

Applied population genetics

Visualisation and understanding

BIO392 practical session

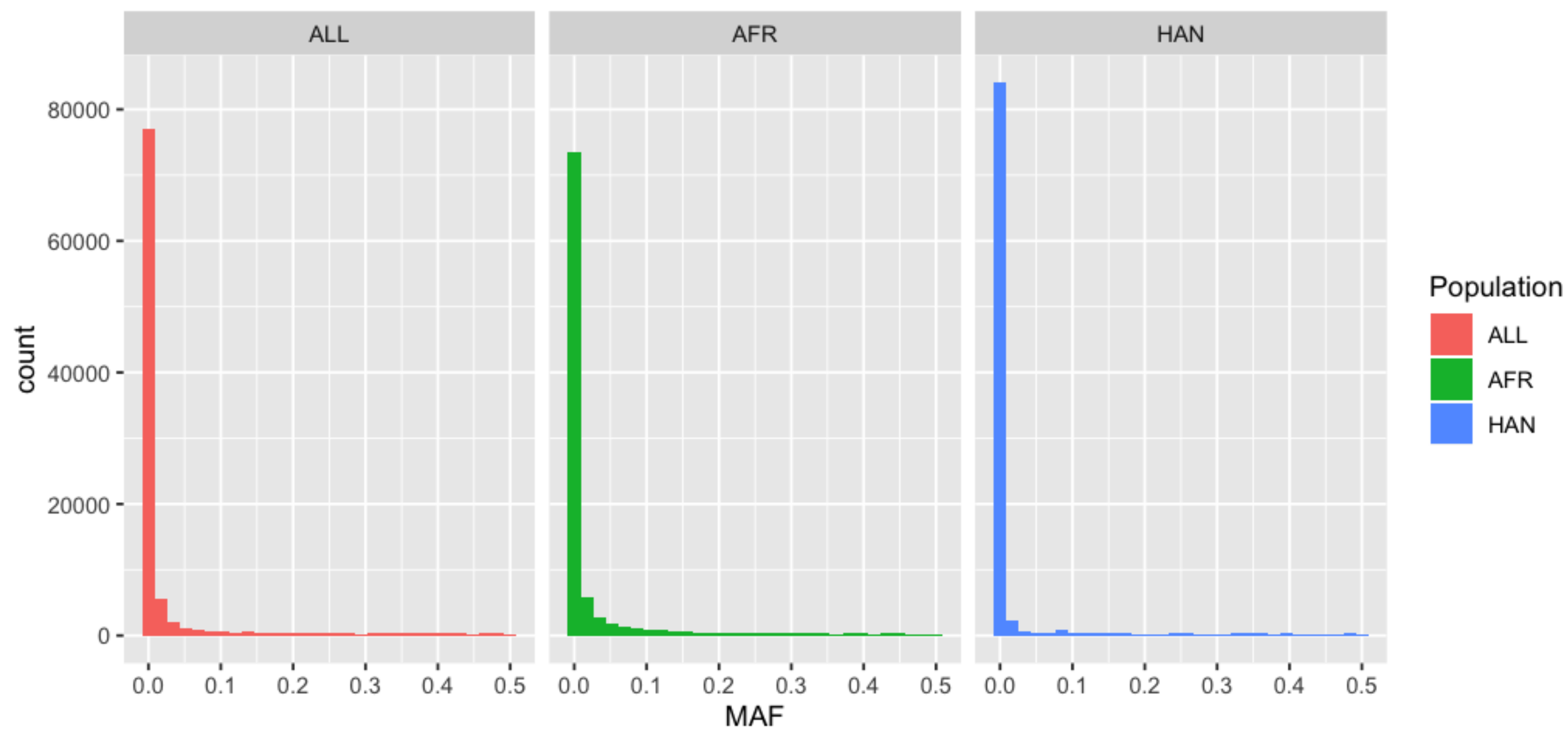
Qingyao Huang

Data sets (from 1000 Genomes)

- All populations together (2504 individuals)
- African population (661 individuals)
- Han Chinese population (103 individuals)
- Only chromosome 20, small chunk 1- 3.2Mb

Histogram of minor allele frequency

- Use the plink.frq file and plot histogram in R
- How do they differ?



