Revisit Population-based and Family-based Genotype Imputation

Ching-Ti Liu[1], Xuan Deng[1], Virginia Fisher[1], Nancy Heard-Costa[2], Hanfei Xu[1], Yanhua Zhou[1], Ramachandran S. Vasan[2,3], L Adrienne Cupples[1,3]

[1] Department of Biostatistics, Boston University School of Public Health, Boston MA 02118

[2] Department of Medicine, Boston University, Boston MA 02118

[3] Framingham Heart Study, Framingham MA 01702

Corresponding authors:

Ching-Ti Liu, Ph.D.

Associate Professor, Department of Biostatistics, School of Public Health, Boston University,

801 Massachusetts Ave CT3, Boston, MA 02118, USA

Tel: +1 6176387752

Fax: +1 6176386484

Email: ctliu@bu.edu

L Adrienne Cupples, Ph.D.

Professor, Department of Biostatistics, School of Public Health, Boston University,

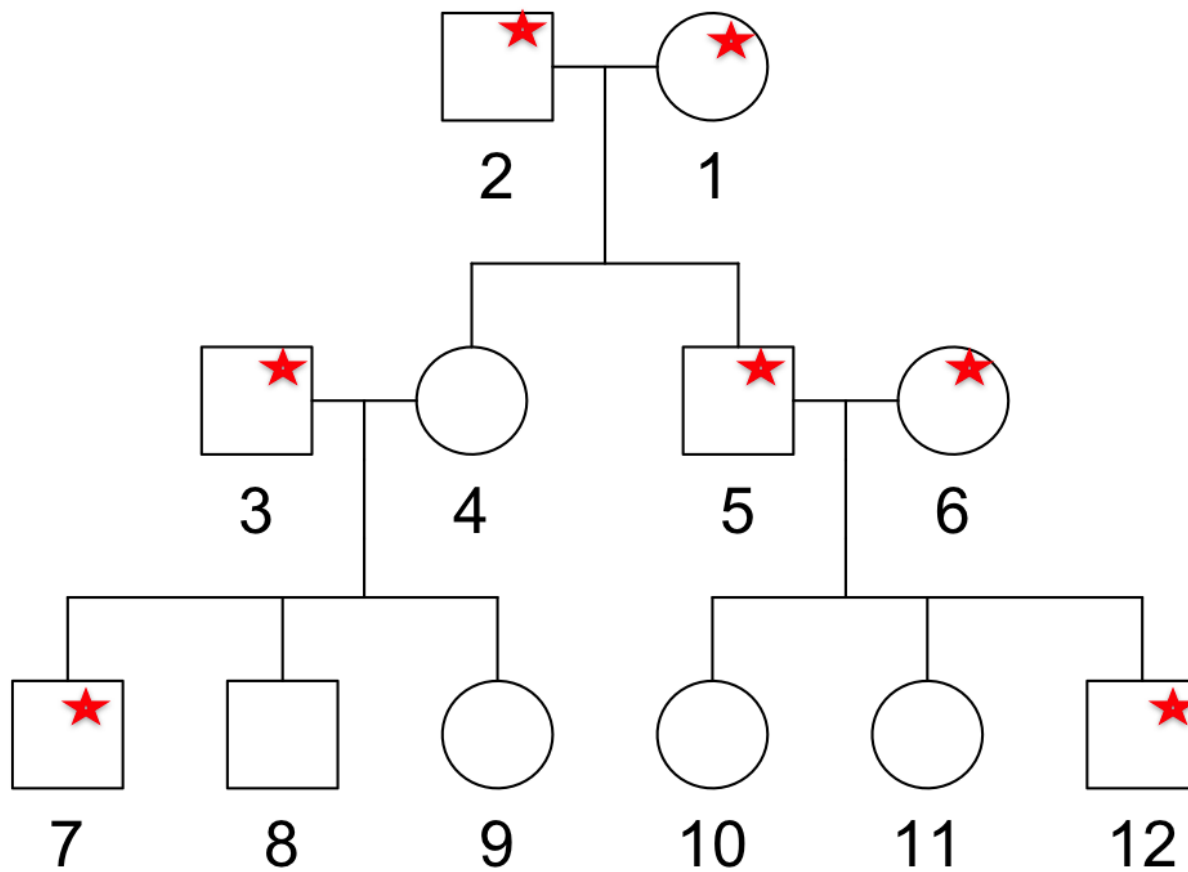801 Massachusetts Ave CT3, Boston, MA 02118, USA

Tel: +1 6176385176

Fax: +1 6176386484

Email: adrienne@bu.edu

Supplemental Figure 1. Pedigree structure and informative individuals for pedigree of size 12.

Supplemental Figure 2. Comparison of imputation accuracy with squared correlation, **Supplemental Figure 2a**,   and imputation quality score, **Supplemental Figure 2b**,   among various imputation strategies under the scenario that genotype information for the first generation was completely imputed based on the second and third generations' genotype information for 1000 families of size 12. The x-axis in the figure is on a logarithmic scale.

Supplemental Figure 1

Informative individuals obtained from GIGI-Picks

Supplemental Figure 2

(a)

Squared Correlation

0.8

0.4

0.0

0.05   0.50   5.00   50.00

Minor Allele Frequency (MAF)

.... Merlin
— — GIGI

(b)

IQS

0.8

0.6

0.4

0.2

0.0

0.05   0.50   5.00   50.00

Minor Allele Frequency (%)

.... Merlin
— — GIGI