R script: 14\_Fig8b\_LDplot.R

Need: 8kb region on Chromosome 16 from337285 To 355042

on linux:

bgzip Documents/GitRepos/BcSolGWAS/data/genome/big\_set\_v97iso\_SNPs\_filtered\_qual30\_dp6\_maf20\_recode.vcf

install vcftools

vcftools --gzvcf Documents/GitRepos/BcSolGWAS/data/genome/big\_set\_v97iso\_SNPs\_filtered\_qual30\_dp6\_maf20.recode.vcf.gz --chr Chromosome16 --recode --out Documents/GitRepos/BcSolGWAS/data/genome/chr16\_analysis/chr16\_analysis

this keeps only 12309 sites, much more manageable.

I’ll also make a shorter fragment with only 8kb around the gene of interest:

vcftools --gzvcf Documents/GitRepos/BcSolGWAS/data/genome/big\_set\_v97iso\_SNPs\_filtered\_qual30\_dp6\_maf20.recode.vcf.gz --chr Chromosome16 --from-bp 342785 --to-bp 349542 --recode --out Documents/GitRepos/BcSolGWAS/data/genome/chr16\_analysis/chr16\_analysis\_seg

then go from abbreviated vcf to DNAbin (with full vcf, this crashes Rstudio on my PC)

but, pegas has a dependency problem on ubuntu → run on PC

<https://knausb.github.io/vcfR_documentation/dnabin.html>

then extract haplotype from DNAbin file

<https://www.rdocumentation.org/packages/pegas/versions/0.10/topics/haplotype>

Plotting: snp.plotter

<https://cran.r-project.org/web/packages/snp.plotter/vignettes/using_snp_plotter.html>

<https://cran.r-project.org/web/packages/snp.plotter/index.html>

[https://academic.oup.com/bioinformatics/article/23/6/774/415503/snp-plotter-an-R-based-SNP-haplotype-association#6145357](https://academic.oup.com/bioinformatics/article/23/6/774/415503/snp-plotter-an-R-based-SNP-haplotype-association" \l "6145357)