

## Prática 3

### Requisitos

-plink

- GCTA

Instale e carregue os seguintes pacotes no R:

"qqman", "ggplot2"

### Atividades

- 1) Selecione a pasta de trabalho.
- 2) Leia o arquivo de fenótipo "Data.txt". Prepare dois arquivos de fenótipo (Sobrevivência Binária e Dia de Morte) que correspondam apenas a animais que possuem genótipos (use "SRS.fam" para fazer a correspondência). Prepare também os arquivos de covariáveis ("Idade" e "Tag\_Weight"). Todos os arquivos devem estar de acordo com o que é suportado pelo GCTA.
- 3) Realizar o controle de qualidade dos dados genômicos (Call-rate SNPs: 90%, MAF: 0,05, HWE (p-valor corrigido por Bonferroni); Call-Rate de amostras 90%) e gerar os arquivos de genótipo finais no formato binário de plink.
- 4) Com o GCTA, construa a matriz de parentesco genômico (GRM) aplicando a função de matriz para dados familiares.
- 5) Estimar a herdabilidade usando o algoritmo REML para o fenótipo de sobrevivência binária usando o GCTA. Use as covariáveis "Idade" e "Tag\_weight" no modelo.
- 6) Implementar a análise de associação (GWAS) para o fenótipo binário de sobrevivência usando o GCTA. Use as covariáveis "Idade" e "Tag\_weight" no modelo linear misto (MLMA).
- 7) Faça um gráfico de Manhattan que mostre o significado dos SNPs transformados em  $-\log_{10}$ . Aplique as linhas de significado cromossômico e genômico para identificar os SNPs com maior associação.

8) Se SNPs significativos forem encontrados, pesquise no site do NCBI ([https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF\\_002163495.1](https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_002163495.1)) genes próximos aos SNPs usando janelas de 200 Kb abaixo e acima do SNP como critério.

9) Repita as etapas 5, 6, 7 e 8 usando o dia da morte como fenótipo.