How to impute chromosome X (Case-Control Studies)

Last changes: 24.08.2024
Swapnil Awasthi, Ph.D. student, swapnil.awasthi@charite.de
Prof. Stephan Ripke, M.D., Ph.D., Group Leader, stephan.ripke@charite.de
GResu (GWAS Research Unit, BIH, Berlin)

These are the instructions for chromosome X imputation of case-control studies. For Chr-X Trio-imputation please refer to this tutorial trios imputation.

Since it uses ricopili pipeline it assumes the datasets being QCed with ricopili (especially the famfile-format)

a. Extract chromosome X from ‘abc-qc.*’ using below plink command

```
plink \
--bfile abc-qc \
--chr X
--make-bed \
--out abc-qc.chrX
```

b. check if there are enough SNPs for imputation.

```
wc -l  abc-qc.chrX.bim
```

There should be at least 5000 SNPs, (if so go to Step ii) if not then revisit unQCed data and **restrict to only those individuals that passed autosomal QC** and then extract chromosome X SNPs (using above command). Please also apply additional QC if necessary.
Eg.

```
plink --bfile  abc.chrX.bim \
--hwe 1e-6 \
--geno 0.02 \
--maf 0.05 \
--make-bed --out abc-qc.chrX
```

c. Replace the code for chr X from “23” to “X” in *bim files using the below command. This step might not be necessary if you used `--chr X` to extract chromosome X in step a.

```
awk '$1="X"' abc-qc.chrX.bim > abc-qc.chrX.bim.tmp
rm abc-qc.chrX.bim
mv abc-qc.chrX.bim.tmp abc-qc.chrX.bim
```

d. Split ‘ms.abc-qc.chrX’ into male and female into a separate directory using plink.

Separate males

  i. `mkdir abc_males`
  ii. `cd abc_males`
  iii. Separate males from ‘ms.abc-qc.chrX′ using plink --filter-males

```
plink
--bfile  ../abc-qc.chrX
--filter-males
--make-bed
--out abc-qc.chrX.mal
```

Separate females

  iv. `cd ../`
  v. `mkdir abc_females`
  vi. `cd abc_females`
  vii. Separate females from ‘ms.abc-qc.chrX’ using plink --filter-females

```
plink
--bfile  ../abc-qc.chrX
--filter-females
--make-bed
--out abc-qc.chrX.fem
```

e. Run the below commands for male and female chr-X imputation in their relevant directories

  i. CHR-X MALES IMPUTATION

```
impute_dirsub
--refdir
/psych/ripke/imputation_references/HRC_EGAZ000012
39289_2016a/chr23/
--minimac3
--minilong
```

```
--impwallinc 6
--out abc_males_caco
--chr23male
```

ii. CHR-X FEMALES IMPUTATION

```
impute_dirsub
--refdir
/psych/ripke/imputation_references/HRC_EGAZ000012
39289_2016a/chr23/
--minimac3
--minilong
--impwallinc 6
--out abc_females_caco
```

Note: Multiple data sets can be split into male and females into separate directory as shown above and can be imputed together.