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

```
;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
In [5]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
In [6]: | ;awk '{print $1}' GenNF.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
         ; 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
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
         ;wc G0.Genotype.ID;wc G1.Genotype.ID;wc G2.Genotype.ID;wc G3.Genotype.ID;wc (
In [28]:
              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 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", calculateInbreedia
         nothing
         y Vecs = make_yVecs("sim.phenotype",ped,numSSBayes)
         Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
         nothing
In [31]:
         vRes
                = 0.912
                = 0.912
         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 (
         @printf("SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = %6.3f\n",
         JCAll = cor1
         SSBRJC from Gibbs - all.ID : correlation = 0.887
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.977
Out[34]: 0.8874110503252064
In [35]: TBV = a[posAi]
         mean(TBV)
Out[35]: 10.199740145833333
In [36]: EBV = aHat3[posAi]
         mean(EBV)
Out[36]: 1.5251149353656235
```

```
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 ) # |
         @printf("SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = %6.3:
         JCAll = cor2
         SSBRJC from Gibbs - genotype.ID : correlation = 0.560
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 0.935
Out[37]: 0.559948484173743
In [38]: TBV = a[posAi]
         mean(TBV)
Out[38]: 11.50541944444444
In [39]: EBV = aHat3[posAi]
         mean(EBV)
Out[39]: 2.8925493565749867
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 = 0.874
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.990
Out[40]: 0.8738312866017834
In [41]: TBV = a[posAi]
         mean(TBV)
Out[41]: 9.898429538461539
In [42]: EBV = aHat3[posAi]
         mean(EBV)
Out[42]: 1.2095531458557702
```

```
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 ei
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.732
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.962
Out[43]: 0.7323531978729594
In [44]: TBV = a[posAi]
         mean(TBV)
Out[44]: 8.683999375
In [45]: EBV = aHat3[posAi]
         mean(EBV)
Out[45]: 0.005556015522373465
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 ep
         @printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G1.ID : correlation = 0.766
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.989
Out[46]: 0.7658380411017554
In [47]: TBV = a[posAi]
         mean(TBV)
Out[47]: 9.382714874999998
In [48]: EBV = aHat3[posAi]
         mean(EBV)
Out[48]: 0.6890557447239849
```

```
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 ei
         @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor5
         SSBRJC from Gibbs - G2.ID : correlation = 0.766
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 0.989
Out[49]: 0.7658992534211417
In [50]: TBV = a[posAi]
         mean(TBV)
Out[50]: 9.982135999999997
In [51]: EBV = aHat3[posAi]
         mean(EBV)
Out[51]: 1.289869786297254
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 ep
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor6
         SSBRJC from Gibbs - G3.ID : correlation = 0.753
         SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = 0.973
Out[52]: 0.753231796413395
In [53]: TBV = a[posAi]
         mean(TBV)
Out[53]: 10.526489125
In [54]: EBV = aHat3[posAi]
         mean(EBV)
Out[54]: 1.837667278547946
```

```
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 ei
         @printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor7
         SSBRJC from Gibbs - G4.ID : correlation = 0.716
         SSBRJC from Gibbs - G4.ID: regression of TBV on GEBV = 0.920
Out[55]: 0.7158789136901965
In [56]: TBV = a[posAi]
         mean(TBV)
Out[56]: 11.062736000000001
In [57]: EBV = aHat3[posAi]
         mean(EBV)
Out[57]: 2.3780148931718204
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 ep
         @printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor8
         SSBRJC from Gibbs - G5.ID : correlation = 0.409
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 0.911
Out[58]: 0.40944131973243164
In [59]: TBV = a[posAi]
         mean(TBV)
Out[59]: 11.560365500000001
In [60]: EBV = aHat3[posAi]
         mean(EBV)
Out[60]: 2.950525893930361
```

```
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 = 0.906
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.049
Out[61]: 0.9063060736702027
In [62]: TBV = a[posAi]
         mean(TBV)
Out[62]: 10.039674999999999
In [63]: EBV = aHat3[posAi]
         mean(EBV)
Out[63]: 1.377070751422532
In [64]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         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 = 0.872
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.011
Out[64]: 0.8720636545138885
In [65]: TBV = a[posAi]
         mean(TBV)
Out[65]: 10.57286
In [66]: EBV = aHat3[posAi]
         mean(EBV)
Out[66]: 1.8947013573209528
```

```
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 = 0.904
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.037
Out[67]: 0.9035665031322626
In [68]: TBV = a[posAi]
         mean(TBV)
Out[68]: 11.076469999999999
In [69]: EBV = aHat3[posAi]
         mean(EBV)
Out[69]: 2.3971575515613135
In [70]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         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 = 0.842
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 0.902
Out[70]: 0.8417749949604869
In [71]: TBV = a[posAi]
         mean(TBV)
Out[71]: 11.59739999999999
In [72]: EBV = aHat3[posAi]
         mean(EBV)
Out[72]: 2.9425813857440613
```

```
In [73]: IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         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 = 0.417
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 0.809
Out[73]: 0.41723115322474985
In [74]: TBV = a[posAi]
         mean(TBV)
Out[74]: 12.042850000000001
In [75]: EBV = aHat3[posAi]
         mean(EBV)
Out[75]: 3.532174242611098
In [76]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         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 = 0.409
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 0.911
Out[76]: 0.40944131973243164
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]: 11.560365500000001
In [80]: | EBV = aHat3[posAi]
         mean(EBV)
Out[80]: 2.950525893930361
```

```
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", corl
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor14
         SSBRJC from Gibbs - G0.noGenotype.ID : correlation = 0.712
         SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 0.957
Out[81]: 0.7115960805745596
In [82]: TBV = a[posAi]
         mean(TBV)
Out[82]: 8.649238461538461
In [83]: EBV = aHat3[posAi]
         mean(EBV)
Out[83]: -0.029611028987887016
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", 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 = 0.988
Out[84]: 0.7504180273999741
In [85]: TBV = a[posAi]
         mean(TBV)
Out[85]: 9.352198333333334
In [86]: EBV = aHat3[posAi]
         mean(EBV)
Out[86]: 0.6581417546573959
```

```
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", cor1!
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor15
         SSBRJC from Gibbs - G2.noGenotype.ID : correlation = 0.752
         SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = 0.988
Out[87]: 0.751785666891681
In [88]: TBV = a[posAi]
         mean(TBV)
Out[88]: 9.954076153846152
In [89]: EBV = aHat3[posAi]
         mean(EBV)
Out[89]: 1.2614777923161242
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", cor1
         @printf("SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor16
         SSBRJC from Gibbs - G3.noGenotype.ID : correlation = 0.739
         SSBRJC from Gibbs - G3.noGenotype.ID: regression of TBV on GEBV = 0.975
Out[90]: 0.7394274270641495
In [91]: TBV = a[posAi]
         mean(TBV)
Out[91]: 10.499029871794875
In [92]: EBV = aHat3[posAi]
         mean(EBV)
Out[92]: 1.8093361475941991
```

```
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", cor1
         @printf("SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor17
         SSBRJC from Gibbs - G4.noGenotype.ID : correlation = 0.705
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 0.928
Out[93]: 0.7050976011130451
In [94]: TBV = a[posAi]
         mean(TBV)
Out[94]: 11.037604871794871
In [95]: EBV = aHat3[posAi]
         mean(EBV)
Out[95]: 2.348421063699019
In [96]: numSSBayes
Out[96]: SSBR.NumSSBayes(54914,45914,9000,40000,39000,1000,0)
```

https://cuda3.ansci.iastate.edu:9999/notebooks/Jupyter/JG3/J-PBLUP.ipynb