Method about the SNP block data structure and circled shuffling algorithm.

In concept, we build SNP-blocks for each individual, link the last SNP-block to the first one to get a circled data structure, perform circle shuffling, re-open the circle, and get a new shuffled SNPs array.

For each SNP-block data structure, it contains two variables:

a) Type of the SNPs within the block (0, 1, or 3).

b) Number of SNPs.

1. Read in the raw SNP data from txt documents. Build SNP-blocks for each individual, then put the SNP-blocks into an ArrayList. Every individual would share 229860 SNPs in total, but they would have different number of SNP-blocks.

2. After we got the array of SNP-blocks, we performed Fisher-Yates shuffle for 100 times for each individual. And then, another random permutation shuffling, pick a random SNP-block in the array, move it to a new ArrayList, repeat till the last SNP-block is moved to the new ArrayList.

3. After shuffling process for the individual, expand the SNP-block data structure to SNPs, pick a random pivot position, break the SNP array into two parts, and switch these two parts, in the end we get a new sequence of SNPs.

Do these 3 steps for all 565 individuals, and then merge all 565 people’s shuffled SNP sequences into a matrix (229860\*565), where each column represents a shuffled data for one individual. Print out the matrix into a txt document for MixScore software to process.

Pass the shuffled matrix together with ***admix.pheno***, ***globalancestry.theta*** to the **MixScore** software.

MixScore parameters:

***nsamples:565***

***nsnps:229860***

***genofile:nothing***

***ancfile:/work/AndrewGroup/ADM\_test/ADM\_Statistic\_Data/CircleShuffle/1matrix.txt***

***phenofile:/work/AndrewGroup/ADM\_test/AdmixScore/admix.pheno***

***thetafile:/work/AndrewGroup/ADM\_test/AdmixScore/globalancestry.theta***

***orfile:1130\_or\_1.out.or***

***outfile:mixscore\_shuffled1.out***

The paper leaded me to circled shuffling: <http://www.nature.com/nbt/journal/v29/n11/pdf/nbt.2023.pdf>

Fish-Yates shuffle wikilink: <https://en.wikipedia.org/wiki/Fisher%E2%80%93Yates_shuffle>