Practical: What you should see

Chris & Tsaone

File preparation *tar –xvzf raw-GWA-data*

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ls -lrt
total 6341960
-rwxr-xr-x@ 1 samieznew  staff    11747611 Jan  6  2010 raw-GWA-data.map
-rw-r--r--@ 1 samieznew  staff         417 Jan  7  2010 high-LD-regions.txt
-rw-r--r--@ 1 samieznew  staff      511429 Jan  7  2010 raw-GWA-data.prune.in
-rw-r--r--@ 1 samieznew  staff  2540057786 Jan 14  2010 raw-GWA-data.ped
-rw-r--r--@ 1 samieznew  staff         718 Feb 16  2010 imiss-vs-het.Rscript
-rw-r--r--@ 1 samieznew  staff         283 Mar  3  2010 plot-IBD.Rscript
-rw-r--r--@ 1 samieznew  staff        1217 Mar  8  2010 run-IBD-QC.pl
-rw-r--r--@ 1 samieznew  staff    10664954 Mar  8  2010 hapmap3r2_CEU.CHB.JPT.YRI.no-at-cg-snps.txt
-rw-r--r--@ 1 samieznew  staff   103669932 Mar  9  2010 hapmap3r2_CEU.CHB.JPT.YRI.founders.no-at-cg-snps.bed
-rw-r--r--@ 1 samieznew  staff    29166116 Mar  9  2010 hapmap3r2_CEU.CHB.JPT.YRI.founders.no-at-cg-snps.bim
-rw-r--r--@ 1 samieznew  staff        8813 Mar  9  2010 hapmap3r2_CEU.CHB.JPT.YRI.founders.no-at-cg-snps.fam
-rw-r--r--@ 1 samieznew  staff           9 Mar  9  2010 pca-populations.txt
-rw-r--r--@ 1 samieznew  staff         327 Mar 11  2010 run-diffmiss-qc.pl
-rw-r--r--@ 1 samieznew  staff         776 Jun  4  2010 plot-pca-results.Rscript
-rw-r--r--  1 samieznew  staff          27 Feb 19  2018 toy.map
-rw-r--r--  1 samieznew  staff       35147 Feb 19  2018 LICENSE
-rwxr-xr-x  1 samieznew  staff       18332 Mar  6  2019 prettify
-rw-r--r--  1 samieznew  staff          58 Jan  5  2024 toy.ped
-rwxr-xr-x  1 samieznew  staff     4344760 Aug 20 06:41 plink
```

./plink --file raw-GWA-data --make-bed --out raw-GWA-data

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ./plink --file raw-GWA-data --make-bed --out raw-GWA-data

PLINK v1.9.0-b.7.11 64-bit (19 Aug 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to raw-GWA-data.log.
Options in effect:
  --file raw-GWA-data
  --make-bed
  --out raw-GWA-data

16384 MB RAM detected; reserving 8192 MB for main workspace.
.ped scan complete (for binary autoconversion).
Performing single-pass .bed write (317503 variants, 2000 people).
--file: raw-GWA-data-temporary.bed + raw-GWA-data-temporary.bim +
raw-GWA-data-temporary.fam written.
317503 variants loaded from .bim file.
2000 people (997 males, 1003 females) loaded from .fam.
2000 phenotype values loaded from .fam.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 2000 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Warning: 11440 het. haploid genotypes present (see raw-GWA-data.hh ); many
commands treat these as missing.
Total genotyping rate is 0.985682.
317503 variants and 2000 people pass filters and QC.
Among remaining phenotypes, 1023 are cases and 977 are controls.
--make-bed to raw-GWA-data.bed + raw-GWA-data.bim + raw-GWA-data.fam ... done.
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Sample QC

Step 1: Identification of individuals with discordant sex information

./plink --bfile raw-GWA-data --check-sex --out raw-GWA-data

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ./plink --bfile raw-GWA-data --check-sex --out raw-GWA-data
PLINK v1.9.0-b.7.11 64-bit (19 Aug 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to raw-GWA-data.log.
Options in effect:
  --bfile raw-GWA-data
  --check-sex
  --out raw-GWA-data

16384 MB RAM detected; reserving 8192 MB for main workspace.
317503 variants loaded from .bim file.
2000 people (997 males, 1003 females) loaded from .fam.
2000 phenotype values loaded from .fam.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 2000 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Warning: 11440 het. haploid genotypes present (see raw-GWA-data.hh ); many
commands treat these as missing.
Total genotyping rate is 0.985682.
317503 variants and 2000 people pass filters and QC.
Among remaining phenotypes, 1023 are cases and 977 are controls.
--check-sex: 8921 Xchr and 0 Ychr variant(s) scanned, 3 problems detected.
Report written to raw-GWA-data.sexcheck .
(base) samieznew@MacBook-Pro-7 Teaching_Material % head raw-GWA-data.sexcheck
 FID  IID       PEDSEX       SNPSEX       STATUS            F
   1    1            1            1           OK       0.9552
   2    2            1            1           OK       0.9997
   3    3            1            1           OK       0.9935
   4    4            1            1           OK            1
   5    5            2            2           OK     0.007004
   6    6            2            2           OK     0.007528
   7    7            2            2           OK    -0.003711
   8    8            2            2           OK     0.007389
   9    9            2            2           OK    -0.009868
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

A list of individuals with discordant sex data

grep PROBLEM raw-GWA-data.sexcheck > raw-GWA-data.sexprobs

awk '{if ($5=="PROBLEM")print }' raw-GWA-data.sexcheck | head

wc -l raw-GWA-data.sexprobs

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % awk '{if ($5=="PROBLEM")print }' raw-GWA-data.sexcheck | head
 772  772           2           0      PROBLEM       0.3084
 853  853           2           0      PROBLEM       0.3666
1920 1920           2           0      PROBLEM       0.4066
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % wc -l raw-GWA-data.sexprobs
       3 raw-GWA-data.sexprobs
```

