

Plink and GRM practical

In this practical, we will:

- Learn about genetic relatedness / relationship matrices by estimating one.
- Learn about some concepts and methods from molecular genetic methods.

Our data

- Simulated dataset based on real data.

The screenshot displays the Nature Genetics journal website interface. At the top left, the journal logo "nature genetics" is visible. To the right, there are links for "Login" and "Cart". A search bar with a "go" button and a link to "Advanced search" is located below the header. The main content area shows a breadcrumb trail: "Journal home > Archive > Letter > Abstract". On the left, a "Journal content" sidebar lists various options, with "Archive" selected. The central part of the page features a "Letter abstract" for the article "A survey of genetic human cortical gene expression", published in *Nature Genetics* 39, 1494 - 1499 (2007). The authors listed are Amanda J Myers, J Raphael Gibbs, Jennifer A Webster, Kristen Rohrer, Alice Zhao, Lauren Marlowe, Mona Kaleem, Doris Leung, Leslie Bryden, Priti Nath, Victoria L Zismann, Keta Joshipura, Matthew J Huentelman, Diane Hu-Lince, Keith D Coon, David W Craig, John V Pearson, Peter Holmans, Christopher B Heward, Eric M Reiman, Dietrich Stephan, and John Hardy. On the right, a "Language selector" sidebar includes a checked option for "Kanji abstract for this article" and a "This issue" section with links to "Table of contents", "Previous abstract", and "Next abstract". An "Article tools" section at the bottom right contains a link for "Full text".

Journal content

- Journal home
- Advance online publication
- Current issue
- Archive**
- Focuses and Supplements
- Press releases
- Free Association

Letter abstract

Nature Genetics 39, 1494 - 1499 (2007)
Published online: 4 November 2007 | doi:10.1038/ng.2007.16

A survey of genetic human cortical gene expression

Amanda J Myers^{1,2,10}, J Raphael Gibbs^{1,3,10}, Jennifer A Webster^{4,5,10}, Kristen Rohrer¹, Alice Zhao¹, Lauren Marlowe¹, Mona Kaleem¹, Doris Leung¹, Leslie Bryden¹, Priti Nath¹, Victoria L Zismann^{4,5}, Keta Joshipura^{4,5}, Matthew J Huentelman^{4,5}, Diane Hu-Lince^{4,5}, Keith D Coon^{4,5,5}, David W Craig^{4,5}, John V Pearson^{4,5}, Peter Holmans⁷, Christopher B Heward⁸, Eric M Reiman^{4,5,9}, Dietrich Stephan^{4,5,9} & John Hardy^{1,3}

Language selector

- Kanji abstract for this article

This issue

- Table of contents
- Previous abstract
- Next abstract

Article tools

- Full text

1	WGACON 1	0 0 1 1	C T G G A G A G T T C C A A G G A A T T C T C C C C G G G T A G
2	WGACON 2	0 0 1 1	C T G G A G A G 0 0 C C G G G G A C C T C C C C C C G G G T A G
3	WGACON 3	0 0 1 1	T T G G A A G G T T C C A G G G A C C T C C C C C T G G G T A G
4	WGACON 4	0 0 2 1	T T G G A A G G T T C C A G G G A A T T C C C C C C G G 0 0 A G
5	WGACON 5	0 0 2 1	T T G G A A G G T T C C A A G T A A T T T T C C C C G T G T A A
6	WGACON 6	0 0 1 1	T T G G A A G G T T C C A A G G A A T T C T C C C C G G G T A A
7	WGACON 7	0 0 1 1	C C G G 0 0 0 0 T T C C A A G T A A T T C T C C C C G G G T G G
8	WGACON 8	0 0 1 1	T T G G A A G G T T C C A G G G A A C C C C C C C G G G T A A
9	WGACON 9	0 0 1 1	T T G G A A G G 0 0 C C A A G G A C C T C C C C C C G G 0 0 A G
10	WGACON 10	0 0 1 1	T T G G A A G G T T C C A A G G A A C T C T C C C C G G G T A G
11	WGACON 11	0 0 2 1	C T G G A G 0 0 T T C C G G T T A A T T T T C C C C G G G T A A
12	WGACON 12	0 0 1 1	T T G G A A G G T T C C A A G T A A T T C C C C C C G G G T G G
13	WGACON 13	0 0 1 1	T T G G A A G G T T C C G G G G A C T T C C C C C C G G 0 0 A A
14	WGACON 14	0 0 2 1	C T G G A A G G T T C C A G G G A A T T C C C C C C G G G T A A
15	WGACON 15	0 0 1 1	C T G G A A A G C T C C A G G G A C C T C C C C C C G G G T G G
16	WGACON 16	0 0 1 1	T T G G A A G G T T C C A A G T A A C C C C C C C G T G T A A
17	WGACON 17	0 0 1 1	T T G G A A G G T T C C A G G G A A T T C C C C C C G G G T A A
18	WGACON 18	0 0 2 1	T T G G A A G G T T C C A A G G A C T T C C C C C C G G G T A A
19	WGACON 19	0 0 2 1	C T G G A G A G 0 0 C C A A G G A A C T C C C C C C G G 0 0 A A
20	WGACON 20	0 0 1 1	T T G G A A G G T T C C A A G G 0 0 T T C C C T C C G G G T A G
21	WGACON 21	0 0 1 1	C C G G G G 0 0 T T C C A A G G A C T T C C C C C C G G G T A A
22	WGACON 22	0 0 1 1	T T G G A A G G T T C C A A G G A A C T C C C C C T G G G T A G
23	WGACON 23	0 0 2 1	C T G G A A G G T T C C 0 0 G G A A T T C C C C C C G G 0 0 A A
24	WGACON 24	0 0 2 1	T T G G A A G G T T C C A G G G A C T T C C C C C C G G G T A G
25	WGACON 25	0 0 1 1	T T G G A A A G T T C C A G G G A A T T C C C C C C G G G T G G
26	WGACON 26	0 0 1 1	C T G G A G A G T T C C A A G G A A T T C T C C C C G G G T A G
27	WGACON 27	0 0 1 1	C T G G A G A G 0 0 C C G G G G A C C T C C C C C C G G G T A G
28	WGACON 28	0 0 1 1	T T G G A A G G T T C C A G G G A C C T C C C C C T G G G T A G
29	WGACON 29	0 0 2 1	T T G G A A G G T T C C A G G G A A T T C C C C C C G G 0 0 A G
30	WGACON 30	0 0 2 1	T T G G A A G G !
31	WGACON 31	0 0 1 1	T T G G A A G G !
32	WGACON 32	0 0 1 1	C C G G 0 0 0 0 !
33	WGACON 33	0 0 1 1	T T G G A A G G !

