Exploring Genetic Diversity for Common Bean Improvement: Dissecting Domestication Traits and Drought Tolerance Strategies through Population Structure and QTL Analysis

Submitted by:

Kate Elizabeth Denning-James

September 2025

University of East Anglia for the degree of Doctor in Philosophy

Earlham Institute

Abstract

Common beans, *Phaseolus vulgaris*, are an important crop for food and nutritional security and the ecosystem services they provide, especially in developing countries. They are adapted to a broad range of ecogeographic conditions, therefore are genetically diverse and have a complex population structure. This includes two wild gene pools, an ancient-relic population, two domesticated gene pools and races within each gene pool. There is frequent gene flow, for example, between the wild and domesticated populations, the domesticated gene pools and between the races.

The two gene pools (Andean and Mesoamerican) overlap in Colombia. However, the Colombian common bean population structure and level of admixture have not been defined. This research is testing the hypothesis that Colombia has a large amount of genetic admixture diversity and introgressed lines, compared to the 'ancestral' Andean and Mesoamerican gene pools.

This genetic diversity provides a reservoir of adaptive genes that could be used in genetic research and crop improvement. Mobilising genetic diversity into elite backgrounds could help to reduce the effects of climate change on yields, protecting food security. Exploring the genetic diversity in landraces, and linking to phenotypes, could improve tolerance to abiotic stresses in elite cultivars and help safeguard future food security through breeding programs.

144 common beans have been sequenced, primarily landraces from Colombia but also wild accessions and heirlooms from neighbouring countries. The sequence data has been analysed to infer the population structure and understand the evolution of the diversity panel. The PhD also involves phenotyping for domestication traits and tolerance to water-deficit to identify molecular markers associated with the traits in common beans via GWAS. The results will support future breeding programs to reduce the effects of climate change on common bean yields.

Access Condition and Agreement

Main body word count: 39,317

Contents

Abstract ***2***

Access Condition and Agreement ***3***

Contents ***4***

Figures ***8***

Tables ***16***

Acknowledgements ***17***

Chapter 1

Literature Review ***18***

1.1 Agrobiodiversity **19**

1.2 Common Beans, *Phaseolus vulgaris* **21**

1.3 Importance of Common Beans **22**

1.3.1 Consumption and economics 22

1.3.2 Nutritional Qualities 23

1.3.3 Ecosystem services 24

1.4 Population Structure of Common Beans **25**

1.4.1 Wild common beans 25

1.4.2 Domestication 27

1.4.3 Domestication syndrome traits 28

1.4.4 Races within gene pools 29

1.4.5 Dissemination of Common Beans 30

1.4.6 Gene flow 30

1.5 Exploring genetic diversity to respond to climate change **32**

1.6 Global water scarcity **32**

1.6.1 Drought definitions 33

1.6.2 Drought stress in agricultural crops 33

1.6.3 Common beans and drought stress 35

1.7 Breeding programmes **36**

1.8 Colombian and Admixed Common Beans **36**

1.9 Aims and Objectives of the Thesis **37**

Chapter 2

Selective breeding for determinacy and photoperiod sensitivity in common bean (Phaseolus vulgaris L.) ***39***

2.1 Abstract **40**

2.2 Introduction **41**

2.2.1 Importance of Common Beans 41

2.2.2 Common Bean population structure 41

2.2.3 The Diversity Panel 42

2.2.4 Determinacy and Photoperiod Insensitivity 44

2.2.5 Admixed diversity in Colombia 45

2.2.6 Project aims 46

2.3 Methods **47**

2.3.1 Genotyping 47

2.3.2 Phenotyping 47

2.3.3 Pre-processing genotype data 49

2.3.4 Population structure analysis 50

2.3.5 Introgression analysis 51

2.3.6 Genome wide association study 52

2.3.7 Selecting significant loci, candidate gene mining and functional annotation 52

2.3.8 Mean allele dosage 53

2.4 Results **54**

2.4.1 Alignment to reference genomes 54

2.4.2 Population structure 55

2.4.3 Introgression analysis 63

2.4.4 Phenotypic variation and correlations 78

2.4.5 Environmental analysis 81

2.4.6 GWAS for Determinacy 82

2.4.7 GWAS for Photoperiod Sensitivity 85

2.4.8 Mean allele dosage from QTLs for determinacy 87

2.4.9 Mean allele dosage from QTLs for photoperiod sensitivity 96

2.5 Discussion **104**

2.5.1 Environmental analysis 104
2.5.2 Alignment and heterozygosity 105
2.5.3 Population Structure and phenotyping 106
2.5.4 Introgressions and population structure 108
2.5.5 GWAS results 110
2.5.6 QTLS and Candidate Genes Associated with Determinacy 110
2.5.7 QTLS and Candidate Genes Associated with Photoperiod Sensitivity 113
2.5.8 Mean allele dosage 116

2.6 Conclusion Error! Bookmark not defined.

Chapter 3

Natural variation in Colombian common beans reveals diverse water deficit response strategies and genetic loci through GWAS and phenotyping 118

3.1 Abstract 119

3.2 Introduction 120
3.2.1 Common Beans, Phaseolus vulgaris 120
3.2.2 Drought stress in agricultural crops 120
3.2.3 Common beans and drought stress 121
3.2.4 Research aims 121