awk '{if ($5=="PROBLEM")print }' raw-GWA-data.sexcheck > fail-sexcheck-qc.txt

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % awk '{if ($5=="PROBLEM")print }' raw-GWA-data.sexcheck > fail-sexcheck-qc.txt
(base) samieznew@MacBook-Pro-7 Teaching_Material % cat fail-sexcheck-qc.txt
 772  772           2           0      PROBLEM       0.3084
 853  853           2           0      PROBLEM       0.3666
1920 1920           2           0      PROBLEM       0.4066
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Step 2: Identification of individuals with elevated missing data rates or outlying heterozygosity rate

./Plink --bfile raw-GWA-data --missing --out raw-GWA-data

The fourth column in the file "raw-GWA-data.imiss" (N_MISS) denotes the number of missing SNPs and the sixth column (F_MISS) denotes the proportion of missing SNPs per individual.

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ./Plink --bfile raw-GWA-data --missing --out raw-GWA-data

PLINK v1.9.0-b.7.11 64-bit (19 Aug 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to raw-GWA-data.log.
Options in effect:
  --bfile raw-GWA-data
  --missing
  --out raw-GWA-data

16384 MB RAM detected; reserving 8192 MB for main workspace.
317503 variants loaded from .bim file.
2000 people (997 males, 1003 females) loaded from .fam.
2000 phenotype values loaded from .fam.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 2000 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Warning: 11440 het. haploid genotypes present (see raw-GWA-data.hh ); many
commands treat these as missing.
Total genotyping rate is 0.985682.
--missing: Sample missing data report written to raw-GWA-data.imiss, and
variant-based missing data report written to raw-GWA-data.lmiss.
(base) samieznew@MacBook-Pro-7 Teaching_Material % head raw-GWA-data.imiss
 FID  IID MISS_PHENO   N_MISS   N_GENO   F_MISS
   1    1          N     4634   317503   0.0146
   2    2          N     3695   317503  0.01164
   3    3          N     3730   317503  0.01175
   4    4          N     3698   317503  0.01165
   5    5          N     3739   317503  0.01178
   6    6          N     3709   317503  0.01168
   7    7          N     3708   317503  0.01168
   8    8          N     3702   317503  0.01166
   9    9          N     3768   317503  0.01187
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Calculating heterozygosity

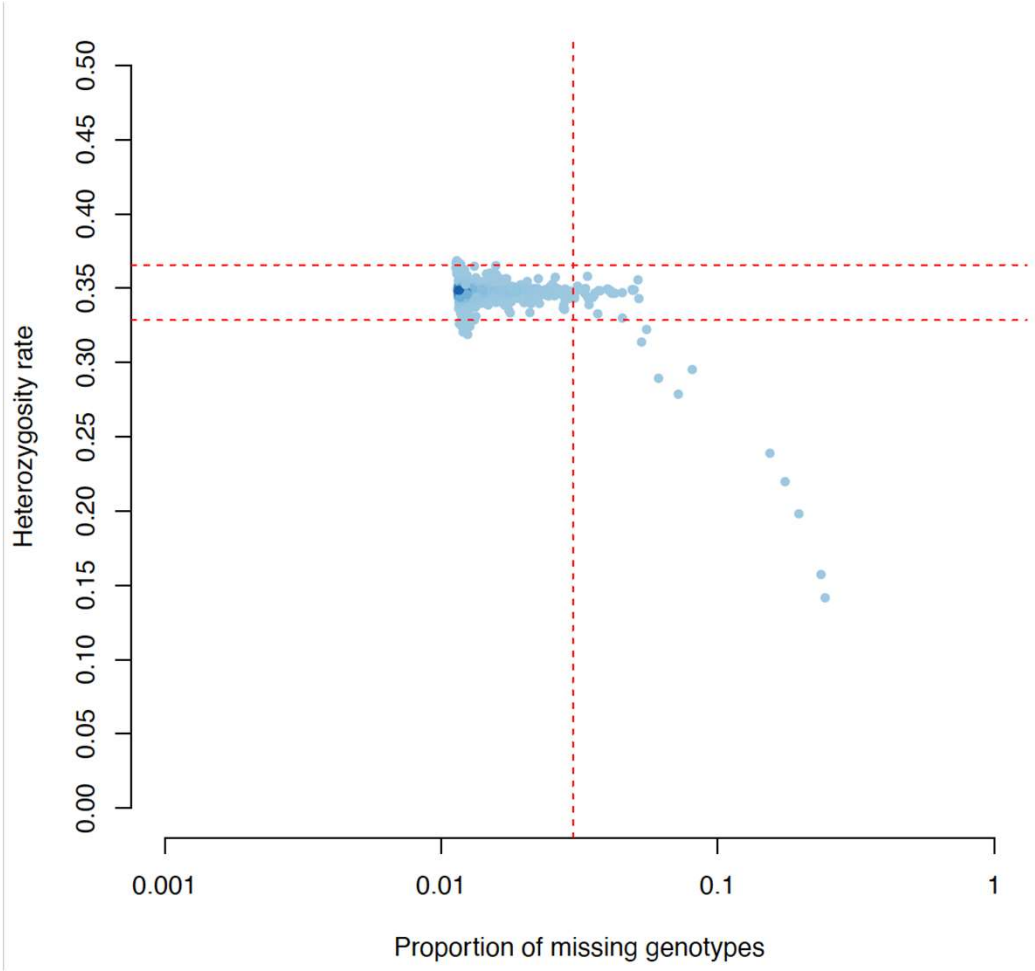
./Plink --bfile raw-GWA-data --het --out raw-GWA-data

