The PLINK example GWAS analysed by PLINK and Sib-pair

David Duffy

Genetic Epidemiology Laboratory

Introduction

- Overview of development of Sib-pair
- PLINK v. Sib-pair

Overview of Sib-pair

An extensible platform for genetic data manipulation and analysis

A platform for methodological experimentation

First code written in 1995. Now all standard Fortran 95, compiles using multiple compilers.

Creeping featurism has continued to today (55000 lines of code + 7000 lines of comments; 9000 LOC in last 12 months)

The Language

- Simple interpreted language, over 200 commands
- Commands for linkage, association, variance components …
- Offers the usual record-wise operations on data – algebra, logical conditions
- Family-centric data operations – subsetting, pruning etc
- Some elementary databasing type operations – merging, editing
- Flexible data export and scripting to use other programs

Using Sib-pair to analyse a GWAS

I have spent a bit of time over the last 1-2 years doing some optimization of the code so that it is not too onerous to use Sib-pair in the analysis of large datasets.

SNP genotype data can be stored internally as 4 bits per genotype, so that large datasets fit into memory. Even the default format for storing genotype is now 4 times smaller than it was.

A binary image of a dataset can be saved and reread from disk. This is much quicker than reading in the original locus and pedigree files. The image is compressed (*gzip*).

The **summary** command allows one to rank and subset out only the results of interest from a large set of tests. This can also generate a Postscript plot, or a .*WIG* file for the UCSC browser. The **keep** and **drop** commands allow one to select loci based on Hardy-Weinberg disequilibrium or allele frequencies.

Sib-pair compared to PLINK: Making a binary file

Creating a binary file for subsequent analysis:

PLINK (35 seconds):

```
plink –noweb –file wgas1 –make-bed –out wgas2
```

Sib-pair (2 minutes 28 seconds):

```
read loc plink wgas1.map
read ped wgas1.ped
set che off
set imp -1
set ple -1
run
write bin wgas1.bin compress
```

The space taken by the resulting files:

```
-rw-r–r– 1 davidD davidD 2.2k Oct 24 13:46 wgas2.fam
-rw-r–r– 1 davidD davidD 7.8M Oct 24 13:46 wgas2.bim
-rw-r–r– 1 davidD davidD 5.3M Oct 24 13:46 wgas2.bed
-rw-r–r– 1 davidD davidD 9.5M Oct 24 13:50 wgas1.bin.gz
```

Sib-pair compared to PLINK: Allele frequencies

Estimating allele frequencies:

PLINK (9 seconds):

```
plink –noweb –bfile wgas2 –freq –out freq1
```

```
 CHR          SNP   A1   A2          MAF  NCHROBS
   1    rs3094315    G    A       0.1236      178
   1    rs6672353    A    G     0.005618      178
   1    rs4040617    G    A       0.1167      180
   1    rs2905036    0    T            0      180
   1    rs4245756    0    C            0      180
   1    rs4075116    C    T      0.05556      180
   1    rs9442385    T    G       0.3933      178
   1    rs6603781    0    G            0      178
…
```

Sib-pair compared to PLINK: Allele frequencies

Estimating allele frequencies:

Sib-pair (14 seconds):

```
read bin wgas1.bin; fre snp
```

OR (15 seconds):

```
read plink wgas2; fre snp
```

```
Marker     NAll  Allele(s)  Freq   Het    Ntyped

rs3094315     2     G (A)   0.1236 0.2179     89 792429 (chr 1)
rs6672353     2     A (G)   0.0056 0.0112     89 817376 (chr 1)
rs4040617     2     G (A)   0.1167 0.2073     90 819185 (chr 1)
rs2905036     1     T       1.0000  -         90 832343 (chr 1)
rs4245756     1     C       1.0000  -         90 839326 (chr 1)
rs4075116     2     C (T)   0.0556 0.1055     90 1043552 (chr 1)
rs9442385     2     T (G)   0.3933 0.4799     89 1137258 (chr 1)
rs6603781     1     G       1.0000  -         89 1198554 (chr 1)
…
```