3.3 Methods 122
3.3.1 Diversity panel and genotyping 122
3.3.2 Experimental design 122
3.3.3 Environmental conditions 125
3.3.4 Phenotyping, data collection and data analysis 127
3.3.5 Genome wide association studies 128
3.3.6 QTL identification and candidate gene prediction 129

3.4 Results 130
3.4.1 Soil conductivity and moisture assessments 130
3.4.2 Phenotypic variation and correlations 133
3.4.3 Drought stress strategies 137
3.4.4 GWAS based on drought responses 142

3.5 Discussion 149
3.5.1 Indicators of drought stress 149

3.5.2 Drought stress strategies 151
3.5.3 Recovery from water deficit 154
3.5.4 Candidate genes 155

3.6 Conclusions 157

Chapter 4

Future work and Further Discussion 158

4.1 Key findings chapter 2 158

4.2 Key findings chapter 3 159

4.3 Future Implications 160

Appendices 163

Appendix 1; Subpopulation and phenotypic annotation for the 144 common bean accessions used in studying determinacy and photoperiod sensitivity. 163

Appendix 2; Genes and non-synonymous genetic variants within each of the identified QTLs associated with determinacy. 163

Appendix 3; Genes and non-synonymous genetic variants within each of the identified QTLs associated with photoperiod sensitivity. 163

Appendix 4; All phenotypes and transformations input into GAPIT to identify significant QTLs for tolerance to water-deficit. 163

Appendix 5; The phenotypes input into the correlation plot, after water application was stopped. 163

Appendix 6; All genes found within the 36 QTLs, their locations, if SNPeff high mutations are present and their annotations. 163

Appendix 7; TAIR and Phytomine annotations for all genes within the QTLs for water deficit response and recovery. 163

Appendix 8; Publication 164

Glossary 180

References 182

Figures

Figure 1; Diagram showing the elements and interactions within agrobiodiversity from Brookfield and Stocking (1999). 19

Figure 2; The split of the wild Mesoamerican and Andean common bean gene pools. Highlighting the asymmetric split and bottlenecks of the wild populations from Schmutz et al. (2014). 26

Figure 3; Distribution of the 127 common beans with location data from the diversity panel. The coordinates of the capital city were used for those without coordinate data. Produced with QGIS. Figure from Denning-James et al. (2025). 43

Figure 4; Frequency distribution of phenotypic traits for the diversity panel (144 accessions). (A) Major seed colour, (B) type of accession, (C) country of origin or collection and (D) estimated seed size. Data collated from corresponding genebank websites and phenotypic studies during 2022 (Appendix 1). 44

Figure 5; Examples of the three different growth habits for common beans. (A) A determinate bush (PHA13862), (B) an indeterminate bush (G5910) and (C) an indeterminate climbing growth habit (G50516K). Photos taken at harvesting during September 2023. Black arrows are equivalent to 20cm. 45

Figure 6; Layout of the glasshouse (144 accessions) from phenotyping trials in summer 2022 (A) and winter 2022 (B). 48

Figure 7; A scatter plot representing the alignment of the diversity panel (144 accessions) to the Mesoamerican and Andean reference genomes. The colours represent the type of accession from gene bank passport data (Appendix 1). 54

Figure 8; Analysis of the population structure of 144 accessions belonging to our diversity panel focusing on Colombia at K=2, Andean or Mesoamerican groups (A), K=4 (B) and K=6 (C) and the cross-validation error for different K values (D) ran using ADMIXTURE (v 1.3.0)(Alexander and Lange 2011). (C-EP) accessions mainly from Peru, then Ecuador and Colombia; (A1) Andean accessions from a variety of South

American countries; (C1) mostly determinate Colombian landraces; (C2) indeterminate Colombian landraces; (M1) mainly medium seeded** from Central America and Colombia; (M2) mainly small seeded** from Central America and Colombia. (Admx_AM) Andean X Mesoamerican hybrids; (Admx_A) and (Admx_M) admixed accessions between subpopulations (ancestry composition q <0.7 at K=6). ** p < 0.01 using a two tailed student t-test with unequal variance. Plots A and C from Denning-James et al. (2025). 56

Figure 9; Analysis of the population structure of 144 accessions belonging to the diversity panel at K=2 (A) and K=6 (B) and the Delta K for K values 2-10 (C) ran using STRUCTURE v 2.3.5 (Pritchard et al. 2000). (A, B) Structured by ADMIXTURE K=6 results as in Figure 8 (Alexander and Lange 2011). 58

Figure 10; Principal component results on all 144 accessions from the diversity panel from GAPIT v.3 (Wang and Zhang 2021). (A) Principal component analysis (PCA) plot of PC1 against PC2 and (B) plot of PC1 against PC3. The colours illustrate the population structure of the diversity panel. PC1 against PC2 coloured by (C) photoperiod insensitivity, (D) determinacy, (E) estimated seed size and (F) country of sample collection (Appendix 1). 60