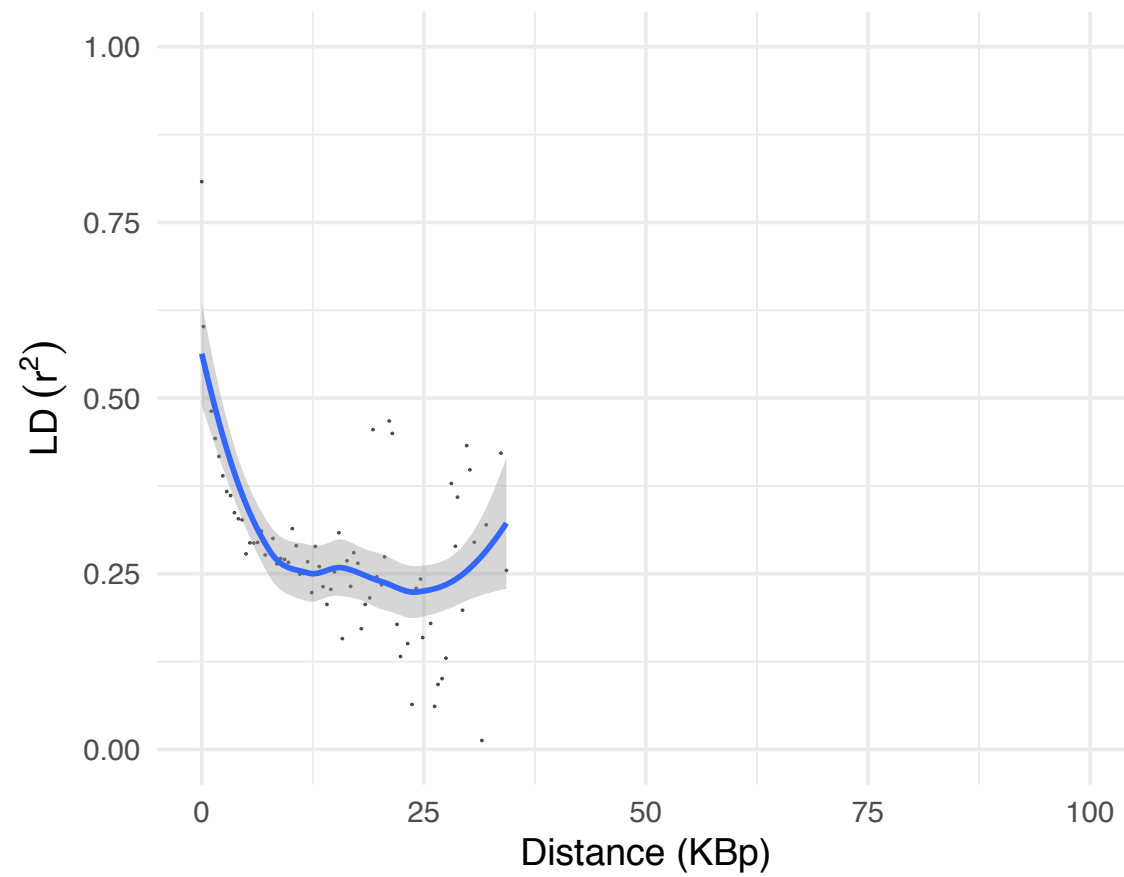
LD decay

- Make use of the R function - `plotDecayLD`
- How do the 3 populations differ?