Sib-pair compared to PLINK: HWE

Testing Hardy-Weinberg equilibrium

PLINK (22 seconds):

```
plink –noweb –bfile wgas2 –hardy –out hwe1
```

```
CHR   SNP         TEST   A1   A2      GENO   O(HET)   E(HET)       P
  1   rs3094315  G    G      ALL   0/22/67   0.2472   0.2166  0.3476
  1   rs3094315  G    G      AFF   0/15/33   0.3125   0.2637  0.5771
  1   rs3094315  G    G    UNAFF    0/7/34   0.1707   0.1562       1
  1   rs6672353  A    A      ALL    0/1/88  0.01124  0.01117       1
  1   rs6672353  A    A      AFF    0/1/48  0.02041   0.0202       1
  1   rs6672353  A    A    UNAFF    0/0/40        0        0       1
  1   rs4040617  G    G      ALL   0/21/69   0.2333   0.2061  0.5994
  1   rs4040617  G    G      AFF   0/14/35   0.2857   0.2449  0.5714
  1   rs4040617  G    G    UNAFF    0/7/34   0.1707   0.1562       1
  1   rs2905036  0    0      ALL    0/0/90        0        0       1


…
```

Sib-pair compared to PLINK: HWE

Testing Hardy-Weinberg equilibrium. Note that Sib-pair calculates two tests of HWE for each SNP (Chi-square and exact test), but only prints the exact P-value here. The usual Sib-pair HWE Chi-square test uses founders and nonfounders, and gene-drops a correct P-value.

Sib-pair (56 seconds):

```
read bin wgas1.bin
set iter 0
hwe
select trait
hwe
unselect
select not trait
hwe
```

```
Marker     Typed  Genos  Chi-square Asy P  Emp P  Iters
rs3094315      89      3        3.1 0.3476 1.0000     0 HWE .
rs6672353      89      3        0.0 1.0000 1.0000     0 HWE .
rs4040617      90      3        2.8 0.5994 1.0000     0 HWE .
rs2905036      90      1        0.0 1.0000 1.0000     0 HWE .
rs4245756      90      1        0.0 1.0000 1.0000     0 HWE .
rs4075116      90      3        0.6 1.0000 1.0000     0 HWE .
rs9442385      89      3        2.1 0.1815 1.0000     0 HWE .
rs6603781      89      1        0.0 1.0000 1.0000     0 HWE .
rs11260562     88      3        0.1 1.0000 1.0000     0 HWE .

…
```

Sib-pair compared to PLINK: filtering

Filtering individuals and markers:

PLINK (16 seconds):

```
plink –bfile wgas2 –maf 0.01 –geno 0.05 –mind 0.05 –hwe 1e-3 –make-bed
–out wgas3
```

```
Before frequency and genotyping pruning, there are 228694 SNPs
90 founders and 0 non-founders found
Writing list of removed individuals to [ wgas3.irem ]
1 of 90 individuals removed for low genotyping ( MIND > 0.05 )
74 markers to be excluded based on HWE test ( p <= 0.001 )
        65 markers failed HWE test in cases
        74 markers failed HWE test in controls
Total genotyping rate in remaining individuals is 0.995473
2728 SNPs failed missingness test ( GENO > 0.05 )
46834 SNPs failed frequency test ( MAF < 0.01 )
After frequency and genotyping pruning, there are 179493 SNPs
After filtering, 48 cases, 41 controls and 0 missing
After filtering, 44 males, 45 females, and 0 of unspecified sex
```

Sib-pair compared to PLINK: filtering

Filtering individuals and markers. In Sib-pair, **select** and **unselect** are for individuals, and **keep** and **drop** are for loci. The PLINK filters are applied independently of each other, while the Sib-pair filtering acts sequentially.