This command will create the file “raw-GWA-data.het”, in which the third column denotes the observed number of homozygous genotypes [O(Hom)] and the fifth column denotes the number of non-missing genotypes [N(NM)] per individual.

heterozygosity rate per individual,
het = Het = (N(NM) − O(Hom))/N(NM)

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ./Plink --bfile raw-GWA-data --het --out raw-GWA-data

PLINK v1.9.0-b.7.11 64-bit (19 Aug 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to raw-GWA-data.log.
Options in effect:
  --bfile raw-GWA-data
  --het
  --out raw-GWA-data

16384 MB RAM detected; reserving 8192 MB for main workspace.
317503 variants loaded from .bim file.
2000 people (997 males, 1003 females) loaded from .fam.
2000 phenotype values loaded from .fam.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 2000 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Warning: 11440 het. haploid genotypes present (see raw-GWA-data.hh ); many
commands treat these as missing.
Total genotyping rate is 0.985682.
317503 variants and 2000 people pass filters and QC.
Among remaining phenotypes, 1023 are cases and 977 are controls.
--het: 305281 variants scanned, report written to raw-GWA-data.het .
(base) samieznew@MacBook-Pro-7 Teaching_Material % head raw-GWA-data.het
 FID  IID       O(HOM)       E(HOM)        N(NM)            F
   1    1       194674    1.979e+05       303970     -0.03007
   2    2       198537    1.985e+05       304880    0.0007205
   3    3       198960    1.984e+05       304845     0.004942
   4    4       198878    1.985e+05       304876     0.003992
   5    5       198530    1.984e+05       304839    0.0009213
   6    6       198723    1.984e+05       304864     0.002588
   7    7       198980    1.985e+05       304871     0.004969
   8    8       199395    1.985e+05       304874     0.008868
   9    9       199111    1.984e+05       304812       0.0066
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

observed heterozygosity rate per individual



Inidivuals that fail heterozygosity check

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % Rscript imiss_vs_het_QC.R
Number of failing individuals: 45
Output written to fail-imisshet-qc.txt
(base) samieznew@MacBook-Pro-7 Teaching_Material % head fail-imisshet-qc.txt
1003 1003
1006 1006
1045 1045
1058 1058
1154 1154
1236 1236
1294 1294
1395 1395
1537 1537
1554 1554
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Step 3a: Identification of duplicated or related individuals

./Plink --bfile raw-GWA-data --extract raw-GWA-data.prune.in --genome --out raw-GWA-data

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ./Plink --bfile raw-GWA-data --extract raw-GWA-data.prune.in --genome --out raw-GWA-data

PLINK v1.9.0-b.7.11 64-bit (19 Aug 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to raw-GWA-data.log.
Options in effect:
  --bfile raw-GWA-data
  --extract raw-GWA-data.prune.in
  --genome
  --out raw-GWA-data

16384 MB RAM detected; reserving 8192 MB for main workspace.
317503 variants loaded from .bim file.
2000 people (997 males, 1003 females) loaded from .fam.
2000 phenotype values loaded from .fam.
--extract: 51194 variants remaining.
Using up to 4 threads (change this with --threads).
Before main variant filters, 2000 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Warning: 1698 het. haploid genotypes present (see raw-GWA-data.hh ); many
commands treat these as missing.
Total genotyping rate is 0.994749.
51194 variants and 2000 people pass filters and QC.
Among remaining phenotypes, 1023 are cases and 977 are controls.
Excluding 1335 variants on non-autosomes from IBD calculation.
IBD calculations complete.
Finished writing raw-GWA-data.genome .
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Step 3b: Identification of duplicated or related individuals

perl run-IBD-QC.pl raw-GWA-data

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % perl run-IBD-QC.pl raw-GWA-data

Reading PLINK .imiss file raw-GWA-data.imiss
Reading PLINK .genome file raw-GWA-data.genome
```

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % head fail-IBD-QC.txt
1952 1952
1953 1953
1954 1954
1955 1955
1957 1957
1959 1959
1961 1961
1963 1963
1965 1965
1967 1967
(base) samieznew@MacBook-Pro-7 Teaching_Material % wc -l fail-IBD-QC.txt
      14 fail-IBD-QC.txt
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Step 4: Removal of all individuals failing sample QC

cat fail-* | sort -k1 | uniq > fail-qc-inds.txt

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % cat fail-* | sort -k1 | uniq > fail-qc-inds.txt

(base) samieznew@MacBook-Pro-7 Teaching_Material % wc -l fail-qc-inds.txt
      62 fail-qc-inds.txt
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Remove individuals failing sample QC

./plink --bfile raw-GWA-data --remove fail-qc-inds.txt --make-bed --out clean-inds-GWA-data

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ./plink --bfile raw-GWA-data --remove fail-qc-inds.txt --make-bed --out clean-inds-GWA-data

PLINK v1.9.0-b.7.11 64-bit (19 Aug 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to clean-inds-GWA-data.log.
Options in effect:
  --bfile raw-GWA-data
  --make-bed
  --out clean-inds-GWA-data
  --remove fail-qc-inds.txt

16384 MB RAM detected; reserving 8192 MB for main workspace.
317503 variants loaded from .bim file.
2000 people (997 males, 1003 females) loaded from .fam.
2000 phenotype values loaded from .fam.
--remove: 1941 people remaining.
Warning: At least 3 duplicate IDs in --remove file.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 1941 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Warning: 10995 het. haploid genotypes present (see clean-inds-GWA-data.hh );
many commands treat these as missing.
Total genotyping rate in remaining samples is 0.986701.
317503 variants and 1941 people pass filters and QC.
Among remaining phenotypes, 989 are cases and 952 are controls.
--make-bed to clean-inds-GWA-data.bed + clean-inds-GWA-data.bim +
clean-inds-GWA-data.fam ... done.
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

You should have these files in your folder now, for the next part

