```
In [1]: include("/home/nicole/Jupyter/SSBRnoJ/src/SSBR.jl")
        using SSBR
In [2]:
        function getPos(ped,IDs)
            posAi = Array(Int64, size(IDs, 1))
            for (i,id) = enumerate(IDs[:,1])
                 posAi[i] = ped.idMap[id].seqID
            end
            return posAi
        end
Out[2]: getPos (generic function with 1 method)
In [3]: | ; cd Data/0.5/G/1
        /home/nicole/Jupyter/JG3/Data/0.5/G/1
In [4]:
        ;ls
        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.JC.txt
        Regression.G5.G.N.txt
        all.ID
        genotype.ID
        noGenotype.ID
        sim.bv
        sim.phenotype
In [5]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
```

```
;awk '{print $1}' GenNF.txt | sort -b > genotype.ID
         ;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
         ; 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]:
In [13]:
         ; 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
         ; awk 'NR >=40001 && NR <=48000 {print $1}' PedAll.txt | sort -b > G5.ID
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
In [20]:
         ; join G4.ID genotype.ID > G4.Genotype.ID
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
         ;join -v1 G3.ID genotype.ID > G3.noGenotype.ID
In [25]:
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 1200 GO.Genotype.ID
          200 200 1200 G1.Genotype.ID
          200 200 1200 G2.Genotype.ID
          200 200 1200 G3.Genotype.ID
          200 200 1200 G4.Genotype.ID
          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", calculateInbreeding
         y Vecs = make_yVecs("sim.phenotype",ped,numSSBayes)
         Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
         nothing
         vRes
                = 0.668
In [31]:
                = 0.668
         vG
         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 = 0.884
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.964
Out[34]: 0.8838331180533318
In [35]: | TBV = a[posAi]
         mean(TBV)
Out[35]: 11.392354354166669
```

```
In [36]: EBV = aHat3[posAi]
         mean(EBV)
Out[36]: 1.3182794726167257
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 = 0.575
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 0.992
Out[37]: 0.5754781277159947
In [38]: TBV = a[posAi]
         mean(TBV)
Out[38]: 12.5253611111111112
In [39]: EBV = aHat3[posAi]
         mean(EBV)
Out[39]: 2.5139946576147834
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 = 0.870
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.970
Out[40]: 0.8704208408664355
In [41]: TBV = a[posAi]
         mean(TBV)
Out[41]: 11.130891256410255
In [42]: EBV = aHat3[posAi]
         mean(EBV)
Out[42]: 1.0423451991556356
```

```
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 = 0.707
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.861
Out[43]: 0.7069254403314218
In [44]: TBV = a[posAi]
         mean(TBV)
Out[44]: 10.1364735
In [45]: EBV = aHat3[posAi]
         mean(EBV)
Out[45]: 0.00667871693655719
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 = 0.771
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.994
Out[46]: 0.7714283009348639
In [47]: TBV = a[posAi]
         mean(TBV)
Out[47]: 10.634187875
In [48]: EBV = aHat3[posAi]
         mean(EBV)
Out[48]: 0.5727716253276725
```

```
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 = 0.759
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 0.982
Out[49]: 0.7591283509671712
In [50]: TBV = a[posAi]
         mean(TBV)
Out[50]: 11.175617749999997
In [51]: EBV = aHat3[posAi]
         mean(EBV)
Out[51]: 1.0978713576250336
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 = 0.758
         SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = 0.973
Out[52]: 0.7584791728438517
In [53]: TBV = a[posAi]
         mean(TBV)
Out[53]: 11.683383124999999
In [54]: EBV = aHat3[posAi]
         mean(EBV)
Out[54]: 1.5934970402644184
```

```
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 = 0.733