Sib-pair (35 seconds):

```
read bin wgas1.bin
select where numtyp <= 217259
show ped
unselect
select where numtyp > 217259
select not trait
set iter 0
drop where hwe 0.001
unselect
drop where max 0.99
keep where num 85
write bin wgas3.bin compress
```

```
Permanently deleted 48459 loci.

Reread   89 pedigrees, 89 individuals (5.06 s).
Dataset occupies 64.170 Mb.
```

Sib-pair compared to PLINK: association

Simple association analysis

PLINK (13 seconds):

```
plink –noweb –bfile wgas3 –assoc –adjust –out assoc1
...
Writing main association results to [ assoc1.assoc ]
Computing corrected significance values (FDR, Sidak, etc)
Genomic inflation factor (based on median chi-squared) is 1.25937
Mean chi-squared statistic is 1.22972
Correcting for 179493 tests
Writing multiple-test corrected significance values to [
assoc1.assoc.adjusted
]
```

OR (55 seconds)

```
plink –noweb –bfile wgas3 –logistic logistic1
```

```
CHR         SNP         BP   A1       F_A       F_U   A2     CHISQ         P        OR
   1   rs3094315    792429    G    0.1489   0.08537    A     1.684    0.1944     1.875
   1   rs4040617    819185    G    0.1354   0.08537    A     1.111    0.2919     1.678
   1   rs4075116   1043552    C   0.04167   0.07317    T    0.8278    0.3629    0.5507
   1   rs9442385   1137258    T    0.3723    0.4268    G    0.5428    0.4613    0.7966
   1  rs11260562   1205233    A   0.02174   0.03659    G    0.3424    0.5585    0.5852
   1   rs6685064   1251215    C    0.3854     0.439    T    0.5253    0.4686    0.8013
...
```

```
CHR        SNP        BP  A1  TEST  NMISS    ODDS     STAT       P
  1  rs3094315    792429   G   ADD     88   2.061    1.381  0.1672
  1  rs4040617    819185   G   ADD     89   1.804     1.12  0.2629
  1  rs4075116   1043552   C   ADD     89  0.5303  -0.9272  0.3538
  1  rs9442385   1137258   T   ADD     88  0.8197   -0.687  0.4921
  1 rs11260562   1205233   A   ADD     87  0.5758  -0.5877  0.5567
  1  rs6685064   1251215   C   ADD     89  0.8409  -0.6398  0.5223
...
```

Sib-pair compared to PLINK: association

Simple association analysis

Sib-pair (35 seconds):

```
read bin wgas3.bin
set iter 0
ass trait
summary 20
```

