/home/nicole/Jupyter/JG3/Data/0.5/G/6

```
In [4]:
         ;ls
         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.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
         ;awk '{print $1}' PedAll.txt | sort -b > all.ID
In [5]:
         ;awk '{print $1}' GenNF.txt | sort -b > genotype.ID
In [6]:
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
         ; 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
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
```

```
;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
                = 0.912
In [31]: vRes
                = 0.912
         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
         4868.321807 seconds (23.05 G allocations: 723.881 GB, 6.73% gc time)
In [32]: betaHat
Out[32]: 1-element Array{Float64,1}:
          8.47299
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 = 0.916
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.976
Out[35]: 0.9158075877223212
In [36]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[36]: 1.7383607242211485
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 = 0.970
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.043
Out[37]: 0.9700324625394183
In [38]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[38]: 3.0833345897701485
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 = 0.880
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.971
Out[39]: 0.8799783584167721
In [40]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[40]: 1.4279821398636865
```

```
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 = 0.734
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.946
Out[41]: 0.7341588331421665
In [42]: GEBV = aHat1[posAi]
         G0GEBV=mean(GEBV)
Out[42]: 0.1488911231787556
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 = 0.773
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.989
Out[43]: 0.7727973072912075
In [44]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[44]: 0.9183050728683545
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 = 0.775
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 0.996
Out[45]: 0.7748149144855352
In [46]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[46]: 1.5317432348756277
```

```
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 = 0.993
Out[47]: 0.7653252677940383
In [48]: GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[48]: 2.0821601456878844
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 = 0.774
         SSBRJC from Gibbs - G4.ID: regression of TBV on GEBV = 0.984
Out[49]: 0.7736735386058026
In [50]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[50]: 2.613906346419445
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 = 0.966
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 1.039
Out[51]: 0.9656823167329581
In [52]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[52]: 3.135158422296823
```

```
In [53]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         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 = 0.972
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.064
Out[53]: 0.9717937785250303
In [54]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[54]: 1.7236122642593061
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 = 1.051
Out[55]: 0.9739750126725593
In [56]: | GEBV = aHat1[posAi]
         mean(GEBV)
Out[56]: 2.1932463752192066
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 = 0.966
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.067
Out[57]: 0.9655804488359822
```

```
In [58]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[58]: 2.675917418739404
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 = cor11
         SSBRJC from Gibbs - G3.Genotype.ID : correlation = 0.958
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 1.020
Out[59]: 0.9578646013341774
In [60]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[60]: 3.1629929967660253
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 = 0.944
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 0.987
Out[61]: 0.9443458009889651
In [62]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[62]: 3.5879505927998268
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 = 0.966
         SSBRJC from Gibbs - G5.Genotype.ID: regression of TBV on GEBV = 1.039
Out[63]: 0.9656823167329581
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]: 11.560365500000001
In [67]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
Out[67]: 3.135158422296823
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 = 0.713
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 0.955
Out[68]: 0.712673713368802
In [69]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[69]: 0.10851365802284407
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 = 0.756
         SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 0.992
Out[70]: 0.7560461462590488
In [71]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[71]: 0.8856142702439738
```

```
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 = 0.760
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 0.997
Out[72]: 0.7603003724422013
In [73]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[73]: 1.5024054352893772
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 = 0.751
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 0.992
Out[74]: 0.7507430918439846
In [75]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[75]: 2.054446482839727
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 = 0.760
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 0.983
Out[76]: 0.7600714901239302
In [77]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[77]: 2.588930852922512
```

In [78]: numSSBayes

Out[78]: SSBR.NumSSBayes(54914,45914,9000,40000,39000,1000,200)