         SSBRJC from Gibbs - G4.ID: regression of TBV on GEBV = 0.956
Out[55]: 0.7326312378767961
In [56]: TBV = a[posAi]
         mean(TBV)
Out[56]: 12.150585625
In [57]: EBV = aHat3[posAi]
         mean(EBV)
Out[57]: 2.073365392763034
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 = 0.436
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 1.002
Out[58]: 0.4360452758554354
In [59]: TBV = a[posAi]
         mean(TBV)
Out[59]: 12.573878249999998
In [60]: EBV = aHat3[posAi]
         mean(EBV)
Out[60]: 2.5654927027836396
```

```
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 = 0.909
         SSBRJC from Gibbs - G0.Genotype.ID: regression of TBV on GEBV = 1.096
Out[61]: 0.9086743275262619
In [62]: TBV = a[posAi]
         mean(TBV)
Out[62]: 11.136659999999997
In [63]: EBV = aHat3[posAi]
         mean(EBV)
Out[63]: 1.1611595458653432
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 = 0.883
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.010
Out[64]: 0.8827726756286254
In [65]: TBV = a[posAi]
         mean(TBV)
Out[65]: 11.721100000000002
In [66]: EBV = aHat3[posAi]
         mean(EBV)
Out[66]: 1.6284937073166537
```

```
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 = 0.898
         SSBRJC from Gibbs - G2.Genotype.ID: regression of TBV on GEBV = 0.992
Out[67]: 0.8975419887614531
In [68]: TBV = a[posAi]
         mean(TBV)
Out[68]: 12.206114999999997
In [69]: EBV = aHat3[posAi]
         mean(EBV)
Out[69]: 2.097290252857807
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 = 0.877
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 0.992
Out[70]: 0.8770321644543825
In [71]: TBV = a[posAi]
         mean(TBV)
Out[71]: 12.621465
In [72]: EBV = aHat3[posAi]
         mean(EBV)
Out[72]: 2.5654025293163603
```

```
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 = 0.516
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.023
Out[73]: 0.5158591313364875
In [74]: TBV = a[posAi]
         mean(TBV)
Out[74]: 13.000779999999999
In [75]: EBV = aHat3[posAi]
         mean(EBV)
Out[75]: 3.0577054459634985
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 = 0.436
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 1.002
Out[76]: 0.4360452758554354
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]: 12.573878249999998
In [80]: EBV = aHat3[posAi]
         mean(EBV)
Out[80]: 2.5654927027836396
```

```
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 = 0.689
         SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 0.856
Out[81]: 0.6885101590290259
In [82]: TBV = a[posAi]
         mean(TBV)
Out[82]: 10.110827692307693
In [83]: EBV = aHat3[posAi]
         mean(EBV)
Out[83]: -0.022923355600078354
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 = 0.755
         SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 0.990
Out[84]: 0.7553504232587579
In [85]: TBV = a[posAi]
         mean(TBV)
Out[85]: 10.60631833333333
In [86]: EBV = aHat3[posAi]
         mean(EBV)
Out[86]: 0.5457018283535962
```

```
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 = 0.742
         SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = 0.977
Out[87]: 0.7417589211237546
In [88]: TBV = a[posAi]
         mean(TBV)
Out[88]: 11.14919474358974
In [89]: EBV = aHat3[posAi]
         mean(EBV)
Out[89]: 1.0722452321062448
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 = 0.743
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 0.973
Out[90]: 0.7431775704666279
In [91]: TBV = a[posAi]
         mean(TBV)
Out[91]: 11.659329743589744
In [92]: EBV = aHat3[posAi]
         mean(EBV)
Out[92]: 1.568576386698984
```

```
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 = 0.723
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 0.964
Out[93]: 0.7231875418801711
In [94]: TBV = a[posAi]
         mean(TBV)
Out[94]: 12.128785769230769
In [95]: EBV = aHat3[posAi]
         mean(EBV)
Out[95]: 2.048125904219432
In [96]: numSSBayes
Out[96]: SSBR.NumSSBayes(54906,45906,9000,40000,39000,1000,0)
```