Potato Color Genetics

immediate

May 10, 2025

0.1 2.1 Genotypic data

....

To ensure data quality and reliability for downstream GWAS analysis, standard filtering criteria were applied. SNPs with a minor allele frequency (MAF) below 1% were excluded to eliminate rare variants that may lead to spurious associations due to insufficient statistical power (Anderson et al., 2010). Individuals with a genotype call missing rate (MIND) greater than 10% were removed to avoid bias introduced by low-quality samples. SNPs with a missing rate (GENO) above 10% were filtered out to maintain marker integrity and reduce imputation noise. Additionally, SNPs with Hardy-Weinberg equilibrium (HWE) p-values below 1e-10 were excluded, as extreme deviations may indicate genotyping errors, population structure artifacts, or selection (Wigginton et al., 2005). These thresholds are commonly used in polyploid GWAS studies and are consistent with established quality control practices (Lu et al., 2013; Pavan et al., 2020).

References

Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nature Protocols*, *5*(9), 1564–1573. https://doi.org/10.1038/nprot.2010.116

Wigginton, J. E., Cutler, D. J., & Abecasis, G. R. (2005). A note on exact tests of hardy-weinberg equilibrium. *The American Journal of Human Genetics*, *76*(5), 887–893. https://doi.org/10.1086/429864

Lu, F., Lipka, A. E., Glaubitz, J., Elshire, R. J., Cherney, J. H., Casler, M. D., Buckler, E. S., & Costich, D. E. (2013). Switchgrass genomic diversity, ploidy, and evolution: Novel insights from a network-based snp discovery protocol. *PLoS Genetics*, *9*(1), e1003215. https://doi.org/10.1371/journal.pgen.1003215

Pavan, S., Delvento, C., Ricciardi, L., Lotti, C., Ciani, E., & DâAgostino, N. (2020). Recommendations for choosing the genotyping method and best practices for quality control in crop genome-wide association studies. *Frontiers in Genetics*, *11*, 447. https://doi.org/10.3389/fgene.2020.00447