```
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/0.5a0/G/9
         /home/nicole/Jupyter/JG3/Data/0.5a0/G/9
 In [4]:
         ;ls
         GenNF.txt
         PedAll.txt
         Phe.txt
         PheAll.txt
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
         ;awk '{print $1,$3}' PheAll.txt > sim.bv
In [9]:
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
In [12]:
         ; awk 'NR >=24001 && NR <=32000 {print $1}' PedAll.txt | sort -b > G3.ID
In [13]:
In [14]: ; awk 'NR >=32001 && NR <=40000 {print $1}' PedAll.txt | sort -b > G4.ID
In [15]:
         ; 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]:
         ; join G2.ID genotype.ID > G2.Genotype.ID
In [18]:
In [19]:
         ; join G3.ID genotype.ID > G3.Genotype.ID
In [20]:
         ; join G4.ID genotype.ID > G4.Genotype.ID
         ; join G5.ID genotype.ID > G5.Genotype.ID
In [21]:
In [22]:
         ;join -v1 G0.ID genotype.ID > G0.noGenotype.ID
         ;join -v1 G1.ID genotype.ID > G1.noGenotype.ID
In [23]:
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
          200 200 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
                7800 46800 GO.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
                = 0.594
         vG
                = 0.594
         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 = 0.898
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.996
Out[34]: 0.8984477308647231
In [35]: TBV = a[posAi]
         mean(TBV)
Out[35]: 2.6163576041666667
In [36]: EBV = aHat3[posAi]
         mean(EBV)
Out[36]: 1.3570703305516008
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 = 0.607
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.043
Out[37]: 0.6072294415193861
In [38]: TBV = a[posAi]
         mean(TBV)
Out[38]: 3.899156333333334
In [39]: | EBV = aHat3[posAi]
         mean(EBV)
Out[39]: 2.635894727337599
```

```
In [40]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',head
         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 = 0.884
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.990
Out[40]: 0.883601340665855
In [41]: TBV = a[posAi]
         mean(TBV)
Out[41]: 2.320327128205128
In [42]: EBV = aHat3[posAi]
         mean(EBV)
Out[42]: 1.0619570082163703
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 ep
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.687
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.822
Out[43]: 0.6873740401533486
In [44]: TBV = a[posAi]
         mean(TBV)
Out[44]: 1.2755743750000004
In [45]: EBV = aHat3[posAi]
         mean(EBV)
Out[45]: 0.002907847128903057
```

```
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 ei
         @printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G1.ID : correlation = 0.779
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 1.003
Out[46]: 0.7793954341584852
In [47]: TBV = a[posAi]
         mean(TBV)
Out[47]: 1.7703276250000002
In [48]: EBV = aHat3[posAi]
         mean(EBV)
Out[48]: 0.5350383587931373
In [49]: IDs = readtable("G2.ID", eltypes =[UTF8String], separator = ' ', header=false
         posAi = getPos(ped,IDs)
         cor5 = cor(a[posAi],aHat3[posAi])[1,1]
         req5 = linreg(aHat3[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G2.ID : correlation = %6.3f\n", cor5 ) # with ep
         @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor5
         SSBRJC from Gibbs - G2.ID : correlation = 0.763
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 1.002
Out[49]: 0.7632421477240279
In [50]: TBV = a[posAi]
         mean(TBV)
Out[50]: 2.344312625
In [51]: EBV = aHat3[posAi]
         mean(EBV)
Out[51]: 1.0905299732112568
```

```
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 ei
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor6
         SSBRJC from Gibbs - G3.ID : correlation = 0.786
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 1.047
Out[52]: 0.7863107875097429
In [53]: TBV = a[posAi]
         mean(TBV)
Out[53]: 2.8908736250000002
In [54]: EBV = aHat3[posAi]
         mean(EBV)
Out[54]: 1.6303954304396329
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 ep
         @printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor7
         SSBRJC from Gibbs - G4.ID : correlation = 0.753