Figure 11; (A) Proportion of heterozygous sites against the percentage of read-pair alignment to the Andean reference genome G19833 for 144 accessions (Schmutz et al., 2014). (B) Principal component results on the 108 Andean accessions from the diversity panel ran using GAPIT v.3 (Wang and Zhang 2021). The colours illustrate the population structure of the diversity panel from ADMIXTURE K=6 (Alexander and Lange 2011). 61

Figure 12; Introgression analysis on all 144 accessions from the diversity panel using VcfHunter (Martin et al. 2023) split by ADMIXTURE K=6 population structure. (A) M1, (B), M2 and admixed M, (C) admixed between Andean and Mesoamerican, (D) A1, (E) admixed A, (F) C1 and C2 and (G) C-EP. The colours represent the 'ancestral' background, and the staggered columns (y-axis) reflects the proportion (0-1) of SNPs in that window across the genome (x-axis) from each background, calculated

for each accession. Accession name and subpopulation names at, K=4 and K=6 population structure from ADMIXTURE, respectively, were labelled. 69

Figure 13; Comparison of genetic landscape metrics, namely, proportion of Andean, Mesoamerican, heterozygous or unknown background content, among subpopulations for the diversity panel (144 accessions). (A) Comparison across the whole genomes, (B) comparisons calculated per chromosome. Grouping by the admixture K=6 population structure. Significance was calculated with the Dunn test and Bonferroni correction (* p <0.05, ** p<0.01, *** p<0.001). 77

Figure 14; Correlation plots produced in R using the 'corrplot' package (Wei and Simko 2021). Phenotypes are agronomic traits measured in 144 common bean genotypes grown at the Norwich Research Park, Norwich, UK in 2022 and 2023, results of the population structure analysis and data from gene bank passport data (IPK and CIAT) (Appendix 1). (A) Associations between discrete and continuous phenotypes assessed with a one-way ANOVA and measured with eta-squared. (B) Correlation between discrete datasets, the association was calculated with Cramer Vs and significance with Chi-squared (Mangiafico. 2025). (C) Pearson correlation coefficients among continuous datasets with normal distribution. (K6) K6 subgroups from ADMIXTURE; (K2) K2 subgroups from ADMIXTURE; (D) Determinacy; (SS) Seed size; (Est_SW) Estimated weight of 100 seeds; (PS) photoperiod sensitivity; (Win_DTF) days to flowering from winter 2023; (Sum_DTF) days to flowering from summer 2022. *Indicates p < 0.05; **indicates p < 0.01; ***indicates p < 0.001. 79

Figure 15; Frequency distribution of seed weight and days to flower traits evaluated in two seasons in a common bean diversity panel (144 accessions). (A) E100_SW, estimated weight of 100 seeds; (B) phenological days to flowering in the summer 2022 (S22_DTF) (and (C) in the winter 2023 (W23_DTF) at the Norwich Research Park, excluding those which did not flower. The distributions were split into the subpopulations from K6 ADMIXTURE. (D) E100_SW***; (E) S22_DTF***; (F) W23_DTF*. Completed a one-way ANOVA for E100_SW, S22_DTF and W23_DTF.

*Indicates p < 0.05; **indicates p < 0.01; ***indicates p < 0.001. Figure from Denning-James et al. (2025). 80

Figure 16; Plot represent (A) the long daylength trial from April 2022 to July 2022 and (B) the short daylength trial from Dec 2022 to March 2023. Data for daylength collected from CW4326, Norwich, UK station (52.643°, 1.237°) (Visual Crossing Corporation 2025) and data for temperature collected from within the glasshouse with the Gemini Tinytag, Ultra 2. The vertical dotted lines represent the first accession to flower, the average time for flowering (removing those which did not) and the last plant to flower. 81

Figure 17; Manhattan plots highlighting markers significantly associated with determinacy on (A) the whole panel (n=144) and (B) the Andean subpanel (n=108). The analyses were completed with GAPIT and the models FarmCPU, BLINK or MLM (Huang et al. 2019; Liu et al. 2016; Wang and Zhang 2021; Zhang et al. 2010). The X-axis represents the genomic position of markers, and the Y-axis is the -log 10 of the P-values for association with the phenotype. The vertical lines correspond to QTLs found by at least two models. Point size correlates to -log10(P-value). Quantile-quantile (QQ) plots are provided for (C) the whole panel, (D) the Andean panel and (E) models not selected for the Andean panel. Plots A-D from Denning-James et al. (2025). 84

Figure 18; Manhattan plots highlighting markers significantly associated with photoperiod insensitivity on (A) the whole panel (n=144) and (B) the Andean subpanel (n=108). The analyses were completed with GAPIT and the models FarmCPU, BLINK or MLM (Huang et al. 2019; Liu et al. 2016; Wang and Zhang 2021; Zhang et al. 2010). The X-axis represents the genomic position of markers, and the Y-axis is the -log 10 of the P-values for association with the phenotype. The vertical lines correspond to QTLs found by at least two models. Point size correlates to -log10(P-value). Quantile-quantile (QQ) plots are provided for (C) the whole panel, (D) the Andean panel, (E) models not selected for the whole panel and (F) models not selected for the Andean panel. Plots A-D from Denning-James et al. (2025). 86