```
-rw-r--r--  1 samieznew  staff      206620 Dec  3 15:13 clean-inds-GWA-data.hh
-rw-r--r--  1 samieznew  staff   154306461 Dec  3 15:13 clean-inds-GWA-data.bed
-rw-r--r--  1 samieznew  staff       32786 Dec  3 15:13 clean-inds-GWA-data.fam
-rw-r--r--  1 samieznew  staff     8788167 Dec  3 15:13 clean-inds-GWA-data.bim
-rw-r--r--  1 samieznew  staff        1196 Dec  3 15:13 clean-inds-GWA-data.log
-rw-r--r--@ 1 samieznew  staff      116676 Dec  3 15:15 GWAS_Practical_CK2.docx
(base) samieznew@MacBook-Pro-7 Teaching Material %
```

Part B: SNP Quality Control

Step 1: Identification of all SNPs with an excessive missing data rate

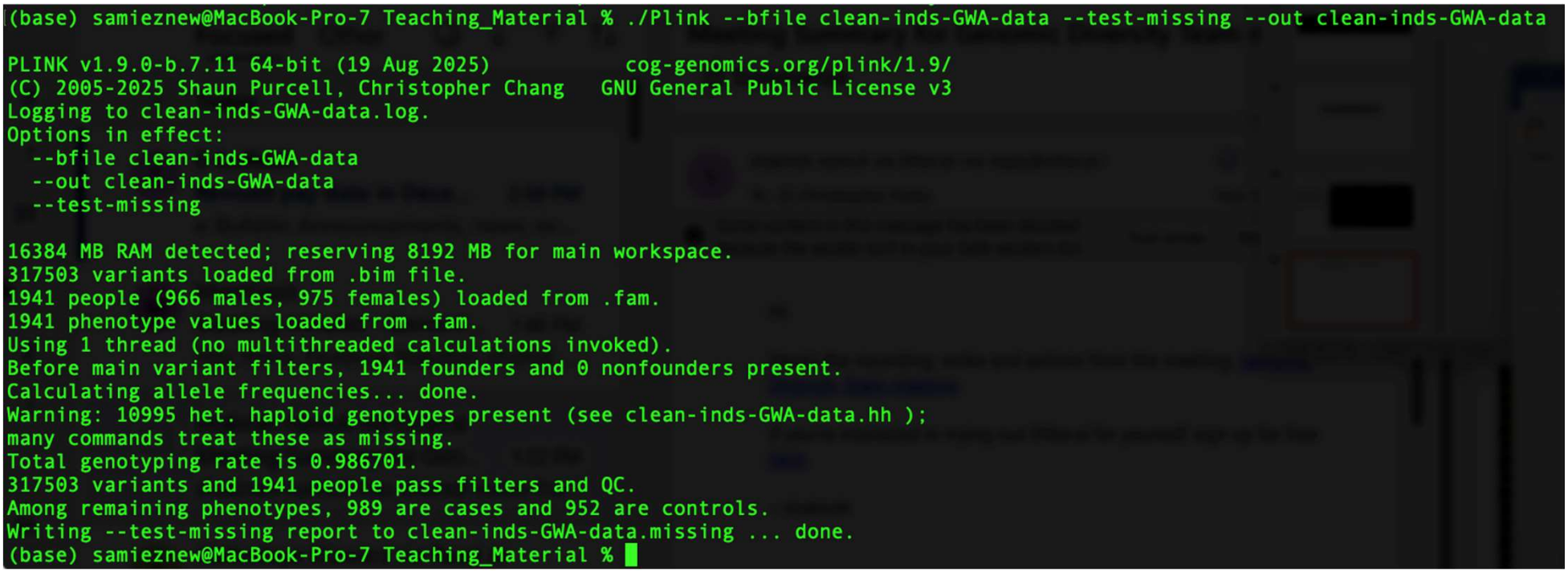
./Plink --bfile clean-inds-GWA-data --missing --out clean-inds-GWA-data

The third column in the file "clean-inds-GWA-data.lmiss" (N_MISS) denotes the number of missing genotypes and the fifth column (F_MISS) denotes the proportion of missing genotypes per SNP

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ./Plink --bfile clean-inds-GWA-data --missing --out clean-inds-GWA-data
PLINK v1.9.0-b.7.11 64-bit (19 Aug 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to clean-inds-GWA-data.log.
Options in effect:
  --bfile clean-inds-GWA-data
  --missing
  --out clean-inds-GWA-data

16384 MB RAM detected; reserving 8192 MB for main workspace.
317503 variants loaded from .bim file.
1941 people (966 males, 975 females) loaded from .fam.
1941 phenotype values loaded from .fam.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 1941 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Warning: 10995 het. haploid genotypes present (see clean-inds-GWA-data.hh );
many commands treat these as missing.
Total genotyping rate is 0.986701.
--missing: Sample missing data report written to clean-inds-GWA-data.imiss, and
variant-based missing data report written to clean-inds-GWA-data.lmiss.
(base) samieznew@MacBook-Pro-7 Teaching_Material % head clean-inds-GWA-data.lmiss
 CHR          SNP   N_MISS   N_GENO   F_MISS
   1   rs3934834        2     1941  0.00103
   1   rs3737728        3     1941 0.001546
   1   rs6687776        3     1941 0.001546
   1   rs9651273        8     1941 0.004122
   1   rs4970405        2     1941  0.00103
   1  rs12726255        2     1941  0.00103
   1   rs2298217        4     1941 0.002061
   1   rs4970357        4     1941 0.002061
   1   rs4970362        2     1941  0.00103
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Step 2: Test SNPs for different genotype call rates between cases and controls

./Plink --bfile clean-inds-GWA-data --test-missing --out clean-inds-GWA-data



Step2 output

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % head clean-inds-GWA-data.missing
 CHR          SNP     F_MISS_A     F_MISS_U            P
   1    rs3934834     0.002022            0       0.4999
   1    rs3737728            0     0.003151       0.1178
   1    rs6687776     0.002022      0.00105            1
   1    rs9651273     0.003033     0.005252       0.4992
   1    rs4970405            0     0.002101       0.2404
   1   rs12726255     0.002022            0       0.4999
   1    rs2298217     0.003033      0.00105       0.6249
   1    rs4970357     0.002022     0.002101            1
   1    rs4970362     0.001011      0.00105            1
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

highlight all SNPs with significant differences in case and control call rates ($p<10^{-5}$) from this output file

perl run-diffmiss-qc.pl clean-inds-GWA-data