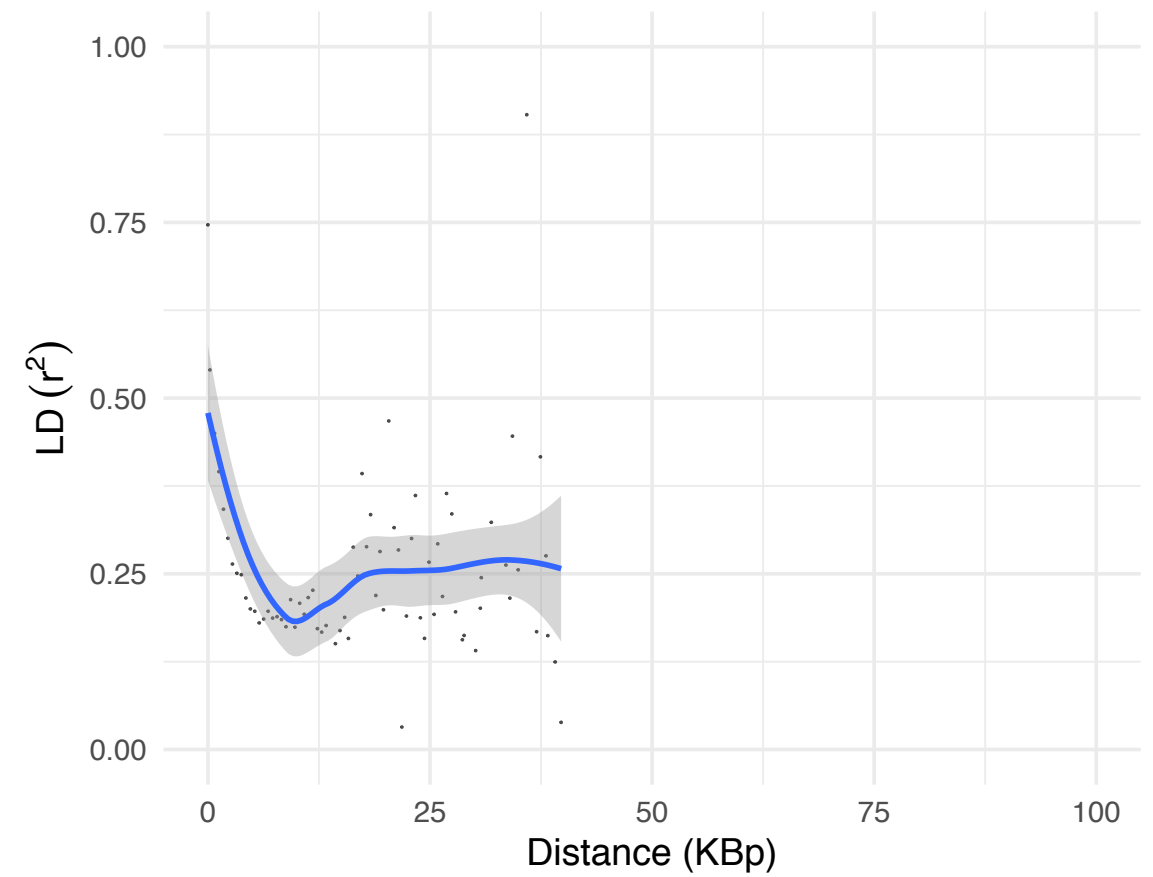
Pairwise LD plot

- Use the region 2 - 3Mbase
- Make use of the R function - `plotPairwiseLD`
- How do the LD plots differ?