Figure 19; Heatmap of SNP mean dosage values for polymorphisms in the coding region of QTLs associated with determinacy from the GWAS of the whole (n=144) or Andean diversity panels. 0 are homozygous reference, 1 are heterozygous, 2 are homozygous alternative and 3 are uncalled genotypes (transformed NAs). Plotted with pheatmap v.1.0.13 with default Euclidean distance to cluster the rows and SNP location to order columns (Kolde 2025). The top row represents the gene which the SNPs fall within. SNPs have been averaged by their determinate trait and population structure at K2 (M; Mesoamerican, A; Andean, Admx; Admixture). SNPs with less than a 0.5 difference in variation between groups have been removed. The plots represent the whole diversity panel for QTLs (A) D1.1, (C) D1.3, (D) D1.4, (E) D1.5, (F) D1.6, (I) D7.1, (K) D8.2, (L) D9.1 and (M) D10.1. Also, the Andean diversity panel for QTLs (B) D1.2, (G) D2.1, (H) D3.1 and (J) D8.1. 95

Figure 20; Heatmap of SNP mean dosage values for polymorphisms in the coding regions in QTLs associated with Photoperiod sensitivity from the GWAS of the whole (n=144) or Andean diversity panels. 0 are homozygous reference, 1 are heterozygous, 2 are homozygous alternative and 3 are uncalled genotypes (transformed NAs). Plotted with pheatmap v.1.0.13 with default Euclidean distance to cluster the rows and SNP location to order columns (Kolde 2025). The top row represents the gene which the SNPs fall within. SNPs have been averaged by their determinate trait and population structure at K2 (M; Mesoamerican, A; Andean, Admx; Admixture). SNPs with less than a 0.5 difference in variation between groups have been removed. The plots represent the whole diversity panel for QTLs (A) PS4.1, (B) PS5.1, (D) PS7.1, (E) PS7.2, (F) PS8.1 and (H) PS9.1. Also, the Andean diversity panel for (C) PS5.2, (G) PS8.2, (I) PS9.2 and (J) PS11.1. 103

Figure 21; Layout of the diversity panel (144 accessions with 3 water deficit replications), for the water deficit experiment at NIAB, Cambridge in 2023. 38 of the accessions were determinate bush, 9 were indeterminate bush, 86 were indeterminate climbing and 11 were prostrate. 123

Figure 22; Temperature (°C), precipitation (mm) and daylength patterns (hrs) during the experiment in Cambridge, UK, from 19/06/2023 (sowing) to 09/09/2023 (harvesting)

(Visual Crossing Corporation 2024). The dotted lines represent key events during the experiment. On 27/07/2023 (38 DAS, week 1) irrigation was stopped to induce the water deficit treatment and on 25/08/2023 (67 DAS, week 5) the plants experienced a mock recovery period due to precipitation. 125

Figure 23; Bar plots of parameters collected from soil moisture sensors (Table 4). For (A) the Soil Matric Potential (kPa), (B) the soil temperature (°C), (C) the bulk electrical conductivity (EC, mS/cm) and (D) for the water content (m3/m3). The colours represent the dates from irrigation (week 0) to harvesting (week 5). Statistical analyses identified significant interactions between stress x growth habit over time. Groups were labelled by a letter and superscript number. Groups sharing the same letter were not significantly different at $\alpha = 0.05$, whereas groups with different letters differed significantly. The superscript numbers in the labels corresponded to the week. 132

Figure 24; Correlation plot comparing phenotypes derived from the poromoter and fluorometer, BBCH (phenological developmental data), D (Determinacy), PS (Photoperiod Sensitivity), Date, VPDleaf (leaf vapour pressure deficit), Tleaf (leaf temperature), Country, Type (wild, heirloom, landrace), Strategy (drought tolerance strategy, Table 5), population structure at K6 and K2 (subgroups from ADMIXTURE), RH (relative humidity), SS (seed size), ETR (electron transport rate), Fs (steady-state chlorophyll fluorescence), GSW (g_{sw}, stomatal conductance), E apparent (E, transpiration), Fm' (maximum fluorescence yield), PhiPS2 (Φ_{PSII}, quantum yield of fluorescence). Data from water deficit treated plants (142 accessions) following the cessation of water application (five recordings between days 38 to 72, Appendix 4 and 5). (A) Correlation between continuous datasets, and correlation between continuous and discrete data, where discrete datasets were put to zero. (B) Correlation between discrete datasets. 133

Figure 25; Bar plots of parameters collected from the porometer and fluorometer separated by growth habit from week 1 to week 4, for accessions experiencing water deficit (Table 3, 142 accessions; 37 determinate bush, 8 indeterminate bush, 97 indeterminate climbing and prostrate). (A) Fm' (maximum fluorescence yield),

(B) Fs (steady-state chlorophyll fluorescence), (C) Tleaf (leaf temperature, °C), (D) GSW (g_{sw}, stomatal conductance, mol $m^{-2}s^{-1}$), (E) E apparent (E, transpiration, mmol $m^{-2}s^{-1}$) and (F) PhiPS2 (Φ_{PSII}, quantum yield of fluorescence, µmol m^{-2} s^{-1}). Groups sharing the same letter are not significantly different at α = 0.05, whereas groups with different letters differed significantly. The superscript numbers correspond to the week. 135

