/home/nicole/Jupyter/JG3/Data/0.5/M/5

```
;ls
In [4]:
        Correlation.G5.M.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
        MarNF.txt
        PedAll.txt
        Phe.txt
        PheAll.txt
        Regression.G5.M.JC.txt
        all.ID
        alphaEstimatesJ
        alphaEstimatesJC
        epsiEstimatesJ
        epsiEstimatesJC
        genotype.ID
        meanOfSNPMAll
        meanOfSNPMG0
        meanOfSNPMG1
        meanOfSNPMG2
        meanOfSNPMG3
        meanOfSNPMG4
        meanOfSNPMG5
        noGenotype.ID
        sim.bv
        sim.phenotype
In [5]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
In [6]:
        ;awk '{print $1}' MarNF.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
         ped,A Mats,numSSBayes = calc Ai("PedAll.txt", "genotype.ID", calculateInbreedia
In [30]:
         nothing
         df
                = read_genotypes("MarNF.txt",numSSBayes)
         M Mats = make MMats(df, A Mats, ped, center=true);
                                                                                  # 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
                = 0.699
                = 0.699
         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
         4707.340248 seconds (23.03 G allocations: 723.502 GB, 6.87% gc time)
In [32]: | betaHat
Out[32]: 1-element Array{Float64,1}:
          11,3521
```

```
alphaHat
In [33]:
Out[33]: 150-element Array{Float64,1}:
          -0.11451
            0.137629
           0.0621145
           0.04685
          -0.0586008
           0.00141141
           0.05628
          -0.0385246
          -0.147184
          -0.239706
          -0.0162426
           0.0923494
          -0.084447
          -0.173824
          -0.0505006
           0.026149
           0.00953304
            0.115686
           0.127669
           0.0170349
           0.0562143
          -0.185126
          -0.166819
          -0.112611
          -0.0202665
In [34]: writedlm("alphaEstimatesC",alphaHat)
```

```
In [35]: epsiHat
Out[35]: 45913-element Array{Float64,1}:
           0.250884
          -0.0395315
          -0.470358
           0.207305
          -0.698409
          -0.116512
           0.131773
          -0.0713086
           0.799447
          -0.113168
          -0.256504
           0.264477
           0.293
          -0.122838
          -0.0255855
          -1.01586
           0.800038
          -0.748779
           0.672569
           0.123757
           0.646485
          -0.467264
          -0.291185
          -0.572184
          -0.129679
In [36]: writedlm("epsiEstimatesC",epsiHat)
In [37]: using DataFrames
In [38]: 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 [39]: 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 = 0.902
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.992
Out[39]: 0.9019819643086863
```

```
In [40]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[40]: 0.14151446873727166
In [41]: 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 = 0.850
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.017
Out[41]: 0.8497919627309943
In [42]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[42]: 1.3065152776530824
In [43]: 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 = 0.870
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.966
Out[43]: 0.870270602710462
In [44]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[44]: -0.12733187178176172
In [45]: 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 ep
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.701
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.854
Out[45]: 0.7005863943092451
```

```
In [46]: | GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[46]: -1.2737789885986521
In [47]: 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 =
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.982
Out[47]: 0.7646325533308528
In [48]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[48]: -0.5774002217610116
In [49]: 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 ei
         @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor5
         SSBRJC from Gibbs - G2.ID : correlation = 0.762
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 1.009
Out[49]: 0.7616460800455284
In [50]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[50]: -0.02924427775704592
In [51]: 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 = 0.748
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 1.025
Out[51]: 0.747881467381551
```

```
In [52]: | GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[52]: 0.4539383512132015
In [53]: 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 =
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 1.033
Out[53]: 0.7525175078025551
In [54]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[54]: 0.9273513399000162
In [55]: 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 = 0.832
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 0.989
Out[55]: 0.8321903861737039
In [56]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[56]: 1.3482206094271219
In [57]: GEBVG5G1=G5GEBV-G1GEBV
Out[57]: 1.9256208311881333
In [58]: GEBVG1G5=[G1GEBV;G2GEBV;G3GEBV;G4GEBV;G5GEBV]
Out[58]: 5-element Array{Float64,1}:
          -0.5774
          -0.0292443
           0.453938
           0.927351
           1.34822
```

```
In [59]: reg8 = linreg(aHat1[posAi], a[posAi])
Out[59]: 2-element Array{Float64,1}:
          11.4587
           0.988866
In [60]: VarGEBV=var(aHat1[posAi])
Out[60]: 0.39123267861652106
In [61]: VarTBV=var(a[posAi])
Out[61]: 0.5524134329322257
In [62]: Cov=cov(aHat1[posAi], a[posAi])
Out[62]: 0.38687651870494344
In [63]: b=Cov/VarGEBV
Out[63]: 0.9888655520111922
In [64]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         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 )
         @printf("SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor9
         SSBRJC from Gibbs - G0.Genotype.ID : correlation = 0.838
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 0.938
Out[64]: 0.8380138201846167
In [65]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[65]: 0.19700533270641746
In [66]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ', head
         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 )
         @printf("SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor9
         SSBRJC from Gibbs - G1.Genotype.ID : correlation =
                                                              0.802
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 0.843
Out[66]: 0.8015513825674975
```

```
In [67]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[67]: 0.6119758402622927
In [68]: 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 =
         SSBRJC from Gibbs - G2.Genotype.ID: regression of TBV on GEBV = 0.887
Out[68]: 0.8085198534572063
In [69]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[69]: 0.9647737791698531
In [70]: 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 = 0.803
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 0.940
Out[70]: 0.8033263912606625
In [71]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[71]: 1.3664750900945595
In [72]: 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 = 0.771
         SSBRJC from Gibbs - G4.Genotype.ID: regression of TBV on GEBV = 0.869
Out[72]: 0.7706745672290699
```

```
In [73]: | GEBV = aHat1[posAi]
         mean(GEBV)
Out[73]: 1.7241330750707127
In [74]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         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 =
         SSBRJC from Gibbs - G5.Genotype.ID: regression of TBV on GEBV = 0.989
Out[74]: 0.8321903861737039
In [75]: writedlm("Correlation.G5.M.C.txt",cor13)
In [76]: | writedlm("Regression.G5.M.C.txt",reg13)
In [77]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
Out[77]: 1.3482206094271219
In [78]: 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 = 0.682
         SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 0.868
Out[78]: 0.6822509415674405
In [79]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[79]: -1.3114914070936539
```

```
In [80]: 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 = 0.750
         SSBRJC from Gibbs - G1.noGenotype.ID: regression of TBV on GEBV = 0.992
Out[80]: 0.7498350713103535
In [81]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[81]: -0.6078970438641733
In [82]: 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 = 0.748
         SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = 1.012
Out[82]: 0.7481203573041324
In [83]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[83]: -0.05473192024235104
In [84]: 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 = 0.733
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 1.020
Out[84]: 0.7330139166554166
In [85]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[85]: 0.4305399732931667
```