/home/nicole/Jupyter/JG3/Data/0.5/Q/4

```
;ls
In [4]:
         Correlation.G5.Q.J.txt
         Correlation.G5.Q.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
         PedAll.txt
         Phe.txt
         PheAll.txt
         OTLNF.txt
         Regression.G5.Q.J.txt
         Regression.G5.Q.JC.txt
         all.ID
         alphaEstimates
         genotype.ID
         meanOfSNPQAll
         meanOfSNPQG0
         meanOfSNPQG1
         meanOfSNPQG2
         meanOfSNPQG3
         meanOfSNPQG4
         meanOfSNPQG5
         noGenotype.ID
         sim.bv
         sim.phenotype
         ;awk '{print $1}' PedAll.txt | sort -b > all.ID
In [5]:
         ;awk '{print $1}' QTLNF.txt | sort -b > genotype.ID
In [6]:
         ;join -v1 all.ID genotype.ID > noGenotype.ID
In [7]:
         ;awk '{print $1,$2}' Phe.txt > sim.phenotype
In [8]:
In [9]:
         ;awk '{print $1,$3}' PheAll.txt > sim.bv
         ; awk 'NR >=1 && NR <=8000 {print $1}' PedAll.txt | sort -b > G0.ID
In [10]:
```

```
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
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 [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
        ; join G3.ID genotype.ID > G3.Genotype.ID
In [19]:
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 G
          200
               200 1200 GO.Genotype.ID
               200 1200 Gl.Genotype.ID
          200
          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", calculateInbreedin
         nothing
         df
                = read_genotypes("QTLNF.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 J
         nothing
In [31]: | vRes
                = 0.759
                = 0.759
         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
         2984.622172 seconds (22.99 G allocations: 722.841 GB, 7.96% gc time)
In [32]: betaHat
Out[32]: 1-element Array{Float64,1}:
          8.72703
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", corl ) # with \epsilon
         @printf("SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = %6.3f\n",
         JCAll = cor1
         SSBRJC from Gibbs - all.ID : correlation = 0.919
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 1.001
Out[35]: 0.9188902115403603
In [36]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[36]: 1.728348511705809
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 ) # N
         @printf("SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = %6.3f
         JCAll = cor2
         SSBRJC from Gibbs - genotype.ID : correlation = 0.951
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.164
Out[37]: 0.9514026014497169
In [38]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[38]: 3.028397713428731
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 = 0.883
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.969
Out[39]: 0.8832388884052378
In [40]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[40]: 1.4283371574620578
```

```
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 er
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.721
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.884
Out[41]: 0.7206418018817838
In [42]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[42]: 0.16622747331397608
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 er
         @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 = 1.009
Out[43]: 0.7782165374252237
In [44]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[44]: 0.917955349867524
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 er
         @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor5
         SSBRJC from Gibbs - G2.ID : correlation = 0.775
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 1.041
Out[45]: 0.7745653593548215
In [46]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[46]: 1.5126337218532153
```

```
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 er
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor6
         SSBRJC from Gibbs - G3.ID : correlation = 0.770
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 1.033
Out[47]: 0.7699188997565225
In [48]: GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[48]: 2.075675894729011
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 er
         @printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor7
         SSBRJC from Gibbs - G4.ID : correlation = 0.769
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 1.058
Out[49]: 0.76886934908365
In [50]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[50]: 2.620559061888861
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 er
         @printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor8
         SSBRJC from Gibbs - G5.ID : correlation = 0.944
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 1.151
Out[51]: 0.943809541304906
In [52]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[52]: 3.0770395685822667
```

```
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 = 0.960
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.126
Out[53]: 0.9596999988441531
In [54]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[54]: 1.741013584425813
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]
         req9 = linreg(aHat1[posAi], a[posAi])[2,1]
         <code>@printf("SSBRJC from Gibbs - G1.Genotype.ID : correlation = 6.3f\n", cor9 )</code>
         @printf("SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = %6.
         JCAll = cor9
         SSBRJC from Gibbs - G1.Genotype.ID : correlation = 0.934
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.090
Out[55]: 0.9342057244589939
In [56]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[56]: 2.172682620972472
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 = 0.931
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.067
Out[57]: 0.9306798932164088
```

```
In [58]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[58]: 2.6671712965215524
In [59]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         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 = 0.921
         SSBRJC from Gibbs - G3.Genotype.ID: regression of TBV on GEBV = 1.054
Out[59]: 0.9205455421931081
In [60]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[60]: 3.1367292955022124
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 = 0.916
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.097
Out[61]: 0.9160278451873831
In [62]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[62]: 3.47871756358017
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 = 0.944
         SSBRJC from Gibbs - G5.Genotype.ID: regression of TBV on GEBV = 1.151
Out[63]: 0.943809541304906
In [64]: writedlm("Correlation.G5.Q.N.txt",cor13)
```

```
In [65]: writedlm("Regression.G5.Q.N.txt",reg13)
In [66]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
Out[66]: 3.0770395685822667
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", cor14
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = %
         JCAll = cor14
         SSBRJC from Gibbs - G0.noGenotype.ID : correlation = 0.701
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 0.900
Out[67]: 0.7008516442099813
In [68]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[68]: 0.12584834225982638
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", cor14
         @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 = 1.020
Out[69]: 0.7632990721662483
In [70]: | GEBV = aHat1[posAi]
         mean (GEBV)
Out[70]: 0.885782855736628
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", cor15
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = %
         JCAll = cor15
         SSBRJC from Gibbs - G2.noGenotype.ID : correlation = 0.758
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 1.043
Out[71]: 0.7575534339614336
```

```
In [72]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[72]: 1.483030194297617
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", cor16
         @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 = 1.027
Out[73]: 0.7522597087362669
In [74]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[74]: 2.048469397273288
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", cor17
         @printf("SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = %
         JCAll = cor17
         SSBRJC from Gibbs - G4.noGenotype.ID : correlation = 0.754
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 1.047
Out[75]: 0.7535066285902534
In [76]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[76]: 2.5985549977429296
In [77]: numSSBayes
Out[77]: SSBR.NumSSBayes(54918,45918,9000,40000,39000,1000,50)
```