OR (2 minutes 35 seconds):

```
read bin wgas3.bin
set iter 0
ass trait snp
summary 20
```

```
Marker      Typed  Allels Chi-square Asy P  Emp P  Iters
----------  ------ ------ ---------- ------ ------ ------
rs3094315       88      2        1.7 0.1944 1.0000      0 AssX2-HWE .
rs4040617       89      2        1.1 0.2919 1.0000      0 AssX2-HWE .
rs4075116       89      2        0.8 0.3629 1.0000      0 AssX2-HWE .
rs9442385       88      2        0.5 0.4613 1.0000      0 AssX2-HWE .
rs11260562      87      2        0.3 0.5585 1.0000      0 AssX2-HWE .
rs6685064       89      2        0.5 0.4686 1.0000      0 AssX2-HWE .
...
```

```
Total number of tests = 180235

Locus      Position   P-value  -log10(P)
---------- ---------- ------- ----------
rs2513514       75.92  0.0000      6.329  75922141 (chr 11)
rs6110115       13.91  0.0000      6.149  13911728 (chr 20)
rs2508756       75.92  0.0000      5.677  75921549 (chr 11)
rs16976702      54.12  0.0000      5.661  54120691 (chr 15)
rs11204005      12.90  0.0000      5.103  12895576 (chr 8)
rs16910850      94.48  0.0000      4.915  94478347 (chr 9)
rs1195747      129.97  0.0000      4.846  129970575 (chr 12)
rs7207095       77.93  0.0000      4.774  77933018 (chr 17)
rs16971118      77.67  0.0000      4.720  77672467 (chr 15)
rs6074704       14.12  0.0000      4.696  14115283 (chr 20)
rs1570484       14.14  0.0000      4.696  14139687 (chr 20)
rs9944528       77.89  0.0000      4.664  77894039 (chr 17)
rs636006        32.43  0.0000      4.642  32426349 (chr 3)
...
```

```
Marker         OR       95% CI              P-value
----------- -------- -------------------- ----------
rs3094315      2.061    0.739 –    5.749 0.8361E-01 792429 (chr 1)
rs4040617      1.804    0.642 –    5.068 0.1314     819185 (chr 1)
rs4075116      1.886    0.493 –    7.207 0.1769     1043552 (chr 1)
rs9442385      1.220    0.692 –    2.151 0.2460     1137258 (chr 1)
rs11260562     1.737    0.276 –   10.948 0.2784     1205233 (chr 1)
rs6685064      1.189    0.699 –    2.022 0.2612     1251215 (chr 1)
...
```

```
Total number of tests = 180235

Locus       Position   P-value  -log10(P)
----------- ---------- ------- ----------
rs1548299         3.64  0.0000      5.263  3640174 (chr 9)
rs2513514        75.92  0.0000      5.187  75922141 (chr 11)
rs6110115        13.91  0.0000      4.841  13911728 (chr 20)
rs2508756        75.92  0.0000      4.753  75921549 (chr 11)
rs16976702       54.12  0.0000      4.711  54120691 (chr 15)
rs11204005       12.90  0.0000      4.474  12895576 (chr 8)
rs9302779         3.87  0.0000      4.387  3873346 (chr 16)
rs17534370       70.30  0.0000      4.341  70297172 (chr 9)
rs11781505      142.00  0.0001      4.266  142002911 (chr 8)
rs11785430      142.08  0.0001      4.266  142078709 (chr 8)
rs1195747       129.97  0.0001      4.146  129970575 (chr 12)
rs16910850       94.48  0.0001      4.126  94478347 (chr 9)
rs1570484        14.14  0.0001      4.124  14139687 (chr 20)
rs6074704        14.12  0.0001      4.124  14115283 (chr 20)
...
```

Sib-pair compared to PLINK: association

Stratified association analysis

PLINK (14 seconds):

```
plink –noweb –bfile wgas3 –mh –within pop.cov –adjust –out cmh1

...
Cochran-Mantel-Haenszel 2x2xK test, K = 2
Writing results to [ cmh1.cmh ]
Computing corrected significance values (FDR, Sidak, etc)
Genomic inflation factor (based on median chi-squared) is 1.0147
Mean chi-squared statistic is 0.998899
Correcting for 179493 tests
Writing multiple-test corrected significance values to [ cmh1.cmh.adjusted
]
```

```
CHR         SNP  A1  A2       BP   CHISQ        P      OR      L95    U95
  1   rs3094315   G   A   792429   1.047   0.3062   1.887   0.5668  6.285
  1   rs4040617   G   A   819185  0.8534   0.3556   1.777   0.5288  5.974
  1   rs4075116   C   T  1043552   1.339   0.2472  0.2929  0.04513  1.901
  1   rs9442385   T   G  1137258  0.4069   0.5236  0.7773    0.363  1.665
  1  rs11260562   A   G  1205233 0.04184   0.8379    1.27   0.1455  11.08
  1   rs6685064   C   T  1251215   1.994   0.1579   1.804   0.7988  4.074


...
```

Sib-pair compared to PLINK: association

Stratified association analysis

Sib-pair (2 minutes 35 seconds):

```
read bin wgas3.bin
set locus group qua
update popcov.dat
set iter 0
ass trait cov group snp
summary 20
```