```
-rw-r--r--  1 samieznew  staff      206620 Dec  3 15:38 clean-inds-GWA-data.hh
-rw-r--r--  1 samieznew  staff    17159409 Dec  3 15:38 clean-inds-GWA-data.missing
-rw-r--r--  1 samieznew  staff        1043 Dec  3 15:38 clean-inds-GWA-data.log
-rw-r--r--@ 1 samieznew  staff     4441473 Dec  3 15:40 QC_CK.pptx
-rw-r--r--  1 samieznew  staff           0 Dec  3 15:42 fail-diffmiss-qc.txt
(base) samieznew@MacBook-Pro-7 Teaching_Material % perl run-diffmiss-qc.pl clean-inds-GWA-data
```

None, maybe because the data is simulated?

Step 3: Removal of all SNPs failing QC

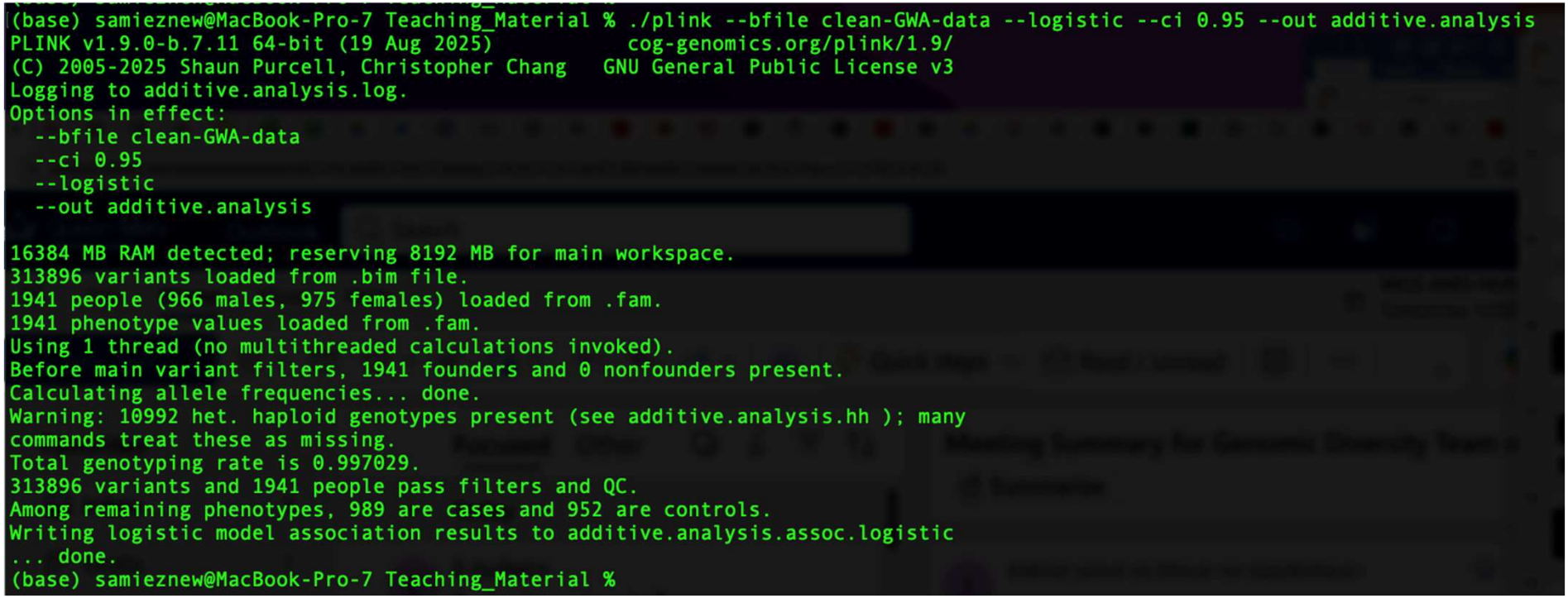
./plink --bfile clean-inds-GWA-data --exclude fail-diffmiss-qc.txt --geno 0.05 --hwe 0.00001 --make-bed --out clean-GWA-data

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % ./plink --bfile clean-inds-GWA-data --exclude fail-diffmiss-qc.txt  --geno 0.05 --hwe 0.00001 --make-bed --out clean-GWA-data

PLINK v1.9.0-b.7.11 64-bit (19 Aug 2025)           cog-genomics.org/plink/1.9/
(C) 2005-2025 Shaun Purcell, Christopher Chang   GNU General Public License v3
Logging to clean-GWA-data.log.
Options in effect:
  --bfile clean-inds-GWA-data
  --exclude fail-diffmiss-qc.txt
  --geno 0.05
  --hwe 0.00001
  --make-bed
  --out clean-GWA-data

16384 MB RAM detected; reserving 8192 MB for main workspace.
317503 variants loaded from .bim file.
1941 people (966 males, 975 females) loaded from .fam.
1941 phenotype values loaded from .fam.
--exclude: 317503 variants remaining.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 1941 founders and 0 nonfounders present.
Calculating allele frequencies... done.
Warning: 10995 het. haploid genotypes present (see clean-GWA-data.hh ); many
commands treat these as missing.
Total genotyping rate is 0.986701.
3341 variants removed due to missing genotype data (--geno).
Warning: --hwe observation counts vary by more than 10%, due to the X
chromosome.  You may want to use a more stringent (i.e. less extreme) --hwe
