```
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/Q/4
        /home/nicole/Jupyter/JG3/Data/0.5/Q/4
In [4]:
        ;ls
        Correlation.G5.Q.JC.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
        PedAll.txt
        Phe.txt
        PheAll.txt
        QTLNF.txt
        Regression.G5.Q.JC.txt
        all.ID
        genotype.ID
        noGenotype.ID
        sim.bv
        sim.phenotype
In [5]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
In [6]: | ;awk '{print $1}' QTLNF.txt | sort -b > genotype.ID
```

```
;join -v1 all.ID genotype.ID > noGenotype.ID
         ;awk '{print $1,$2}' Phe.txt > sim.phenotype
         ;awk '{print $1,$3}' PheAll.txt > sim.bv
 In [9]:
         ; 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 [12]:
         ; awk 'NR >=16001 && NR <=24000 {print $1}' PedAll.txt | sort -b > G2.ID
In [13]:
         ; 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 [16]:
         ; join G0.ID genotype.ID > G0.Genotype.ID
In [17]:
         ; join G1.ID genotype.ID > G1.Genotype.ID
         ; 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
         ; join -v1 G4.ID genotype.ID > G4.noGenotype.ID
In [26]:
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
         ped,A Mats,numSSBayes = calc Ai("PedAll.txt", "genotype.ID", calculateInbreedia
In [30]:
         nothing
         df
                = read_genotypes("QTLNF.txt",numSSBayes)
         M Mats = make MMats(df, A Mats, ped, center=true);
                                                                                  # with
         y Vecs = make yVecs("sim.phenotype",ped,numSSBayes);
         Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
         X_Mats, W_Mats = make_XWMats(Z_Mats,M_Mats,numSSBayes)
                                                                                  # no
         nothing
In [31]:
         vRes
                = 0.759
                = 0.759
         vG
         nIter = 50000
         @time aHat1,alphaHat,betaHat,epsiHat =
         ssGibbs(M_Mats,y_Vecs,Z_Mats,X_Mats,W_Mats,A_Mats, numSSBayes,vRes,vG,nIter,
         nothing
         This is iteration 5000
         This is iteration 10000
         This is iteration 15000
         This is iteration 20000
         This is iteration 25000
         This is iteration 30000
         This is iteration 35000
         This is iteration 40000
         This is iteration 45000
         This is iteration 50000
         3359.054100 seconds (22.99 G allocations: 722.840 GB, 8.46% gc time)
In [32]: betaHat
Out[32]: 1-element Array{Float64,1}:
          10.2874
In [33]: using DataFrames
```

```
In [34]: 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 [35]: | IDs = readtable("all.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor1 = cor(a[posAi],aHat1[posAi])[1,1]
         reg1 = linreg(aHat1[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 = 0.919
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 1.001
Out[35]: 0.9189174249672409
In [36]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[36]: 0.1680225983052738
In [37]: IDs = readtable("genotype.ID", eltypes =[UTF8String], separator = ' ',header:
         posAi = getPos(ped,IDs)
         cor2 = cor(a[posAi], aHat1[posAi])[1,1]
         reg2 = linreg(aHat1[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 = 0.952
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 1.164
Out[37]: 0.9515644077958035
In [38]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[38]: 1.4682085668852523
In [39]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor3 = cor(a[posAi],aHat1[posAi])[1,1]
         reg3 = linreg(aHat1[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 = 0.883
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.969
Out[39]: 0.8832577268718352
```

```
In [40]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[40]: -0.132020317520875
In [41]: IDs = readtable("G0.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor4 = cor(a[posAi],aHat1[posAi])[1,1]
         reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G0.ID : correlation = %6.3f\n", cor4 ) # with e;
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.721
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.884
Out[41]: 0.7206481942572285
In [42]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[42]: -1.3940629521383343
In [43]: IDs = readtable("G1.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor4 = cor(a[posAi],aHat1[posAi])[1,1]
         req4 = linreg(aHat1[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 = 0.778
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 1.009
Out[43]: 0.7782358738241512
In [44]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[44]: -0.6425720285204569
In [45]: IDs = readtable("G2.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor5 = cor(a[posAi],aHat1[posAi])[1,1]
         reg5 = linreg(aHat1[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 = 0.775
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 1.041
Out[45]: 0.7746248384308367
```

```
In [46]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[46]: -0.0477593370031644
In [47]: IDs = readtable("G3.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor6 = cor(a[posAi],aHat1[posAi])[1,1]
         reg6 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G3.ID : correlation = %6.3f\n", cor6 ) # with e;
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor6
         SSBRJC from Gibbs - G3.ID : correlation = 0.770
         SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = 1.033
Out[47]: 0.770036931407005