OR

```
read bin wgas3.bin
set locus japan aff
japan=n
select ped JA*
japan=y
unselect
set iter 0
ass trait cov japan snp
summary 20
```

```
Marker         OR       95% CI              P-value
----------- -------- -------------------- ----------
rs3094315      2.061    0.739 –    5.749 0.8361E-01 792429 (chr 1)
rs4040617      1.804    0.642 –    5.068 0.1314     819185 (chr 1)
rs4075116      1.886    0.493 –    7.207 0.1769     1043552 (chr 1)
rs9442385      1.220    0.692 –    2.151 0.2460     1137258 (chr 1)
rs11260562     1.737    0.276 –   10.948 0.2784     1205233 (chr 1)
rs6685064      1.189    0.699 –    2.022 0.2612     1251215 (chr 1)
...
```

```
Total number of tests = 180235

Locus       Position   P-value   -log10(P)
----------- ---------- -------  ----------
rs1548299         3.64  0.0000       5.263  3640174 (chr 9)
rs2513514        75.92  0.0000       5.187  75922141 (chr 11)
rs6110115        13.91  0.0000       4.841  13911728 (chr 20)
rs2508756        75.92  0.0000       4.753  75921549 (chr 11)
rs16976702       54.12  0.0000       4.711  54120691 (chr 15)
rs11204005       12.90  0.0000       4.474  12895576 (chr 8)
rs9302779         3.87  0.0000       4.387  3873346 (chr 16)
rs17534370       70.30  0.0000       4.341  70297172 (chr 9)
rs11781505      142.00  0.0001       4.266  142002911 (chr 8)
rs11785430      142.08  0.0001       4.266  142078709 (chr 8)
rs1195747       129.97  0.0001       4.146  129970575 (chr 12)
rs16910850       94.48  0.0001       4.126  94478347 (chr 9)
rs1570484        14.14  0.0001       4.124  14139687 (chr 20)
rs6074704        14.12  0.0001       4.124  14115283 (chr 20)
...
```

Sib-pair compared to PLINK: filtering

Extracting a single SNP

PLINK (3 seconds):

```
plink –noweb –bfile wgas3 –recode –snp rs11204005 –out tophit
```

Sib-pair (5 seconds):

```
read bin wgas3.bin
keep trait rs11204005
write sib rs11204005.ped
write locus sib rs11204005.in rs11204005.ped
```

Sib-pair compared to PLINK: logistic regression

Logistic regression

PLINK (0.007 seconds):

```
plink –noweb –file tophit –logistic –covar pop.cov
```

```
 CHR          SNP         BP   A1    TEST  NMISS    ODDS   STAT          P
   8   rs11204005 12895576    A     ADD     89 0.06667  -4.33  1.489e-05
   8   rs11204005 12895576    A    COV1     89   79.15   4.68  2.871e-06
```

Sib-pair compared to PLINK: logistic regression

Logistic regression

Sib-pair (0.17 seconds):

```
set locus trait                affection         .
set locus rs11204005           marker          12.895600 12895576 (chr 8)
read pedigree rs11204005.ped
run
set loc japan aff
japan = n
sel ped J*
japan=y
uns
reg trait = rs11204005 japan
ass trait snp cov japan
```

```
------------------------------------------------
Binomial regression analysis of trait “trait”
------------------------------------------------
    Variable         Beta    Stand Error        t-Value
  -----------------------------------------------------
  Intercept       -4.5510         0.9950         4.5737 ***
  rs11204005       2.7080         0.6254         4.3303 ***
  japan            4.3713         0.9341         4.6799 ***

No. usable observations =     89      (100.0%)
Number of affecteds     =     48

Null deviance           =    122.8291
Number of iterations    =      7
Model LR Chi-square     =     64.5816 (df=   2)
Akaike Inf. Criterion   =     64.2474


--------------------------------------------------
Allelic association testing for trait “trait”
--------------------------------------------------
NOTE:  Covariates are: “japan”.

Marker            OR       95% CI              P-value
-------------- -------- -------------------- ----------
rs11204005       15.000    7.264 –   30.976 0.4975E-02 12895576 (chr 8)
```

