**Sample-level filtering**

A

Initial sample

n = 487'409

Ambiguous sex

n = 85

Sample after ambiguous sex filtering

n =

Data-fields filtering

n = 57’450

SNPs after call-rate filtering

p =

Sample after filtering based on UK Biobank data-fields

n =

Call-rate < 95%

n =

SNPs after MAF filtering

p =

r2 < 0.04

p =

Sample after call-rate filtering

n =

SNPs after LD pruning

p =

SNPs after HWE filtering

p =

HWE > 10-6

p =

SNPs after HWE filtering

p =

**SNP-level filtering**

B

Initial SNPs

p = 730’150

MAF < 1%

p =

Call-rate < 95%

p =

Duplicates

p =

SNPs after removing duplicates

p =