p-value threshold for X chromosome variants: male samples are ignored there, so
the same degree of HWE violation corresponds to a less-extreme p-value than it
does elsewhere in the genome.
--hwe: 16 variants removed due to Hardy-Weinberg exact test.
314146 variants and 1941 people pass filters and QC.
Among remaining phenotypes, 989 are cases and 952 are controls.
--make-bed to clean-GWA-data.bed + clean-GWA-data.bim + clean-GWA-data.fam ...
done.
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Practical Two: Basic analysis of genome-wide association studies

Genome-wide association analysis

./plink --bfile clean-GWA-data --logistic --ci 0.95 –out additive.analysis

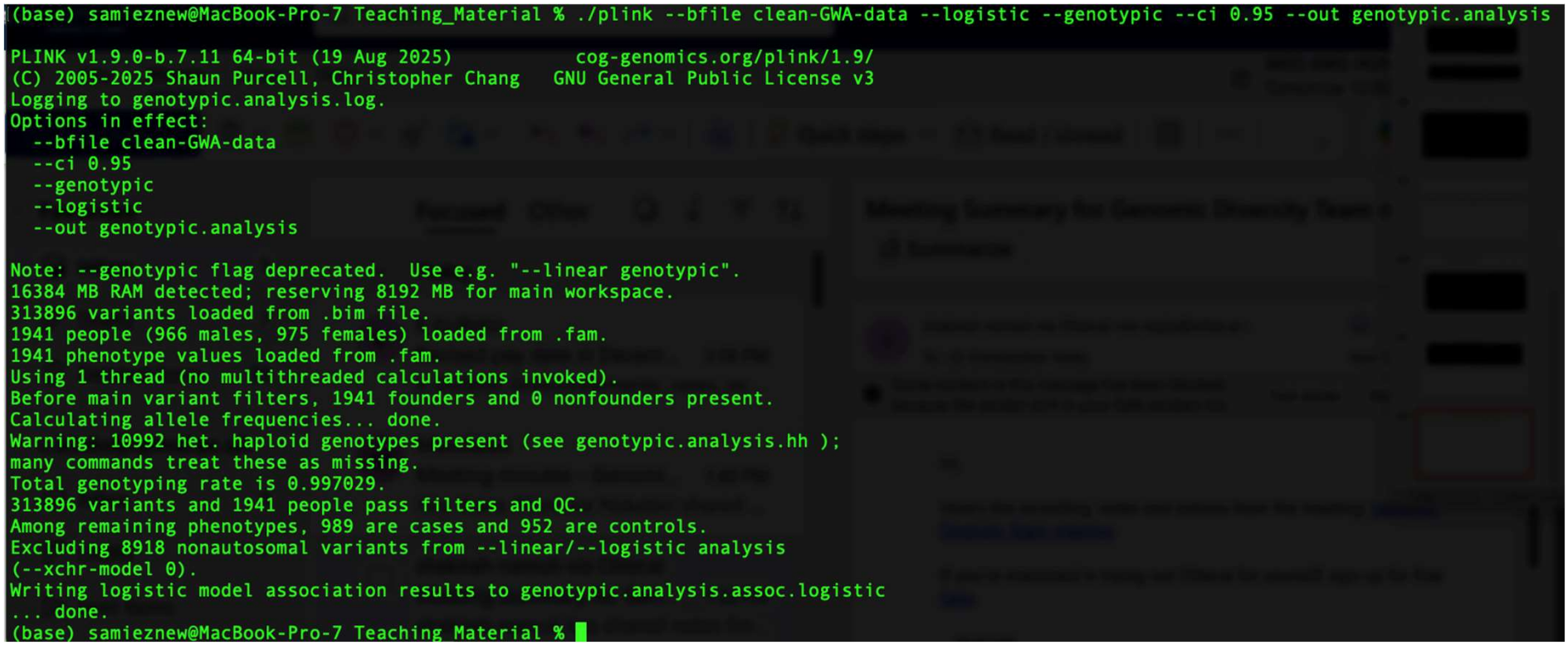


Output file (additive model)

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % head additive.analysis.assoc.logistic
 CHR         SNP         BP   A1       TEST    NMISS         OR         SE      L95      U95         STAT            P
   1   rs3934834    1045729    T        ADD     1939     0.9005     0.0993   0.7412    1.094       -1.056       0.2912
   1   rs3737728    1061338    A        ADD     1938      1.107    0.07213   0.9607    1.275        1.404       0.1602
   1   rs6687776    1070488    T        ADD     1938     0.9241     0.1143   0.7387    1.156      -0.6907       0.4898
   1   rs9651273    1071463    A        ADD     1933      1.119    0.06641   0.9828    1.275        1.699      0.08941
   1   rs4970405    1088878    G        ADD     1939      1.066     0.1452   0.8018    1.417       0.4393       0.6605
   1  rs12726255    1089873    G        ADD     1939      1.035     0.1177   0.8214    1.303       0.2888       0.7727
   1   rs2298217    1104902    T        ADD     1937     0.7615     0.1232   0.5981   0.9695       -2.212      0.02699
   1   rs4970357    1116987    C        ADD     1937      1.109     0.1239   0.8697    1.413       0.8328        0.405
