

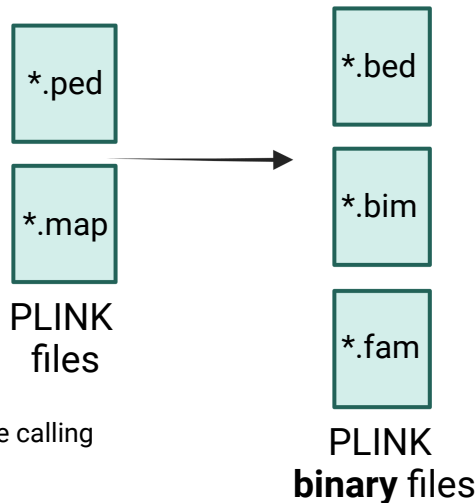
alfa study

ALZHEIMER FAMILIES

1 Collect blood samples from ALFA study participants (N = 2,640)

2 Genome-wide genotyping (Neurochip Illumina)

3 Data acquisition (N = 2,527)



*.bed
Contains binary version of the SNP info of the *.ped file

*.bim

Chr	SNP	BP	A1	A2
1	rs1	870000	C	T
1	rs2	880000	A	G
...				

*.fam

FID	IID	PID	MID	Sex	P
1	1	0	0	1	1
2	2	0	0	2	0
...					

including Manual genotype calling

5 Genetic Imputation Analysis

Michigan Imputation Server

Free Next-Generation Genotype Imputation Service

HRC r.1.1 Panel - European Population

post-Imputation QC

Data available in *.vcf format
*.vcf.dose.gz
*.vcf.info.gz

Data available in PLINK binary format
*.bim *.bed *.fam

*.bed (qc)

Contains binary version of the SNP info of the *.ped file

*.bim (qc)

Chr	SNP	BP	A1	A2
1	rs1	870000	C	T
1	rs2	880000	A	G
...				

*.fam (qc)

FID	IID	PID	MID	Sex	P
1	1	0	0	1	1
2	2	0	0	2	0
...					

updated PLINK binary files after QC

4 Quality Control (QC)

Sample (Individuals):

- Rate of Missingness
- Sex Discordance
- Heterozygosity
- Ancestry
- Causal Mutations

Genetic Variants:

- Rate of Missingness
- Hardy Weinberg
- Minor Allele Frequency
- Monomorphic variants

Supplementary Figure S1. Pipeline of genetic data processing.