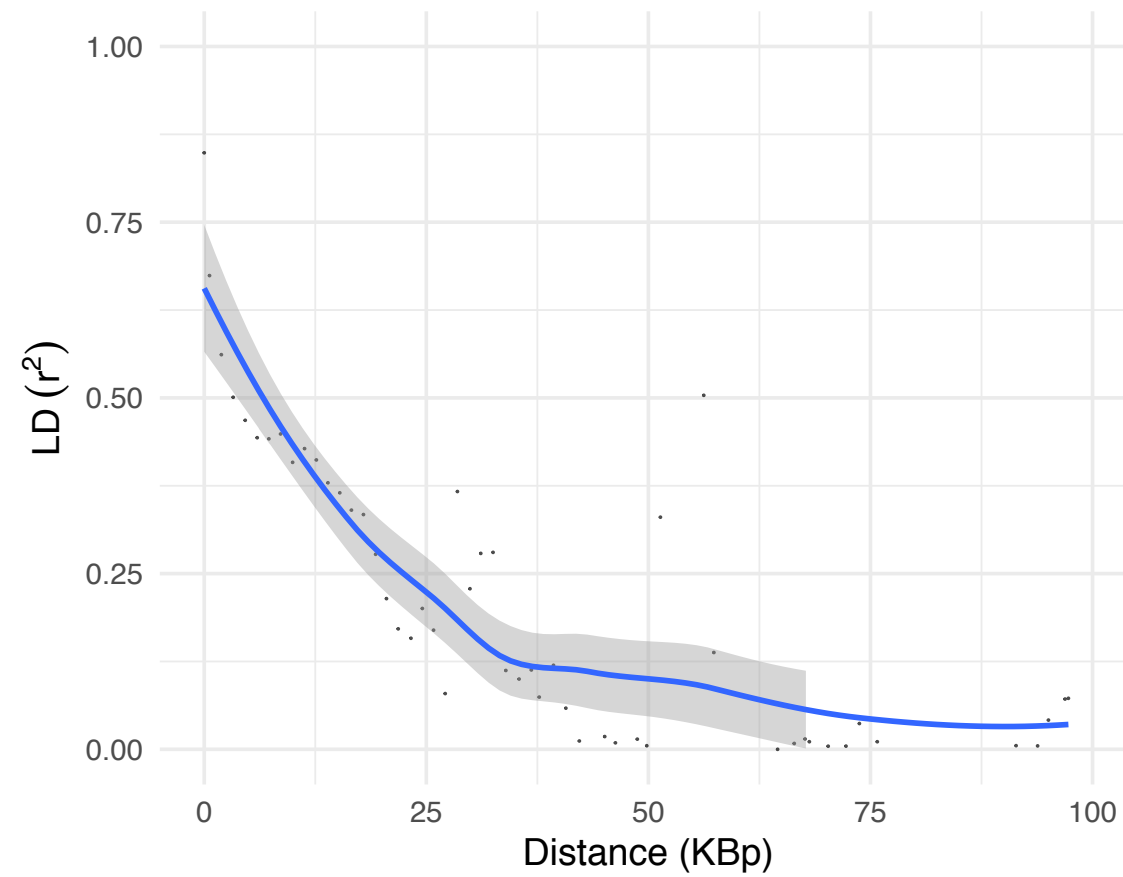
LD decay All



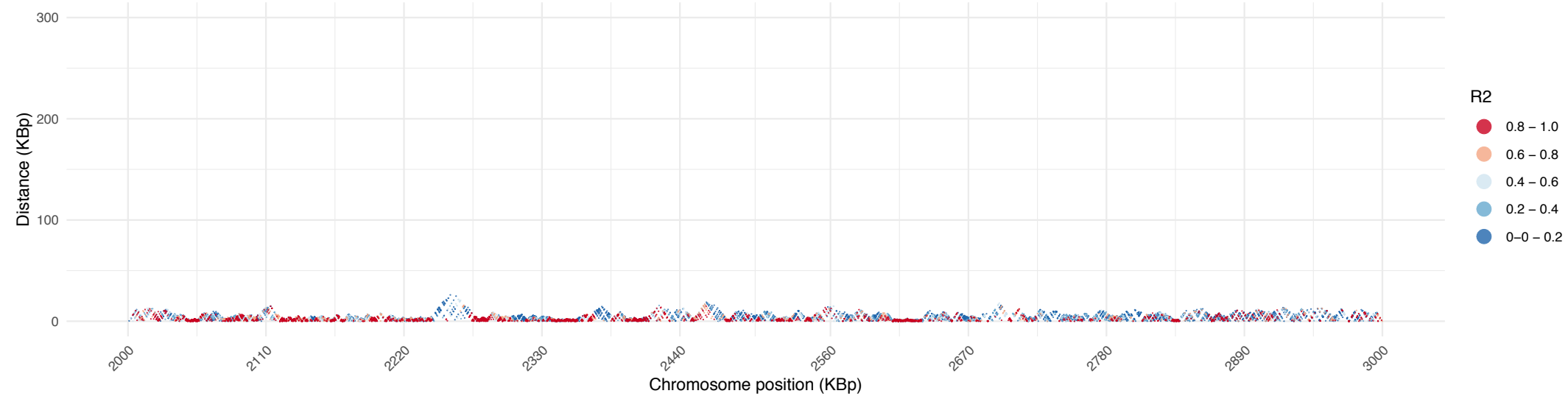
LD decay African



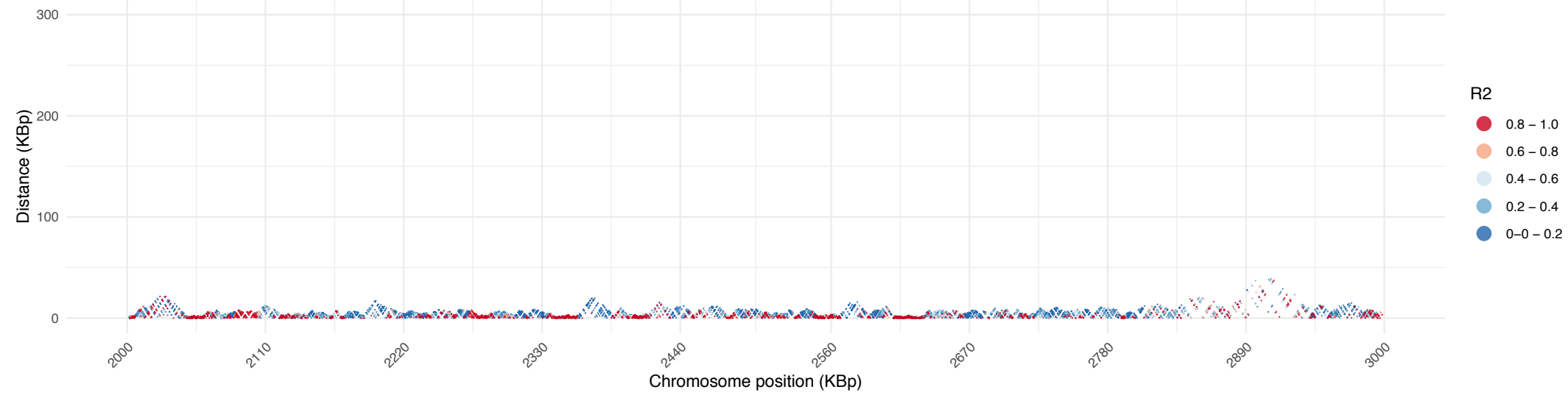
LD decay Han Chinese



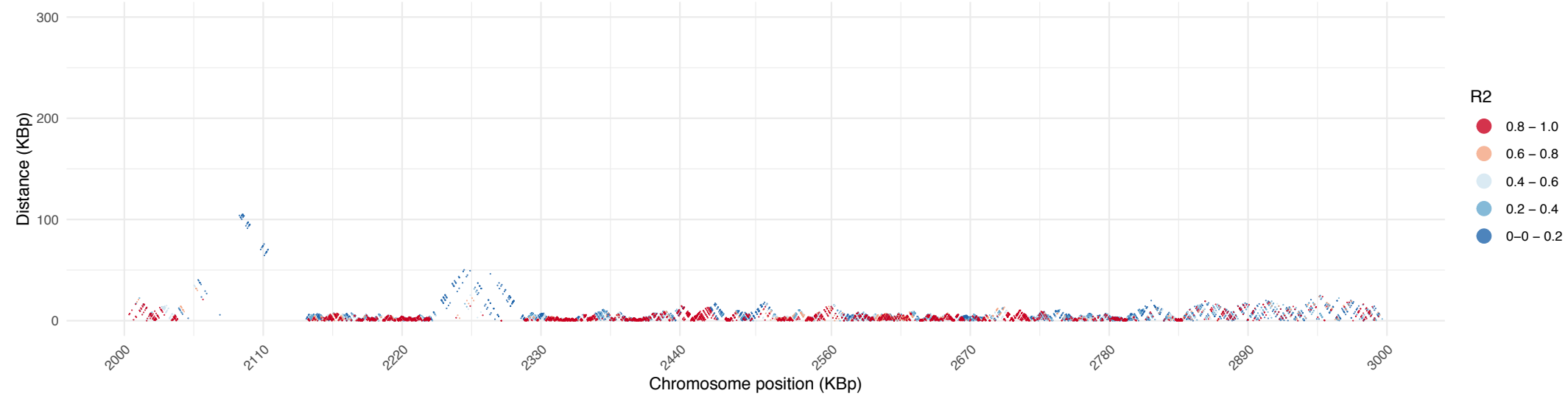
LD plot All



LD plot Arican



LD plot Han Chiense



Sum-up

- Minor allele frequency (MAF) - proportion rare variants
- LD decay - rate shows diversity
- Pair-wise LD plot - linked regions