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

```
;ls
In [4]:
        Correlation.G5.G.C.txt
        Correlation.G5.G.J.txt
        Correlation.G5.G.JC.txt
        Correlation.G5.G.N.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
        Regression.G5.G.N.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]:
        ;join -v1 all.ID genotype.ID > noGenotype.ID
        ;awk '{print $1,$2}' Phe.txt > sim.phenotype
In [8]:
```

```
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
        ; awk 'NR >=8001 && NR <=16000 {print $1}' PedAll.txt | sort -b > G1.ID
In [11]:
        ; 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 [14]:
         ; 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
        ; join G4.ID genotype.ID > G4.Genotype.ID
In [20]:
In [21]: |; join G5.ID genotype.ID > G5.Genotype.ID
In [22]:
        ; join -v1 G0.ID genotype.ID > G0.noGenotype.ID
In [23]: | ;join -v1 G1.ID genotype.ID > G1.noGenotype.ID
        ; join -v1 G2.ID genotype.ID > G2.noGenotype.ID
In [24]:
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 G1.Genotype.ID
          200
               200 1200 G2.Genotype.ID
          200
          200
               200 1200 G3.Genotype.ID
              200 1200 G4.Genotype.ID
          200
          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", calculateInbreedir
         nothing
         y Vecs = make_yVecs("sim.phenotype",ped,numSSBayes)
         Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
         nothing
        vRes
                = 0.759
In [31]:
         vG
                = 0.759
         aHat3=SSBR.PBLUP(y_Vecs,Z_Mats,A_Mats,numSSBayes,vRes,vG);
In [32]: using DataFrames
In [33]: 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 [34]: IDs = readtable("all.ID", eltypes =[UTF8String], separator = ' ', header=false
         posAi = getPos(ped,IDs)
         cor1 = cor(a[posAi],aHat3[posAi])[1,1]
         reg1 = linreg(aHat3[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - all.ID : correlation = %6.3f\n", cor1 ) # with \epsilon
         @printf("SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = %6.3f\n",
         JCAll = cor1
         SSBRJC from Gibbs - all.ID : correlation = 0.895
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.989
Out[34]: 0.8949350914114739
In [35]: TBV = a[posAi]
         mean(TBV)
Out[35]: 10.471534229166668
In [36]: EBV = aHat3[posAi]
         mean(EBV)
Out[36]: 1.4838002798718413
```

```
In [37]: IDs = readtable("genotype.ID", eltypes =[UTF8String], separator = ' ',header=
         posAi = getPos(ped,IDs)
         cor2 = cor(a[posAi], aHat3[posAi])[1,1]
         reg2 = linreg(aHat3[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 = 0.574
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 0.982
Out[37]: 0.5740746779016379
In [38]: TBV = a[posAi]
         mean(TBV)
Out[38]: 11.82475322222222
In [39]: EBV = aHat3[posAi]
         mean(EBV)
Out[39]: 2.8397722613140397
In [40]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',heade
         posAi = getPos(ped,IDs)
         cor3 = cor(a[posAi],aHat3[posAi])[1,1]
         reg3 = linreg(aHat3[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 = 0.881
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.985
Out[40]: 0.8811653687884569
In [41]: TBV = a[posAi]
         mean(TBV)
Out[41]: 10.159252923076924
In [42]: EBV = aHat3[posAi]
         mean(EBV)
Out[42]: 1.1708836687697954
```

```
In [43]: IDs = readtable("G0.ID", eltypes =[UTF8String], separator = ' ',header=false)
         posAi = getPos(ped,IDs)
         cor4 = cor(a[posAi], aHat3[posAi])[1,1]
         reg4 = linreg(aHat3[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 = 0.908
Out[43]: 0.7212422580935137