   1   rs4970362    1134661    A        ADD     1939     0.9531    0.06863   0.8332     1.09      -0.6995       0.4843
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Test for association with disease under a genotypic model

./plink --bfile clean-GWA-data --logistic --genotypic --ci 0.95 --out genotypic.analysis

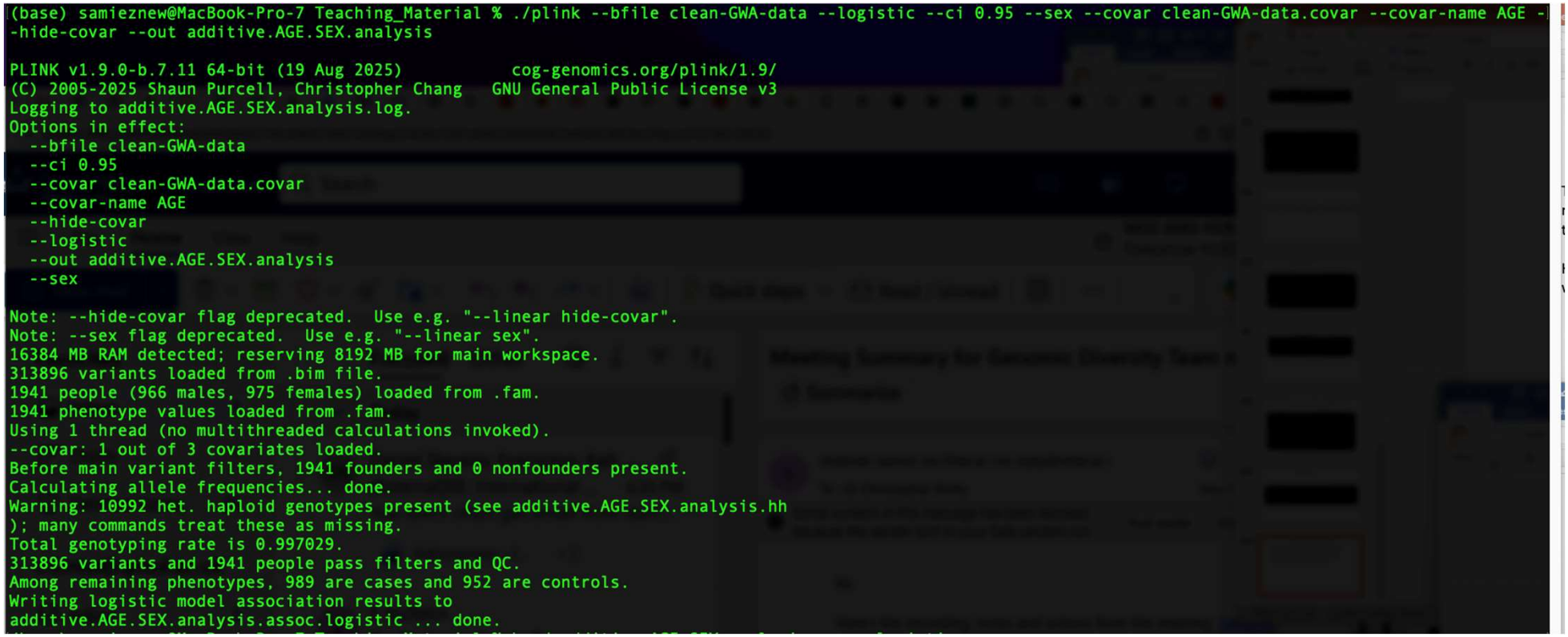


Output genotypic model

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % head genotypic.analysis.assoc.logistic
 CHR         SNP         BP   A1       TEST    NMISS         OR       SE      L95      U95         STAT            P
   1   rs3934834    1045729    T        ADD     1939      1.116   0.1927   0.7647    1.627       0.5675       0.5704
   1   rs3934834    1045729    T     DOMDEV     1939     0.7538    0.217   0.4927    1.153       -1.303       0.1927
   1   rs3934834    1045729    T   GENO_2DF     1939         NA       NA       NA       NA        2.822       0.2439
   1   rs3737728    1061338    A        ADD     1938      1.135  0.08909   0.9534    1.352        1.424       0.1544
   1   rs3737728    1061338    A     DOMDEV     1938     0.9453   0.1145   0.7552    1.183      -0.4913       0.6232
   1   rs3737728    1061338    A   GENO_2DF     1938         NA       NA       NA       NA        2.206       0.3318
   1   rs6687776    1070488    T        ADD     1938      1.033   0.2442   0.6401    1.667       0.1329       0.8943
   1   rs6687776    1070488    T     DOMDEV     1938     0.8696   0.2707   0.5116    1.478      -0.5164       0.6056
   1   rs6687776    1070488    T   GENO_2DF     1938         NA       NA       NA       NA         0.744       0.6893
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Test for association with disease allowing for covariates

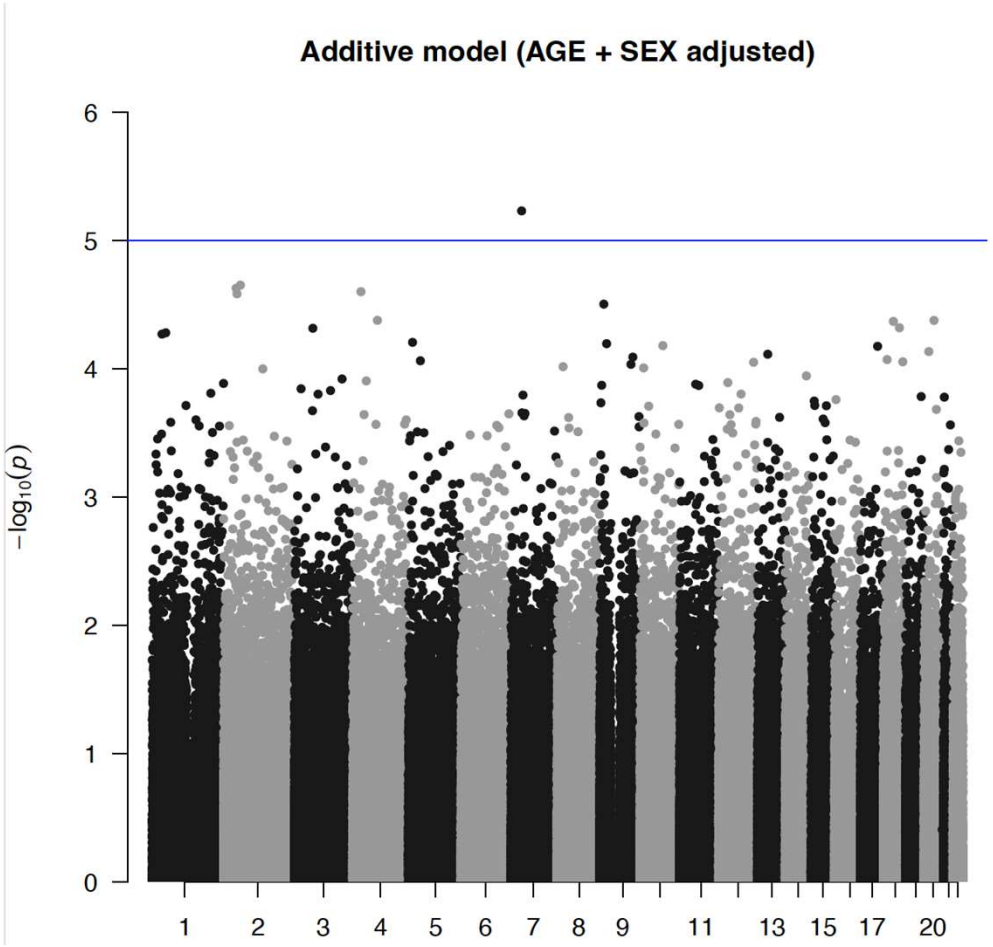
./plink --bfile clean-GWA-data --logistic --ci 0.95 --sex --covar clean-GWA-data.covar --covar-name AGE --hide-covar --out additive.AGE.SEX.analysis



Outputfile after covariate adjustment