Figure 26; (A, B) Histograms for comparative analysis of the diversity panel undergoing water deficit stress treatment (n=142). Histograms for (A) pod weight (g) and (B) foliar weight (g). (C, D) Violin plots for five weeks of recording water-deficit and recovery (Table 3) for (C) BBCH (development) and (D) ETR (electron transport rate, µmol m^{-2} s^{-1}). Grouped by the drought tolerance strategies identified in Table 5. 140

Figure 27; Violin plots comparing leaf porometer and fluorometer responses between water deficit (Week 4) and recovery (Week 5) for all accessions (n=142). (A) ETR (electron transport rate, µmol m^{-2} s^{-1}), (B) Tleaf (leaf temperature, °C), (C) E apparent (E, transpiration, mmol $m^{-2}s^{-1}$), (D) PhiPS2 (Φ_{PSII}, quantum yield of fluorescence, µmol m^{-2} s^{-1}), (E) BBCH (development), (F) Fs (steady-state chlorophyll fluorescence), (G) RH (relative humidity, %), (H) Gsw (g_{sw}, stomatal conductance, mmol $m^{-2}s^{-1}$) and (I) Fm' (maximum fluorescence yield). Significance calculated with Wilcoxen signed-rank test for non-parametric paired data (Rosner et al. 2006) (NS; not significant, * p <0.05, ** p<0.01, *** p<0.001). 141

Figure 28; Manhattan plots highlighting markers significantly associated with the traits; (A) Gsw (g_{sw}, stomatal conductance, porometer), (B) E (transpiration, porometer), (C) development (BBCH and harvesting), (D) ETR (electron transport rate, fluorometer), VPDleaf (leaf vapour pressure deficit) and Tleaf (leaf temperature), (E) 'responses' to drought (weeks 1 and 2) and (F) 'recovery' (weeks 4 and 5). The difference (diff) was calculated to see how the traits changed over time and to see differences in the magnitude of change between weeks (Table 3). Week 1 (W1), week 2 (W2), week 3 (W3), week 4 (W4) and week 5 (W5). The analyses were completed with GAPIT and the models FarmCPU, BLINK, MLMM or MLM on 142 accessions (Huang et al. 2019; Liu et al. 2016; Segura et al. 2012; Wang and Zhang

2021; Zhang et al. 2010). The X-axis represents the genomic position of markers, and the Y-axis is the -log 10 of the P-values for association with the phenotype. The vertical lines correspond to QTLs associated with at least two phenotypes. Point size correlates to -log10(P-value). Quantile-quantile (QQ) plots are provided in Figure 29. 143

Figure 29; Quantile-quantile (QQ) plots are provided for Manhattan plots in Figure 28. (A) Gsw (g_{sw}, stomatal conductance, porometer), (B) *E* (transpiration, porometer), (C) development (BBCH and harvesting), (D) ETR (electron transport rate, fluorometer), VPDleaf (leaf vapour pressure deficit) and Tleaf (leaf temperature), (E) 'responses' to drought (weeks 1 and 2) and (F) 'recovery' (weeks 4 and 5). The difference (diff) was calculated to see how the traits changed over time and to see differences in the magnitude of change between weeks (Table 3). Week 1 (W1), week 2 (W2), week 3 (W3), week 4 (W4) and week 5 (W5). The analyses were completed with GAPIT and the models FarmCPU, BLINK, MLMM or MLM on 142 accessions (Huang et al. 2019; Liu et al. 2016; Segura et al. 2012; Wang and Zhang 2021; Zhang et al. 2010). The X-axis represents the expected -log 10 of the P-values and the Y-axis is the observed -log 10 of the P-values for association with the phenotype. The Y-axis in plots A-C and E-F have breaks to improve visualisation of the distribution and extreme values. 146

Tables

Table 1; Phenotypic characteristics associated with each subpopulation for the diversity panel (144 accessions). Table from Denning-James et al. (2025). 62

Table 2; 23 QTLs associated with determinacy and photoperiod sensitivity. Table from Denning-James et al. (2025). 83

Table 3; Key dates within the water deficit experiment (June 2023 to September 2023), the stage of the trial and the number of days after sowing (DAS). 124

Table 4; Distribution of soil sensors among the growth habits and treatments during the trial. 126

Table 5; The drought stress strategies within the diversity panel were identified based on phenotypic characteristics from comparative analysis on accessions under water-deficit conditions (142 accessions). Including photoperiod sensitivity and growth habit data (photoperiod insensitive and photoperiod sensitive). Data from week 4 (four weeks since irrigation was removed) was selected as it reflected the maximum length of water deficit before the 'recovery' event. 138

Table 6; The 36 QTLs for g_{sw} (porometer), *E* apparent (porometer), development (BBCH and yield), ETR (electron transport rate, fluorometer), VPDLeaf (leaf vapour pressure deficit), Tleaf (leaf temperature), water deficit 'response' (weeks 1 and 2) and 'recovery' from water deficit (weeks 4 and 5). 147

Acknowledgements

This PhD was supported by the Biotechnology and Biological Sciences Research Council (UKRI-BBSRC) through the Norwich Research Park Doctoral Training programme. I am grateful to my primary supervisor, Dr José De Vega, for giving me the opportunity to pursue this PhD and for being consistently calm, positive, and encouraging throughout the process. I also appreciate the advice and perspectives of my secondary supervisors, Dr Caspar Chater and Dr Anthony Hall.