In [44]: TBV = a[posAi]
         mean(TBV)
Out[44]: 9.017859375
In [45]: EBV = aHat3[posAi]
         mean(EBV)
Out[45]: 0.002465054960012266
In [46]: IDs = readtable("G1.ID", eltypes =[UTF8String], separator = ' ',header=false)
         posAi = getPos(ped,IDs)
         cor4 = cor(a[posAi],aHat3[posAi])[1,1]
         reg4 = linreg(aHat3[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 = 0.775
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 1.002
Out[46]: 0.7745964955220455
In [47]: TBV = a[posAi]
         mean(TBV)
Out[47]: 9.608738375000001
In [48]: EBV = aHat3[posAi]
         mean(EBV)
Out[48]: 0.6344625424304675
```

```
In [49]: IDs = readtable("G2.ID", eltypes =[UTF8String], separator = ' ',header=false)
         posAi = getPos(ped,IDs)
         cor5 = cor(a[posAi],aHat3[posAi])[1,1]
         reg5 = linreg(aHat3[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 = 0.772
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 1.016
Out[49]: 0.7724334527067311
In [50]: TBV = a[posAi]
         mean(TBV)
Out[50]: 10.183037250000002
In [51]: EBV = aHat3[posAi]
         mean(EBV)
Out[51]: 1.202866143015828
In [52]: IDs = readtable("G3.ID", eltypes =[UTF8String], separator = ' ',header=false)
         posAi = getPos(ped,IDs)
         cor6 = cor(a[posAi],aHat3[posAi])[1,1]
         reg6 = linreg(aHat3[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 = 0.763
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 0.987
Out[52]: 0.7627675032122285
In [53]: TBV = a[posAi]
         mean(TBV)
Out[53]: 10.773109375
In [54]: EBV = aHat3[posAi]
         mean(EBV)
Out[54]: 1.7876206471642657
```

```
In [55]: IDs = readtable("G4.ID", eltypes =[UTF8String], separator = ' ',header=false)
         posAi = getPos(ped,IDs)
         cor7 = cor(a[posAi],aHat3[posAi])[1,1]
         reg7 = linreg(aHat3[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 = 0.728
         SSBRJC from Gibbs - G4.ID: regression of TBV on GEBV = 0.983
Out[55]: 0.7284631787285037
In [56]: TBV = a[posAi]
         mean(TBV)
Out[56]: 11.359387625
In [57]: EBV = aHat3[posAi]
         mean(EBV)
Out[57]: 2.375624017184832
In [58]: IDs = readtable("G5.ID", eltypes =[UTF8String], separator = ' ',header=false)
         posAi = getPos(ped,IDs)
         cor8 = cor(a[posAi],aHat3[posAi])[1,1]
         reg8 = linreg(aHat3[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 = 0.407
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 0.938
Out[58]: 0.4072218569151642
In [59]: TBV = a[posAi]
         mean(TBV)
Out[59]: 11.887073375
In [60]: EBV = aHat3[posAi]
         mean(EBV)
Out[60]: 2.899763274475641
```

```
In [61]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat3[posAi])[1,1]
         reg9 = linreg(aHat3[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 = 0.910
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.037
Out[61]: 0.9099508305165308
In [62]: TBV = a[posAi]
         mean(TBV)
Out[62]: 10.192884999999997
In [63]: EBV = aHat3[posAi]
         mean(EBV)
Out[63]: 1.269275259658478
In [64]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat3[posAi])[1,1]
         reg9 = linreg(aHat3[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 = 0.911
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.000
Out[64]: 0.9107117414397974
In [65]: TBV = a[posAi]
         mean(TBV)
Out[65]: 10.736695
In [66]: EBV = aHat3[posAi]
         mean(EBV)
Out[66]: 1.7615578687453666
```

```
In [67]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor10 = cor(a[posAi],aHat3[posAi])[1,1]
         reg10 = linreg(aHat3[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 = 0.895
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 0.954
Out[67]: 0.8952042706536661
In [68]: TBV = a[posAi]
         mean(TBV)
