/home/nicole/Jupyter/JG3/Data/0.5/M/9

```
In [4]:
        ;ls
        Correlation.G5.M.J.txt
        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.J.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
        ;awk '{print $1,$3}' PheAll.txt > sim.bv
In [9]:
```

```
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
                = 0.547
                = 0.547
         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
         4520.332939 seconds (23.05 G allocations: 724.232 GB, 7.16% gc time)
In [32]: | betaHat
Out[32]: 1-element Array{Float64,1}:
          11.2806
```

```
alphaHat
In [33]:
Out[33]: 150-element Array{Float64,1}:
          -0.0795695
           0.137686
          -0.0220487
            0.0191924
          -0.000652375
            0.00246583
           0.0615674
           0.0237011
          -0.0767516
          -0.0190863
          -0.0842026
          -0.222416
           0.0057416
           0.072624
           0.216184
          -0.111877
          -0.0638976
          -0.0831753
           0.00465668
          -0.0374027
           0.014706
            0.0221958
            0.101201
            0.0757881
            0.0647147
In [34]: writedlm("alphaEstimatesC",alphaHat)
```

```
In [35]: epsiHat
Out[35]: 45966-element Array{Float64,1}:
          -0.0916295
           0.380812
           0.222244
          -0.380199
          -0.215554
           0.528038
          -0.732845
          -1.40156
           0.0676259
          -0.548458
          -0.460816
           0.381173
          -0.425598
          -0.334774
          -0.335523
          -0.329484
           0.381096
          -0.0213621
          -0.472043
          -0.440734
          -0.623435
           0.0737569
          -0.779512
           0.365264
           0.279394
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 = 0.891
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.995
Out[39]: 0.8907817302401693
```

```
In [40]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[40]: 0.12024635086060864
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 = 0.783
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 0.989
Out[41]: 0.7833775882609103
In [42]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[42]: 1.0785353266291156
In [43]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',heade
         posAi = getPos(ped,IDs)
         cor3 = cor(a[posAi],aHat1[posAi])[1,1]
         req3 = 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 = 0.864
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.956
Out[43]: 0.8636273912125694
In [44]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[44]: -0.10089725893212372
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 = 0.670
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.776
Out[45]: 0.6703832944477273
```

```
In [46]: | GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[46]: -1.1181462035828031
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 = 0.762
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.969
Out[47]: 0.7623289340152928
In [48]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[48]: -0.49896515825614146
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 = 0.752
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 1.006
Out[49]: 0.7524474899886774
In [50]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[50]: -0.012561622263623897
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 = 0.745
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 1.002
Out[51]: 0.7446736892540488
```

```
In [52]: | GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[52]: 0.42864942704103026
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 = 1.011
Out[53]: 0.7271937049516161
In [54]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[54]: 0.814411798138286
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 = 0.759
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 0.942
Out[55]: 0.7592398206255224
In [56]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[56]: 1.1080898640869046
In [57]: GEBVG5G1=G5GEBV-G1GEBV
Out[57]: 1.607055022343046
In [58]: GEBVG1G5=[G1GEBV;G2GEBV;G3GEBV;G4GEBV;G5GEBV]
Out[58]: 5-element Array{Float64,1}:
          -0.498965
          -0.0125616
           0.428649
           0.814412
           1.10809
```

```
In [59]: reg8 = linreg(aHat1[posAi], a[posAi])
Out[59]: 2-element Array{Float64,1}:
          11.492
           0.941744
In [60]: VarGEBV=var(aHat1[posAi])
Out[60]: 0.25958100952664703
In [61]: VarTBV=var(a[posAi])
Out[61]: 0.39937518948266976
In [62]: Cov=cov(aHat1[posAi], a[posAi])
Out[62]: 0.2444589698024204
In [63]: b=Cov/VarGEBV
Out[63]: 0.9417444297955306
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 = 0.822
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 0.962
Out[64]: 0.821787758780432
In [65]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[65]: 0.2433446670136901
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 =
                                                              0.764
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 0.837
Out[66]: 0.7638054726036252
```

```
In [67]: GEBV = aHat1[posAi]
         mean (GEBV)
Out[67]: 0.5577794047454644
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 = 0.759
         SSBRJC from Gibbs - G2.Genotype.ID: regression of TBV on GEBV = 0.829
Out[68]: 0.7593396434830063
In [69]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[69]: 0.8788854414000086
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 = 0.711
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 0.752
Out[70]: 0.7111437012078966
In [71]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[71]: 1.1491383055352669
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 = 0.672
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 0.770
Out[72]: 0.6720724673192905
```

```
In [73]: | GEBV = aHat1[posAi]
         mean(GEBV)
Out[73]: 1.3813473161395922
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 = 0.942
Out[74]: 0.7592398206255224
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]: 1.1080898640869046
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 = 0.652
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 0.788
Out[78]: 0.6522388296732772
In [79]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[79]: -1.1530562259057904
```

```
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 = 0.748
         SSBRJC from Gibbs - G1.noGenotype.ID: regression of TBV on GEBV = 0.976
Out[80]: 0.7476140536063028
In [81]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[81]: -0.5260611726920801
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 = 0.737
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 1.003
Out[82]: 0.7366685066397491
In [83]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[83]: -0.035419239280640086
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 = 0.732
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 0.993
Out[84]: 0.7319177028529962
In [85]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[85]: 0.41017535323348575
```