/home/nicole/Jupyter/JG3/Data/0.1/G/7

```
;ls
In [4]:
        Correlation.G5.G.C.txt
        Correlation.G5.G.J.txt
        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
        PedAll.txt
        Phe.txt
        PheAll.txt
        Regression.G5.G.C.txt
        Regression.G5.G.J.txt
        Regression.G5.G.JC.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
         ; awk 'NR >=16001 && NR <=24000 {print $1}' PedAll.txt | sort -b > G2.ID
In [12]:
         ; awk 'NR >=24001 && NR <=32000 {print $1}' PedAll.txt | sort -b > G3.ID
         ; 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
         ; join G0.ID genotype.ID > G0.Genotype.ID
In [16]:
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
         ; join G5.ID genotype.ID > G5.Genotype.ID
In [21]:
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 [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
               200 1200 GO.Genotype.ID
          200
               200 1200 G1.Genotype.ID
               200 1200 G2.Genotype.ID
          200
              200 1200 G3.Genotype.ID
               200 1200 G4.Genotype.ID
          200
          8000 8000 48000 G5.Genotype.ID
```

```
;wc G0.noGenotype.ID;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype
In [29]:
                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);
                                                                                  # 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
                = 0.823
                = 7.410
         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
         2449.979692 seconds (23.03 G allocations: 723.384 GB, 7.51% gc time)
In [32]: betaHat
Out[32]: 1-element Array{Float64,1}:
          9.13963
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 = 0.530
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.252
Out[35]: 0.5298721309816339
In [36]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[36]: 1.503682852355321
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 = 0.869
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 0.818
Out[37]: 0.8685880142723112
In [38]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[38]: 3.50799998661877
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 = 0.427
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.202
Out[39]: 0.42706964172118955
In [40]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[40]: 1.0411481290637559
```

```
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 = 0.691
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.514
Out[41]: 0.691412275871113
In [42]: GEBV = aHat1[posAi]
         G0GEBV=mean(GEBV)
Out[42]: 0.2056924964526647
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 = 0.333
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.129
Out[43]: 0.33315374715641555
In [44]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[44]: 0.8591385769427935
In [45]: IDs = readtable("G2.ID", eltypes =[UTF8String], separator = ' ', header=false
         posAi = getPos(ped,IDs)
         cor5 = cor(a[posAi],aHat1[posAi])[1,1]
         req5 = 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 = 0.353
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 0.139
Out[45]: 0.35315425675299345
In [46]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[46]: 1.1900076039780079
```

```
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 = 0.329
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 0.130
Out[47]: 0.32923081850067343
In [48]: GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[48]: 1.4322987298498011
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 = 0.329
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 0.133
Out[49]: 0.3287918854547358
In [50]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[50]: 1.7948954183512096
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 = 0.867
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 0.817
Out[51]: 0.8667948970336402
In [52]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[52]: 3.5400642885574483
```

```
In [53]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat1[posAi])[1,1]
         \#TBV = a[posAi]
         \#GEBV = aHat1[posAi]
         \#reg = linreg(X,Y)
         \#reg = linreg(GEBV, TBV)[2,1]
         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 = 0.850
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 0.819
Out[53]: 0.8501721823994521
In [54]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[54]: 2.844604526721444
In [55]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat1[posAi])[1,1]
         \#TBV = a[posAi]
         #GEBV = aHat1[posAi]
         \#reg = linreg(X,Y)
         \#reg = linreg(GEBV, TBV)[2,1]
         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 =
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 0.914
Out[55]: 0.8839298021737286
In [56]: | GEBV = aHat1[posAi]
         mean(GEBV)
Out[56]: 2.9413968973785796
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 = 0.872
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 0.890
Out[57]: 0.8720492858041373
```

```
In [58]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[58]: 3.173880017313144
In [59]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         corl1 = 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 = corl1
         SSBRJC from Gibbs - G3.Genotype.ID : correlation = 0.886
         SSBRJC from Gibbs - G3.Genotype.ID: regression of TBV on GEBV = 0.904
Out[59]: 0.8857863993486114
In [60]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[60]: 3.3853589879995285
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 = 0.823
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 0.807
Out[61]: 0.8227656213314781
In [62]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[62]: 3.9121874261339973
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 = 0.867
         SSBRJC from Gibbs - G5.Genotype.ID: regression of TBV on GEBV = 0.817
Out[63]: 0.8667948970336402
In [64]: | writedlm("Correlation.G5.G.N.txt",cor13)
```

```
In [65]: writedlm("Regression.G5.G.N.txt",reg13)
In [66]: TBVG5Gall = a[posAi]
         TBVG5G=mean(TBVG5Gall)
Out[66]: 11.095323875000002
In [67]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
Out[67]: 3.5400642885574483
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 = 0.672
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 0.519
Out[68]: 0.6719948670547959
In [69]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[69]: 0.138028085420132
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 = 0.320
         SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 0.124
Out[70]: 0.31964892684019675
In [71]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[71]: 0.8057473379572606
```

```
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 = 0.342
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 0.134
Out[72]: 0.3415902925021689
In [73]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[73]: 1.1391390805591586
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 = 0.317
         SSBRJC from Gibbs - G3.noGenotype.ID: regression of TBV on GEBV = 0.125
Out[74]: 0.3172362347046961
In [75]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[75]: 1.3822202616921158
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 = 0.316
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 0.128
Out[76]: 0.3158413237426793
In [77]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[77]: 1.7406058796901125
```

In [78]: numSSBayes

Out[78]: SSBR.NumSSBayes(54878,45878,9000,40000,39000,1000,200)