CSE 280A
Assignment 3
List 6
Andrey Bzikadze

Q: "Is it possible to devise a statistic based on long haplotype lengths?"

As I asked on Piazza, I'm not too sure what should the answer look like.

Generally if reads do not capture more than one SNV, then the analysis becomes much more complicated — we have almost no information on dependency of various SNV. Something can be done if the coverage is close to uniform —

mut. \ pos:	i	j
0	25	50
1	75	50

=> at least 25 indiv. have both i and j mutation

But this analysis is very limited. If the coverage is far from the uniform, the intersections will be even less powerful!