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

```
;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
                = 0.699
In [31]:
         vG
                = 0.699
         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 = 0.884
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.975
Out[34]: 0.8837159378609749
In [35]: TBV = a[posAi]
         mean(TBV)
Out[35]: 11.514132458333334
In [36]: EBV = aHat3[posAi]
         mean(EBV)
Out[36]: 1.3608531455769752
```

```
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 = 0.549
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 0.971
Out[37]: 0.5487914932280055
In [38]: TBV = a[posAi]
         mean(TBV)
Out[38]: 12.736877111111111
In [39]: EBV = aHat3[posAi]
         mean(EBV)
Out[39]: 2.6281006947094694
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 = 0.871
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.981
Out[40]: 0.8709760496446026
In [41]: TBV = a[posAi]
         mean(TBV)
Out[41]: 11.231960615384617
In [42]: EBV = aHat3[posAi]
         mean(EBV)
Out[42]: 1.0684114034694763
```

```
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 = 0.703
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.877
Out[43]: 0.7025869059763595
In [44]: TBV = a[posAi]
         mean(TBV)
Out[44]: 10.187546249999999
In [45]: EBV = aHat3[posAi]
         mean(EBV)
Out[45]: -0.0012330787761268382
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 = 0.770
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.986
Out[46]: 0.7695038529551359
In [47]: TBV = a[posAi]
         mean(TBV)
Out[47]: 10.7345095
In [48]: EBV = aHat3[posAi]
         mean(EBV)
Out[48]: 0.5933826959650165
```

```
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 = 0.770
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 1.008
Out[49]: 0.7698933888906929
In [50]: TBV = a[posAi]
         mean(TBV)
Out[50]: 11.27873
In [51]: EBV = aHat3[posAi]
         mean(EBV)
Out[51]: 1.115851127843941
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 = 0.752
         SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = 1.000
Out[52]: 0.7517874706985195
In [53]: TBV = a[posAi]
         mean(TBV)
Out[53]: 11.789662125
In [54]: EBV = aHat3[posAi]
         mean(EBV)
Out[54]: 1.6209014121375802
```

```
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 = 0.726
         SSBRJC from Gibbs - G4.ID: regression of TBV on GEBV = 0.975
Out[55]: 0.7260201658999998
In [56]: TBV = a[posAi]
         mean(TBV)
Out[56]: 12.30248425
In [57]: EBV = aHat3[posAi]
         mean(EBV)
Out[57]: 2.1516343337973476
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 = 0.402
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 0.958
Out[58]: 0.4015215812384653
In [59]: TBV = a[posAi]
         mean(TBV)
Out[59]: 12.791862625000002
In [60]: EBV = aHat3[posAi]
         mean(EBV)
Out[60]: 2.6845823824940918
```

```
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 = 0.886
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 0.981
Out[61]: 0.8864881440255429
In [62]: TBV = a[posAi]
         mean(TBV)
Out[62]: 11.279599999999999
In [63]: EBV = aHat3[posAi]
         mean(EBV)
Out[63]: 1.1994332438728719
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 = 0.880
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 0.981
Out[64]: 0.8800984505023802
In [65]: TBV = a[posAi]
         mean(TBV)
Out[65]: 11.827124999999997
In [66]: EBV = aHat3[posAi]
         mean(EBV)
Out[66]: 1.6604189877017697
```

```
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 = 0.886
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 0.993
Out[67]: 0.8860873110479595
In [68]: TBV = a[posAi]
         mean(TBV)
Out[68]: 12.28475999999997
In [69]: EBV = aHat3[posAi]
         mean(EBV)
Out[69]: 2.1217750695517226
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 = 0.861
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 1.016
Out[70]: 0.8606697020077982
In [71]: TBV = a[posAi]
         mean(TBV)
Out[71]: 12.79993999999998
In [72]: EBV = aHat3[posAi]
         mean(EBV)
Out[72]: 2.6805189121374697
```

```
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 = 0.412
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 0.847
Out[73]: 0.41197431610069063
In [74]: TBV = a[posAi]
         mean(TBV)
Out[74]: 13.29354
In [75]: EBV = aHat3[posAi]
         mean(EBV)
Out[75]: 3.219089748898632
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 = 0.402
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 0.958
Out[76]: 0.4015215812384653
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]: 12.791862625000002
In [80]: EBV = aHat3[posAi]
         mean(EBV)
Out[80]: 2.6845823824940918
```

```
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 = 0.683
         SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 0.872
Out[81]: 0.682806542557785
In [82]: TBV = a[posAi]
         mean(TBV)
Out[82]: 10.15954487179487
In [83]: EBV = aHat3[posAi]
         mean(EBV)
Out[83]: -0.03201939474148579
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 = 0.753
         SSBRJC from Gibbs - G1.noGenotype.ID: regression of TBV on GEBV = 0.982
Out[84]: 0.7534989102649137
In [85]: TBV = a[posAi]
         mean(TBV)
Out[85]: 10.70649371794872
In [86]: EBV = aHat3[posAi]
         mean(EBV)
Out[86]: 0.5660227910486897
```

```
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 = 1.009
Out[87]: 0.7556524101668907
In [88]: TBV = a[posAi]
         mean(TBV)
Out[88]: 11.252934358974361
In [89]: EBV = aHat3[posAi]
         mean(EBV)
Out[89]: 1.0900582062616901
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 = 0.736
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 1.004
Out[90]: 0.736167290567994
In [91]: TBV = a[posAi]
         mean(TBV)
Out[91]: 11.763757564102566
In [92]: EBV = aHat3[posAi]
         mean(EBV)
Out[92]: 1.5937317326504032
```

```
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 = 0.714
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 0.981
Out[93]: 0.7140784999955377
In [94]: TBV = a[posAi]
         mean(TBV)
Out[94]: 12.277072564102564
In [95]: EBV = aHat3[posAi]
         mean(EBV)
Out[95]: 2.1242636821280843
In [96]: numSSBayes
Out[96]: SSBR.NumSSBayes(54913,45913,9000,40000,39000,1000,0)
```