gwas_plinkdata.ped

- 150 individuals
- 23 MZ pairs, 21 DZ pairs, 62 unrelated individuals

1) Change working directory

First, open Unix terminal. Copy the files to a folder of your preference.

Then, we need to make sure we are working in the right folder:

```
cd C:/Users/joseMor/Documents
```

2) Check file format

PLINK-friendly formats (.bed and .ped):

- PED: pedigree information standard
- BED: compressed (binary) version of PED

```
plink --file gwas_plinkdata --make-bed  
--out gwas_plinkdata
```

3) Clean the data (quality control)

- PLINK includes several options to *clean* genetic data.
- This means filtering out low quality data or outliers.
- We are going to run a very basic quality control (to learn more: next year at the IBG GWAS workshop!)

```
plink --bfile gwas_plinkdata --geno 0.05 --mind 0.05 --hwe  
1e-6 --maf 0.1 --make-bed --out gwas_plinkdata_clean
```

4) Estimate the genetic relatedness matrix

Output style (A) -- matrix

```
plink --bfile gwas_plinkdata_clean --make-rel  
triangle
```

Take a look to the results in Unix...

```
zless -S plink.rel
```

4) Estimate the genetic relatedness matrix

- Open the file 'GRM_highlighted.pdf' to see how it looks.
- Zoom in to find which individuals are likely to be MZ twins (in green), DZ twins (in red), and genetically unrelated (not highlighted).

4) Estimate the genetic relatedness matrix

0.994539

-0.02333

1.00109

-0.02266

-0.02092

0.995978

-0.01955

-0.01997

-0.01365

1.02602

-0.02774

-0.0197

-0.0211

-0.0292

0.985955

-0.02373

-0.02732

-0.02697

-0.0151

-0.01919

-0.02781

-0.02424

-0.02777

-0.0025

-0.01934

-0.01162

-0.01133

-0.01948

-0.02476

-0.01486

-0.01204

-0.0176

-0.02524

-0.02264

-0.02792

4) Estimate the genetic relatedness matrix

Output style (B) – relatedness pair by pair

```
plink --bfile gwas_plinkdata_clean -make-grm-  
gz no-gz
```

Take a look to the results in Unix...

```
zless -S plink.grm
```

4) Estimate the genetic relatedness matrix

- Open the file 'grel_highlighted.xls' to see how it looks.
- You can sort the data by genetic relatedness and find which pairs are likely to be MZ twins (in green), DZ twins (in red), or unrelated individuals (not highlighted).

4) Estimate the genetic relatedness matrix

ID1	ID2	common SNPs	gen relatedness
2	1	247262	-0.0233
3	1	250987	-0.0227
3	2	247011	-0.0209
4	1	247683	-0.0195
26	3	251499	0.9960
73	51	247061	0.5048

Thank you!