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 1.017
Out[55]: 0.7534480467378328
In [56]: TBV = a[posAi]
         mean(TBV)
Out[56]: 3.45945825
In [57]: EBV = aHat3[posAi]
         mean(EBV)
Out[57]: 2.18962793612043
```

```
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 ei
         @printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor8
         SSBRJC from Gibbs - G5.ID : correlation = 0.458
         SSBRJC from Gibbs - G5.ID: regression of TBV on GEBV = 1.079
Out[58]: 0.4582692145364429
In [59]: TBV = a[posAi]
         mean(TBV)
Out[59]: 3.957599125
In [60]: EBV = aHat3[posAi]
         mean(EBV)
Out[60]: 2.6939224376162447
In [61]: IDs = readtable("G0.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 - 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 = 0.912
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 1.009
Out[61]: 0.9116671895817529
In [62]: TBV = a[posAi]
         mean(TBV)
Out[62]: 2.31163
In [63]: EBV = aHat3[posAi]
         mean(EBV)
Out[63]: 1.100025684841908
```

```
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 = 0.905
         SSBRJC from Gibbs - G1.Genotype.ID: regression of TBV on GEBV = 1.013
Out[64]: 0.9045554051164307
In [65]: TBV = a[posAi]
         mean(TBV)
Out[65]: 2.9020149999999996
In [66]: EBV = aHat3[posAi]
         mean(EBV)
Out[66]: 1.6388041929940314
In [67]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         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 = 0.914
         SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = 1.019
Out[67]: 0.9142290546323221
In [68]: TBV = a[posAi]
         mean(TBV)
Out[68]: 3.4562049999999997
In [69]: EBV = aHat3[posAi]
         mean(EBV)
Out[69]: 2.172720628686574
```

```
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 = 0.908
         SSBRJC from Gibbs - G3.Genotype.ID: regression of TBV on GEBV = 0.990
Out[70]: 0.9078760386021065
In [71]: TBV = a[posAi]
         mean(TBV)
Out[71]: 4.02339
In [72]: EBV = aHat3[posAi]
         mean(EBV)
Out[72]: 2.7388924733031565
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 = 0.581
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.179
Out[73]: 0.5814601805966787
In [74]: TBV = a[posAi]
         mean(TBV)
Out[74]: 4.46483
In [75]: EBV = aHat3[posAi]
         mean(EBV)
Out[75]: 3.207922245716489
```

```
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 = 0.458
         SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = 1.079
Out[76]: 0.4582692145364429
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]: 3.957599125
In [80]: EBV = aHat3[posAi]
         mean(EBV)
Out[80]: 2.6939224376162447
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 = 0.665
         SSBRJC from Gibbs - G0.noGenotype.ID: regression of TBV on GEBV = 0.808
Out[81]: 0.6645884311022129
In [82]: TBV = a[posAi]
         mean(TBV)
Out[82]: 1.2490088461538464
In [83]: | EBV = aHat3[posAi]
         mean(EBV)
Out[83]: -0.025223379479122714
```

```
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 = 0.762
         SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 1.000
Out[84]: 0.7618142345506285
In [85]: TBV = a[posAi]
         mean(TBV)
Out[85]: 1.7413100000000001
In [86]: EBV = aHat3[posAi]
         mean(EBV)
Out[86]: 0.5067366707367043
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 = 0.743
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 0.998
Out[87]: 0.7428957205477477
In [88]: TBV = a[posAi]
         mean(TBV)
Out[88]: 2.3158025641025644
In [89]: EBV = aHat3[posAi]
         mean(EBV)
Out[89]: 1.062781494865736
```

```
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 = 0.768
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 1.051
Out[90]: 0.7684601012996081
In [91]: TBV = a[posAi]
         mean(TBV)
Out[91]: 2.861834743589743
In [92]: EBV = aHat3[posAi]
         mean(EBV)
Out[92]: 1.6019724293405684
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 = 0.741
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 1.019
Out[93]: 0.7411429498692089
In [94]: TBV = a[posAi]
         mean(TBV)
Out[94]: 3.4336794871794876
In [95]: EBV = aHat3[posAi]
         mean(EBV)
Out[95]: 2.163517825617967
In [96]: numSSBayes
Out[96]: SSBR.NumSSBayes(54952,45952,9000,40000,39000,1000,0)
```