My work has benefited from the support of colleagues across EI, NBI, CIAT, Agrosavia, NIAB and Kew, as well as from other collaborators and co-authors. In particular, I am thankful to Dr Andres Cortes, Dr Matthew Blair, Dr Tom Wood and Dr Claudia Lowe for generously sharing their insights. I am also grateful to the researchers and technicians at CIAT and Agrosavia, who welcomed me warmly, showed me Colombia, and shared their expertise.

Past and present members of the JDV lab group have been a constant source of support, and I've appreciated all the coffee breaks and lab outings. I am especially grateful to Cam, James, and Juliana for always listening and checking in. I also want to thank Laura from IBM, who instilled confidence and belief in me, her support really meant a lot. Many people also helped during sowing and harvesting, including Catherine from JIC Horticultural Services, as well as Juliana, Lu, Sofia, Jess, and Johnny.

My PhD would not have been the same without the laughs at lunches, pub quizzes and other weekend adventures with Mia, Jess, Sofia, Insect George, Lu, James and Becky. I am deeply grateful to them for their support and always cheering me up. The ESB also organised many activities, which provided welcome breaks along the way. Norwich itself has been made special by the community I found here, including Carole, Lizzie, and Gillian, and their wonderful pets Mimi, Nemo, and Goofy, who all helped me maintain a healthy work-life balance.

Finally, I owe the greatest thanks to my family and friends beyond Norwich. For always being there for me, cheering me on, sending care packages, and encouraging me to keep going. I am especially grateful to my parents and my partner Jack, who supported me through the stressful times and reminded me to always be kind to myself.

Chapter 1

Literature Review

Sections 1.4.2, 1.4.3 and 1.8 published at:

Kate E Denning-James, Caspar Chater, Andrés J Cortés, Matthew W Blair, Diana Peláez, Anthony Hall, José J De Vega, Genome-wide association mapping dissects the selective breeding of determinacy and photoperiod sensitivity in common bean (Phaseolus vulgaris L.), G3 Genes|Genomes|Genetics, Volume 15, Issue 6, June 2025, jkaf090, https://doi.org/10.1093/g3journal/jkaf090

1.1 Agrobiodiversity

Agrobiodiversity is a subset of biodiversity that comprises all diversity related to agriculture (Figure 1) such as landraces and crop wild relatives (CWR) (Brookfield and Stocking 1999). Agricultural ecosystems provide us with goods (food, pharmaceuticals, forage, and bioenergy) and services, such as soil and water quality, and carbon sequestering (Power 2010).

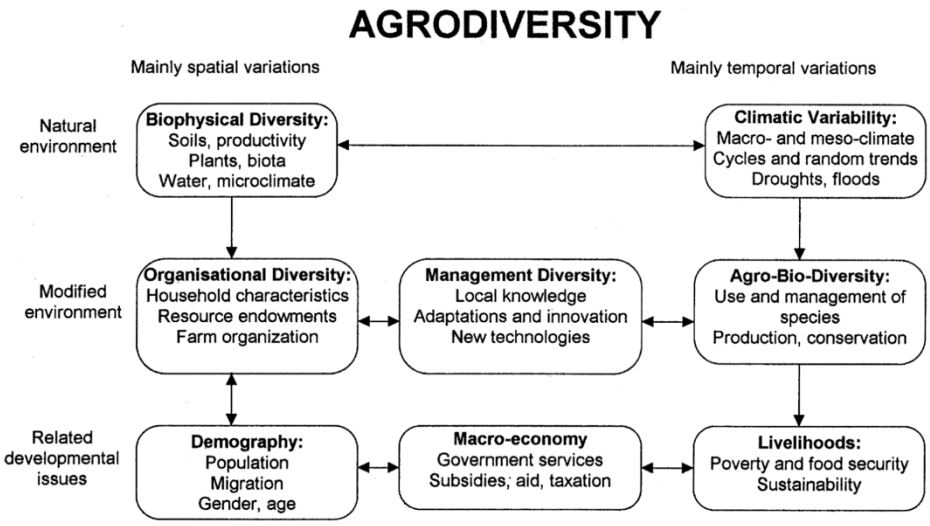


Figure 1; Diagram showing the elements and interactions within agrobiodiversity from Brookfield and Stocking (1999).

CWR are a wild taxon related to crop varieties. The degree of relatedness can vary, from the same species to the same genus or family (Maxted et al. 2006). They usually have higher levels of genetic diversity, having not experienced the genetic bottleneck of domestication (Warschefsky et al. 2014). CWR adaptive genes can be hybridised into elite varieties with breeding programs. However, this can be challenging due to linkage drag - whereby the genetic background of CWR leads to a fitness reduction due to deleterious variants in domesticated traits that were not filtered out during domestication.