```
(base) samieznew@MacBook-Pro-7 Teaching_Material % head additive.AGE.SEX.analysis.assoc.logistic
 CHR         SNP         BP   A1       TEST    NMISS         OR         SE      L95      U95         STAT            P
   1   rs3934834    1045729    T        ADD     1939     0.8834     0.1196   0.6988    1.117       -1.037       0.2998
   1   rs3737728    1061338    A        ADD     1938       1.11    0.08714   0.9358    1.317        1.199       0.2307
   1   rs6687776    1070488    T        ADD     1938     0.9652     0.1367   0.7384    1.262      -0.2589       0.7957
   1   rs9651273    1071463    A        ADD     1933      1.207    0.08006    1.032    1.412        2.352      0.01869
   1   rs4970405    1088878    G        ADD     1939      1.081     0.1742   0.7682    1.521       0.4467       0.6551
   1  rs12726255    1089873    G        ADD     1939      1.113     0.1417   0.8434     1.47       0.7578       0.4486
   1   rs2298217    1104902    T        ADD     1937      0.824     0.1468    0.618    1.099       -1.319       0.1871
   1   rs4970357    1116987    C        ADD     1937       1.09       0.15   0.8122    1.462       0.5735       0.5663
   1   rs4970362    1134661    A        ADD     1939      1.004    0.08198   0.8547    1.179      0.04498       0.9641
(base) samieznew@MacBook-Pro-7 Teaching_Material %
```

Manhattan. Interpretation?



END