/home/nicole/Jupyter/JG3/Data/0.5/G/3

```
;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
                = 0.664
                = 0.664
         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 = 0.882
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.978
Out[34]: 0.8824953881456271
In [35]: TBV = a[posAi]
         mean(TBV)
Out[35]: 11.135083416666665
In [36]: EBV = aHat3[posAi]
         mean(EBV)
Out[36]: 1.2976544212425094
```

```
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 = 0.549
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 0.951
Out[37]: 0.5493480005956689
In [38]: TBV = a[posAi]
         mean(TBV)
Out[38]: 12.27303822222223
In [39]: EBV = aHat3[posAi]
         mean(EBV)
Out[39]: 2.463378187392464
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 = 0.868
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.980
Out[40]: 0.8679961320422401
In [41]: TBV = a[posAi]
         mean(TBV)
Out[41]: 10.872478461538462
In [42]: EBV = aHat3[posAi]
         mean(EBV)
Out[42]: 1.0286412444386737
```

```
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 = 0.700
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.877
Out[43]: 0.6995711216889441
In [44]: TBV = a[posAi]
         mean(TBV)
Out[44]: 9.84533075
In [45]: EBV = aHat3[posAi]
         mean(EBV)
Out[45]: 0.00151502395740011
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 = 0.773
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.984
Out[46]: 0.7725871354461374
In [47]: TBV = a[posAi]
         mean(TBV)
Out[47]: 10.40474075
In [48]: EBV = aHat3[posAi]
         mean(EBV)
Out[48]: 0.5736346558759015
```

```
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 = 0.750
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 0.979
Out[49]: 0.7496561352919332
In [50]: TBV = a[posAi]
         mean(TBV)
Out[50]: 10.945704874999999
In [51]: EBV = aHat3[posAi]
         mean(EBV)
Out[51]: 1.0959091680294455
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 = 0.738
         SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = 0.972
Out[52]: 0.7382760419479034
In [53]: TBV = a[posAi]
         mean(TBV)
Out[53]: 11.422414125000001
In [54]: EBV = aHat3[posAi]
         mean(EBV)
Out[54]: 1.5744754256824816
```

```
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 = 0.712
         SSBRJC from Gibbs - G4.ID: regression of TBV on GEBV = 0.935
Out[55]: 0.7115018929422184
In [56]: TBV = a[posAi]
         mean(TBV)
Out[56]: 11.871169375
In [57]: EBV = aHat3[posAi]
         mean(EBV)
Out[57]: 2.0280087925129235
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 = 0.398
         SSBRJC from Gibbs - G5.ID: regression of TBV on GEBV = 0.905
Out[58]: 0.3984199348489252
In [59]: TBV = a[posAi]
         mean(TBV)
Out[59]: 12.321140625
In [60]: EBV = aHat3[posAi]
         mean(EBV)
Out[60]: 2.512383461396904
```

```
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 = 0.900
         SSBRJC from Gibbs - G0.Genotype.ID: regression of TBV on GEBV = 1.084
Out[61]: 0.8998236346247229
In [62]: TBV = a[posAi]
         mean(TBV)
Out[62]: 10.961675
In [63]: EBV = aHat3[posAi]
         mean(EBV)
Out[63]: 1.1557836474378849
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 = 0.879
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.014
Out[64]: 0.8793293474424341
In [65]: TBV = a[posAi]
         mean(TBV)
Out[65]: 11.46430000000001
In [66]: EBV = aHat3[posAi]
         mean(EBV)
Out[66]: 1.6351374369606535
```

```
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 = 0.851
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 0.952
Out[67]: 0.8514746133453182
In [68]: TBV = a[posAi]
         mean(TBV)
Out[68]: 11.912560000000003
In [69]: EBV = aHat3[posAi]
         mean(EBV)
Out[69]: 2.0660397984886765
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 = 0.896
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 1.026
Out[70]: 0.8964757382936054
In [71]: TBV = a[posAi]
         mean(TBV)
Out[71]: 12.34098
In [72]: EBV = aHat3[posAi]
         mean(EBV)
Out[72]: 2.5054662192726234
```

```
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 = 0.465
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 0.929
Out[73]: 0.46531491648031875
In [74]: TBV = a[posAi]
         mean(TBV)
Out[74]: 12.761580000000002
In [75]: EBV = aHat3[posAi]
         mean(EBV)
Out[75]: 2.9942528746248795
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 = 0.398
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 0.905
Out[76]: 0.3984199348489252
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]: 12.321140625
In [80]: | EBV = aHat3[posAi]
         mean(EBV)
Out[80]: 2.512383461396904
```

```
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 = 0.677
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 0.865
Out[81]: 0.6774609051804127
In [82]: TBV = a[posAi]
         mean(TBV)
Out[82]: 9.81670653846154
In [83]: EBV = aHat3[posAi]
         mean(EBV)
Out[83]: -0.02808160741389438
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 = 0.757
         SSBRJC from Gibbs - G1.noGenotype.ID: regression of TBV on GEBV = 0.982
Out[84]: 0.7565074097241409
In [85]: TBV = a[posAi]
         mean(TBV)
Out[85]: 10.377572564102563
In [86]: EBV = aHat3[posAi]
         mean(EBV)
Out[86]: 0.5464166358480872
```

```
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 = 0.734
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 0.978
Out[87]: 0.7339185879850127
In [88]: TBV = a[posAi]
         mean(TBV)
Out[88]: 10.92091371794872
In [89]: EBV = aHat3[posAi]
         mean(EBV)
Out[89]: 1.0710340236586962
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 = 0.721
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 0.970
Out[90]: 0.7213941215088447
In [91]: TBV = a[posAi]
         mean(TBV)
Out[91]: 11.398861153846154
In [92]: EBV = aHat3[posAi]
         mean(EBV)
Out[92]: 1.550603866872478
```

```
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 = 0.699
         SSBRJC from Gibbs - G4.noGenotype.ID: regression of TBV on GEBV = 0.936
Out[93]: 0.6990894082914335
In [94]: TBV = a[posAi]
         mean(TBV)
Out[94]: 11.8483383333333334
In [95]: EBV = aHat3[posAi]
         mean(EBV)
Out[95]: 2.0032333032280016
In [96]: numSSBayes
Out[96]: SSBR.NumSSBayes(54866,45866,9000,40000,39000,1000,0)
```