Out[68]: 11.359055000000001
In [69]: EBV = aHat3[posAi]
         mean(EBV)
Out[69]: 2.3737098381020076
In [70]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor11 = cor(a[posAi],aHat3[posAi])[1,1]
         reg11 = linreg(aHat3[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 = 0.867
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 0.946
Out[70]: 0.8674310601259959
In [71]: TBV = a[posAi]
         mean(TBV)
Out[71]: 11.936440000000003
In [72]: EBV = aHat3[posAi]
         mean(EBV)
Out[72]: 2.9674354465970767
```

```
In [73]: IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor12 = cor(a[posAi],aHat3[posAi])[1,1]
         reg12 = linreg(aHat3[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 = 0.471
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 0.926
Out[73]: 0.4712850515779606
In [74]: TBV = a[posAi]
         mean(TBV)
Out[74]: 12.405885000000001
In [75]: EBV = aHat3[posAi]
         mean(EBV)
Out[75]: 3.427242367003219
In [76]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor13 = cor(a[posAi],aHat3[posAi])[1,1]
         reg13 = linreg(aHat3[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 = 0.407
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 0.938
Out[76]: 0.4072218569151642
In [77]: writedlm("Correlation.G5.G.PBLUP.txt",cor13)
In [78]: | writedlm("Regression.G5.G.PBLUP.txt", reg13)
In [79]:
         TBV = a[posAi]
         mean(TBV)
Out[79]: 11.887073375
In [80]: EBV = aHat3[posAi]
         mean(EBV)
Out[80]: 2.899763274475641
```

```
In [81]: IDs = readtable("G0.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat3[posAi])[1,1]
         reg14 = linreg(aHat3[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 = 0.904
Out[81]: 0.7009792299634106
In [82]: TBV = a[posAi]
         mean(TBV)
Out[82]: 8.987730512820514
In [83]: EBV = aHat3[posAi]
         mean(EBV)
Out[83]: -0.030017257980974023
In [84]: IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat3[posAi])[1,1]
         reg14 = linreg(aHat3[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 = 0.759
         SSBRJC from Gibbs - G1.noGenotype.ID: regression of TBV on GEBV = 1.002
Out[84]: 0.7592151448718636
In [85]: TBV = a[posAi]
         mean(TBV)
Out[85]: 9.57981641025641
In [86]: EBV = aHat3[posAi]
         mean(EBV)
Out[86]: 0.6055626622685472
```

```
In [87]: IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor15 = cor(a[posAi],aHat3[posAi])[1,1]
         reg15 = linreg(aHat3[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 = 0.756
         SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = 1.019
Out[87]: 0.7558533747385638
In [88]: TBV = a[posAi]
         mean(TBV)
Out[88]: 10.15288294871795
In [89]: EBV = aHat3[posAi]
         mean(EBV)
Out[89]: 1.1728445098084903
In [90]: IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor16 = cor(a[posAi],aHat3[posAi])[1,1]
         reg16 = linreg(aHat3[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 = 0.745
         SSBRJC from Gibbs - G3.noGenotype.ID: regression of TBV on GEBV = 0.988
Out[90]: 0.7450944608421075
In [91]: TBV = a[posAi]
         mean(TBV)
Out[91]: 10.743280384615385
In [92]: EBV = aHat3[posAi]
         mean(EBV)
Out[92]: 1.7573689856403474
```

```
In [93]: | IDs = readtable("G4.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor17 = cor(a[posAi],aHat3[posAi])[1,1]
         reg17 = linreg(aHat3[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 = 0.715
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 0.983
Out[93]: 0.7153795097824844
In [94]: TBV = a[posAi]
         mean(TBV)
Out[94]: 11.33255435897436
In [95]: EBV = aHat3[posAi]
         mean(EBV)
Out[95]: 2.3486594441125663
In [96]: numSSBayes
Out[96]: SSBR.NumSSBayes(54918,45918,9000,40000,39000,1000,0)
```