Sib-pair compared to PLINK: logistic regression

Interaction within logistic regression

PLINK

```
plink –file tophit –logistic –covar pop.cov –interaction
```

```
...
 CHR         SNP         BP A1      TEST  NMISS    ODDS      STAT          P
   8 rs11204005 12895576  A       ADD     89  0.2918    -0.645     0.5189
   8 rs11204005 12895576  A      COV1     89   319.1     2.655   0.007937
   8 rs11204005 12895576  A ADDxCOV1     89  0.3366   -0.7811     0.4348
```

Sib-pair compared to PLINK: logistic regression

Interaction within logistic regression

Sib-pair. This is not a canned procedure in Sib-pair, but is easily automated.

```
# Genotypic analysis interaction
macro interaction
 llm %1 %2 %3 %1*%2 %1*%3 %2*%3
;;;;
```

```
interaction trait rs11204005 japan
...
Model: Intercept  trait(2)  rs11204005(3)  japan(2)  trait(2)*rs11204005(3)
       trait(2)*japan(2) japan(2)*rs11204005(3)

    Term                            Beta  Stand Error     Exp(Beta)     t-Value
---------------------------------------------------------------------------
    Intercept                     2.3920       0.3016        10.935       7.931 ***
    trait(2)                     -5.1292       1.2687         0.006       4.043 ***
    rs11204005(2)                 0.5595       0.3774         1.750       1.483 +
    rs11204005(3)                -1.0235       0.5841         0.359       1.752 +
    japan(2)                     -0.5895       0.5025         0.555       1.173 .
    trait(2)*rs11204005(2)        3.5295       1.1854        34.108       2.978 *
    trait(2)*rs11204005(3)        5.7167       1.3889       303.896       4.116 ***
    trait(2)*japan(2)             4.6967       1.1258       109.584       4.172 ***
    japan(2)*rs11204005(2)       -2.5075       1.1266         0.081       2.226 +
    japan(2)*rs11204005(3)       -3.4291       1.2858         0.032       2.667 *

 No. of complete observations =    89 ( -0.0%)

                  Model LRTS =     0.30
          Degrees of freedom =     2
             Nominal P-value =     0.8598
```

```
# Allelic additive analysis interaction macro
macro interaction
 macro p <- ple
 set ple -2 silent
 macro a <- alleles %2
 out %% silent
 doanalysis %1 %2 %a %3
 out
 echo
 echo Interaction analysis trait=%1 marker=%2 covariate=%3
 file print /LRT/ %%
 file delete %%
 set ple %p silent
;;;;
macro doanalysis
 set loc dose qua
 dose=x
 if (%2 == “%3/%3”) then dose=0
 if (%2 == “%3/%4”) then dose=1
 if (%2 == “%4/%4”) then dose=2
 set loc inter qua
 inter=x
 inter = dose*%5
 reg %1 = dose %5 inter
 reg %1 = dose %5
 lrt
;;;;
```

```
interaction trait rs11204005 japan

Interaction analysis trait=trait marker=rs11204005 covariate=japan

LRTS           0.6601    1  0.4165
```

Sib-pair compared to PLINK: logistic regression

More subsetting examples

PLINK:

```
plink –file tophit –filter-males –logistic
plink –file tophit –filter-females –logistic
plink –file tophit –logistic –sex –interaction
```

Sib-pair:

```
read bin gwas3.bin
select male
ass trait snp
unselect
select female
ass trait snp
unselect
set loc fem aff
fem=female
interaction trait rs11204005 fem
```

Sib-pair compared to PLINK: filtering on interSNP LD

Subsetting to obtain SNPs in LD, and plotting clusters

PLINK: (1 minute 17 seconds)

```
plink –bfile wgas3 –indep-pairwise 50 10 0.2 –out prune1
```

Sib-pair (3 minutes for one chromosome)

```
read bin wgas3.bin
drop where r2 0.2
```

Sib-pair compared to PLINK

Individual MDS plot

PLINK

```
plink –bfile wgas3 –extract prune1.prune.in –genome –out ibs1
plink –bfile wgas3 –read-genome ibs1.genome –cluster –ppc 1e-3 –cc –mds-plot
2 –out strat1
```