Landraces are cultivated crops which have undergone artificial selection, known as domestication. The landraces become genetically and morphologically different to their CWR, and are associated with traditional farming systems and ecogeographic environments (Di Vittori et al. 2017). However, they are more genetically diverse than the modern, homogenous, elite,

commercial varieties, as the latter result from a discrete number of founders (founder effect) (Savic et al. 2021). Landraces can usually tolerate local biotic and abiotic stress better than elite varieties; this results in more stable yields under low input agricultural systems (Zeven 1998). Heirlooms or vintage species are old varieties with defined phenotypic traits that have been consciously maintained by farmers through generations. A large proportion of genetic diversity is stored in landraces or heirloom varieties.

Crop domestication is dynamic and continuous, leading to domestication syndrome traits (traits commonly found after domestication, e.g. shattering and determinacy) and often pleiotropic effects and polyploidisation. However, domestication causes genetic bottlenecks, a reduction in genetic diversity (genetic erosion) and can increase deleterious recessive mutations in the population, which are maintained because of the prevalence of inbreeding (Mackay et al. 2021; Meyer et al. 2012; Singh et al. 2019). Autogamous species, such as common beans tend to be more strongly affected by genetic erosion due to lower levels of recombination (De Ron et al. 2015). Genetic erosion can also continue after domestication during dispersion, migration, and adaptive radiation at secondary centres. Domesticated crops also have higher levels of linkage disequilibrium, a function of recombination, population size and the breeding system and increased by genetic drift, hitchhiking and epistatic selection (Rossi et al. 2009). Consequently, it is important to understand the domestication process to understand the structure and complexity of the genomes.

The genetic diversity from landraces and CWR can be mobilised into the background of elite varieties to reduce the effects of biotic and abiotic stress on yields, protecting food security. Breeding or genetic engineering are some of the solutions to creating more climate resilient, higher yielding and pest or disease resistant varieties (FAO 2009). Plant genetic resources provide the reservoir of adaptive and productive genes (usually free of deleterious mutations) needed to develop improved cultivars with desirable characteristics and are crucial to sustaining genetic gain. However, mobilising genetic diversity is complex (McCouch et al. 2020). In pre-breeding, exotic material (landraces, CWRs, heirlooms) is crossed into the genetic background of elite productive varieties. However, the trait introgression can be masked by regulatory interactions (dominance, epistasis), linkage drag, pleiotropy or quantitative inheritance.

The genetic diversity of crops originates from ancestral populations. However, landraces and their genetic diversity are under threat of genetic extinction from replacement with homogenous elite varieties and our increasing dependence on a smaller number of crops for food security (Villa et al. 2005). Elite, commercial varieties are genetically homogenous due to economic and consumer pressure. They rely on humans and advances in agricultural technology for high yields and survival such as machinery, fertilisers, pesticides, and herbicides (Casanas et al. 2017; Kwak et al. 2012).

1.2 Common Beans, *Phaseolus vulgaris*

Common beans, *Phaseolus vulgaris*, are one of the domesticated species in the *Leguminosae* or *Fabaceae* family, the third largest flowering plant family (Azani et al. 2017). The genus *Phaseolus* has ~70-80 species, five of which are domesticated (Chacon-Sanchez et al. 2021; Freytag and Debouck 2002).

P. vulgaris are diploids (2n = 2x = 22), with a relatively short life cycle (60-120 days) and small genome of ~600Mb (Kwak et al. 2009; Schmutz et al. 2014). Due to their small genome and synteny, they can be used as a model for the complex genomes of soybeans, pigeon pea or cowpea (Nadeem et al. 2021; Schmutz et al. 2014). Common beans are autogamous and annual; however, wild species can be perennial and allogamous (Chacon-Sanchez et al. 2021; Debouck et al. 1993; Schier et al. 2019). The gene pools of common beans grow in a large variety of environments in the neotropics. There are large differences in their life history traits, morphology and genetics (Beebe et al. 2012; Bitocchi et al. 2017; Broughton et al. 2003; Gepts and Debouck 1991). Common beans can be grown under a variety of cultivation practices, such as in, monoculture, associations, interplanting or rotations (Kwak et al. 2009).

Common bean genetic resources are found in gene banks worldwide (e.g. CIAT and IPK) (Nadeem et al. 2021). Breeding programmes remain important to protect common bean yields and global productivity. Past breeding programmes have developed elite cultivars across different regions and market classes, targeting yields, seed characteristics, and abiotic and biotic stress tolerance. The conservation of the accessions is important for the genetic improvement of common bean yields and to overcome abiotic and biotic stress, currently done through classical breeding. Molecular tools and genomic approaches can be used alongside conventional breeding to accelerate the process. However, common bean yields are still

heavily impacted by different stresses such as climate change, consequently phenotyping common bean landraces and CWR remains important for future yield stability.

Common beans show a high level of genomic synteny with other legume species such as *Glycine max* (soybean), due to their shared evolutionary history within the *Fabaceae* family (Schmutz et al. 2014). This evolutionary relationship allows comparative genetics within the family so that model legumes, such as soybean, can be used to infer gene function and help with candidate gene identification in common beans. However, there is still a lack of data available in legumes therefore, gene annotation and candidate gene identification often requires more distantly related model species such as *Arabidopsis*. This can reduce the confidence in candidate genes when relying on distant homology (Zhou et al. 2025).

1.3 Importance of Common Beans

Common beans help achieve developmental goals by reducing poverty (1.3.1), improving health and nutrition (1.3.2), and improving ecosystem resilience (1.3.3).

