



### Flowchart of data merging

SL: Sri Lanka, SGVP: Singapore Genome Variation Project, 1KG: 1000 genomes project, GW: genome-wide, PCA: principal components analysis,  $F_{ST}$ : Wright's Fixation Index.

\*Mapped to wrong position in CHS, \*\*Based on discordance of minor allele frequencies (MAF) between SGVP CHS and 1KG CHS (see Supplemental Digital Content 3 for more details), †6 SNPs were flagged based on  $F_{ST} > 0.05$  between SL and STU, and re-genotyped using Sanger

sequencing or Taqman genotyping. Those with discordant genotypes were removed for PCA and STRUCTURE analysis, but MAFs of 3 SNPs from Taqman genotyping were included for  $F_{ST}$  analyses. See Supplemental Digital Contents 4 & 5 for more details.