Sib-pair compared to PLINK: merging in additional data

Merge in additional “fine mapping” genotyping

PLINK

```
plink –bfile wgas3 –snp rs11204005 –window 100 –merge extra.ped extra.map \
      –make-bed –out followup
plink –bfile followup –mh –within pop.cov –out followup-cmh
```

Sib-pair

```
read bin wgas3.bin
which rs11204005
keep trait 108321 -- 108421
include extras.in
update extras.dat
set loc japan aff
japan=n
sel ped J*
japan=y
unselect
# reorder data by map position
order trait japan $mm
pack
ass trait cov japan
write bin followup.bin compress
```

Sib-pair v. PLINK and haplo.stats: Haplotype association

SNP haplotype association analysis

PLINK (0.1 seconds)

```
plink –bfile followup2 –chap –hap-snps rs2460915-rs2460338
–each-versus-others
```

```
HAPLO      FREQ     OR(A)    SPEC(A)    OR(N)
  -------    ------   -------  ---------  -------
    TGTAG    0.1712   (-ref-)  1.247e-07  (-ref-)
    AGTAG    0.0111 1.117e-09    0.07573     |
    TATAG   0.06471     32.39   0.004304     |
    AGGAG    0.1441     2.201   0.004375     |
    AAGAG   0.04592     21.65    0.04734     |
    AAGGG   0.03485     823.6    0.00586     |
    AGGAC   0.02286     1.726     0.2204     |
    AGGGC   0.04373    0.7162    0.01523     |
    AAGGC    0.4382     23.52   1.83e-07     |
  -------    ------   -------    ------- --------

Model comparison test statistics:

                           Alternate       Null
                  -2LL :       63.33      122.8

  Likelihood ratio test: chi-square = 59.5
                         df = 8
                         p = 5.835e-10

[ haplo.glm() LRTS=60.24186 ]
```

Sib-pair (85 seconds)

```
dis rs2460915 rs7835221 rs2460911 rs11204005 rs2460338 trait
```

Poor old Sib-pair! In its defence, I should add it fits a “full” model testing for extragametic association, as well as the association model testing for differences in haplotype frequency. The former model does not scale up well.

```
Haplotype   n           y           [haplo.group() results]
---------   ---------   ---------
A A G A C   0.0000      0.0112      [ -       0.01133 ]
A G G A C   0.0366      0.0109      [ 0.01322 0.06798 ]
A G T A C   0.0122      0.0000      [ 0.01220 -       ]
A A G G C   0.2537      0.5993      [ 0.25375 0.59611 ]
A G G G C   0.0754      0.0140      [ 0.07537 0.01389 ]
T G G G C   0.0123      0.0000      [ 0.00566 0.01234 ]
A A G A G   0.0132      0.0672      [ 0.01322 0.06798 ]
T A G A G   0.0000      0.0104      [ -       0.01053 ]
A G G A G   0.2307      0.0773      [ 0.23069 0.07817 ]
T A T A G   0.0135      0.1126      [ 0.01352 0.11388 ]
A G T A G   0.0245      0.0000      [ 0.02454 -       ]
T G T A G   0.3278      0.0332      [ 0.32779 0.03359 ]
A A G G G   0.0000      0.0638      [ -       0.06346 ]

trait               41          48

            Number of loci =         5
 No. genotyped individuals =        89
 No. obs. unique genotypes =        30
  Stratified LD Chi-square =        44.30 (df= 470, P=1.0000)
    Association Chi-square =        76.92 (df=   7, P=0.0000)

NOTE:  Degrees of freedom calculation for association test assumes only
          8 haplotypes to be present in the population.

[ the LRTS based on haplo.em() is 76.66 ]
```

Sib-pair v. PLINK: Conclusions

- Similar range of simpler association analysis
- Sib-pair is still slower, but not annoyingly so (some exceptions)
- Sib-pair offers greater interactivity/programmability