In [48]: GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[48]: 0.5153903463490601
In [49]: IDs = readtable("G4.ID", eltypes =[UTF8String], separator = ' ', header=false
         posAi = getPos(ped,IDs)
         cor7 = cor(a[posAi],aHat1[posAi])[1,1]
         req7 = linreg(aHat1[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 = 0.769
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 1.058
Out[49]: 0.7688909418967874
In [50]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[50]: 1.060262933487035
In [51]: IDs = readtable("G5.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor8 = cor(a[posAi],aHat1[posAi])[1,1]
         reg8 = linreg(aHat1[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 = 0.944
         SSBRJC from Gibbs - G5.ID: regression of TBV on GEBV = 1.151
Out[51]: 0.9439878164773186
```

```
In [52]: writedlm("Correlation.G5.Q.C.txt",cor13)
         LoadError: UndefVarError: cor13 not defined
         while loading In[52], in expression starting on line 1
In [53]: writedlm("Regression.G5.Q.C.txt",reg13)
         LoadError: UndefVarError: reg13 not defined
         while loading In[53], in expression starting on line 1
In [54]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[54]: 1.5168766276575034
In [55]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat1[posAi])[1,1]
         reg9 = linreg(aHat1[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 = 0.960
         SSBRJC from Gibbs - G0.Genotype.ID: regression of TBV on GEBV = 1.125
Out[55]: 0.9598273295756209
In [56]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[56]: 0.18015627220327965
In [57]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ', head
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat1[posAi])[1,1]
         reg9 = linreg(aHat1[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 =
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 1.090
Out[57]: 0.9345478115679285
In [58]: | GEBV = aHat1[posAi]
         mean(GEBV)
Out[58]: 0.61216050312711
```

```
In [59]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor10 = cor(a[posAi],aHat1[posAi])[1,1]
         reg10 = linreg(aHat1[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 = 0.931
         SSBRJC from Gibbs - G2.Genotype.ID: regression of TBV on GEBV = 1.067
Out[59]: 0.9309533119758823
In [60]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[60]: 1.1067560508201622
In [61]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ', head
         posAi = getPos(ped,IDs)
         cor11 = cor(a[posAi],aHat1[posAi])[1,1]
         reg11 = linreg(aHat1[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 = 0.921
         SSBRJC from Gibbs - G3.Genotype.ID: regression of TBV on GEBV = 1.054
Out[61]: 0.9210215779582085
In [62]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[62]: 1.5764889371168627
In [63]: IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor12 = cor(a[posAi],aHat1[posAi])[1,1]
         reg12 = linreg(aHat1[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 =
         SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = 1.096
Out[63]: 0.9163662496507724
In [64]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[64]: 1.9187586402687966
```

```
In [65]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor13 = cor(a[posAi],aHat1[posAi])[1,1]
         reg13 = linreg(aHat1[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 = 0.944
         SSBRJC from Gibbs - G5.Genotype.ID: regression of TBV on GEBV = 1.151
Out[65]: 0.9439878164773186
In [66]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
Out[66]: 1.5168766276575034
In [67]: IDs = readtable("G0.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat1[posAi])[1,1]
         reg14 = linreg(aHat1[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 = 0.701
         SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 0.900
Out[67]: 0.7008487460406642
In [68]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[68]: -1.4344275476342732
In [69]: IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat1[posAi])[1,1]
         reg14 = linreg(aHat1[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 = 0.763
         SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = 1.020
Out[69]: 0.7633140272864326
In [70]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[70]: -0.674744657537061
```

```
In [71]: IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor15 = cor(a[posAi],aHat1[posAi])[1,1]
         reg15 = linreg(aHat1[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 = 0.758
         SSBRJC from Gibbs - G2.noGenotype.ID: regression of TBV on GEBV = 1.043
Out[71]: 0.7576138625800222
In [72]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[72]: -0.07736229566530096
In [73]: IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor16 = cor(a[posAi],aHat1[posAi])[1,1]
         reg16 = linreg(aHat1[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 = 0.752
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 1.027
Out[73]: 0.7523803077616238
In [74]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[74]: 0.4881826901755267
In [75]: IDs = readtable("G4.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor17 = cor(a[posAi],aHat1[posAi])[1,1]
         reg17 = linreg(aHat1[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 = 0.754
         SSBRJC from Gibbs - G4.noGenotype.ID : regression of TBV on GEBV = 1.047
Out[75]: 0.7535217952866002
In [76]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[76]: 1.0382502230567334
```

In [77]: numSSBayes

Out[77]: SSBR.NumSSBayes(54918,45918,9000,40000,39000,1000,50)