.699
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 1.042
Out[75]: 0.6991862372584179
In [76]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[76]: 0.7665853556213372
In [77]: numSSBayes
Out[77]: SSBR.NumSSBayes(54942,45942,9000,40000,39000,1000,200)
```