1.3.1 Consumption and economics

A second green revolution is needed to produce 70% more food before 2050 for an estimated population size of 9.6 billion, to guarantee food and nutritional security. Insecurity will disproportionately effect developing countries (Alexandratos and Bruinsma 2012). In this context, common beans have been labelled as one of the important crops to mediate the effects of climate change and protect food and nutritional security. *Phaseolus vulgaris* is the most economically important species in the *Phaseolus* genus (Arriagada et al. 2021; Bitocchi et al. 2017). However, the worldwide production of common beans has only increased by 0-2% per year, relying on more land, not increases in yield from research developments (Foyer et al. 2016).

Common beans are cultivated mainly for their mature grain, but the immature seeds, pods and leaves can also be eaten (Blair et al. 2010a; Ganesan and Xu 2017). Common beans include, Carioca, navy, black, pinto, pink and yellow beans, to name a few (Rawal and Navarro 2019). Different countries produce different common beans depending on demand.

1.3.2 Nutritional Qualities

Common Beans are one of the most important grain legumes for human consumption worldwide (Broughton et al. 2003). They have a high nutritional content of proteins and minerals, which is especially important in developing countries, but deficiencies in human diets are a major concern worldwide (Lisciani et al. 2024; WHO 2023). Globally, the highest producers of common bean include India, Brazil, Myanmar and Tanzania (FAOSTAT 2023). In South and Central America, the highest producers are Brazil, Argentina and Mexico. Colombia produced ~124,000 tonnes of dried beans in 2023, most of which was for domestic consumption, and provides an important source of nutrition within the country.

Common beans have been labelled as one of the essential crops to mediate climate change due to their lower environmental impact, and protection of food and nutritional security (Foyer et al. 2016). They are considered the second most important source of protein in human diets and third most important source of calories in regions of Africa and Latin America. The protein in ~100g of common beans provides ~20-25% of the recommended daily intake (mainly phaseolin) for humans, and the composition meets the requirements of the WHO and FAO. Common beans provide an affordable source of proteins (comparative to animal protein) in developing countries with a long storage life (Castro-Guerrero et al. 2016; Patto et al. 2015). In countries such as Mexico and Brazil they are the primary source of protein. Their amino acids can be complemented with sulphur containing amino acids in cereal (Broughton et al. 2003). Also, the seeds and seed coats contain high levels of carbohydrates (unsaturated fatty acids), protein, minerals, fibres, micronutrients and vitamins (Blair et al. 2013b; Ganesan and Xu 2017).

They have bioactive components important in human metabolism, as well as key nutraceutical properties. They can have positive effects on cardiovascular diseases, obesity and diabetes (Suarez-Martinez et al. 2016). These are chronic and degenerative diseases that cause high levels of mortality worldwide. Common beans also have anti-carcinogenic properties, as the phenolics, lectins and protease inhibitors are anti-mutagenic, anti-inflammatory and anti-proliferative (Bernardi et al. 2023; Ganesan and Xu 2017; Jha et al. 2015; Suarez-Martinez et al. 2016). Subsequently, common beans are classified as a functional food, as they provide nutrition as well as other physiological benefits to consumers. However, the protective properties and nutritional composition of common beans can depend on the variety, cultivar and seed size (Blair et al. 2013b; Caproni et al. 2020; Ganesan and Xu 2017; Jha et al. 2015).

Biofortification could improve the nutritional value of common beans, to aid the reduction of nutritional deficiencies in poorer communities and worldwide, especially where common beans

are commonly consumed (Caproni et al. 2020). Biofortification is improving the nutritional value of crops via biotechnology or plant breeding. For example, in common beans, the level of bioavailable iron, zinc, folate, and protein could be improved through breeding programs. Also, the levels and bioavailability of nutrients could be improved by understanding the distribution of micronutrients within the seed (cotyledon, seed coat and embryo), and the accumulation of anti-nutrients (e.g. tannins, phytic acid, chelating compounds) (Blair et al. 2013b; Caproni et al. 2020; Jha et al. 2015; Kachinski et al. 2022).

1.3.3 Ecosystem services

Common beans have economic and environmental importance due to their symbiotic relationship with nitrogen-fixing *Mycorrhizal* bacteria, called *Rhizobia*. The bacteria fix atmospheric nitrogen, release ammonia into the soil, enhance nitrogen levels, and reduce the need for expensive chemical fertilisers, whilst improving yields (Cusworth et al. 2021; Mupangwa et al. 2021; Mylona et al. 1995; Phiri and Njira 2022). This supports sustainable agriculture due to a reduced dependency on fertilisers (Castro-Guerrero et al. 2016). This is particularly important for developing countries that may not have access to fertilisers (Kamfwa et al. 2019; Phiri and Njira 2022).

Furthermore, they are important for other ecosystem services, for example supporting pollinators (UK bumblebee populations). However, common beans seem to encourage fewer types of pollinators than other crops such as oilseed rape (Garratt et al. 2014). They also encourage soil fertility by increasing populations of earthworms during arable rotations (Scullion et al. 2002; Stopnisek and Shade 2021).

Finally, common beans are high in Biotin. This is important for cellular functions in humans, but also in livestock. Common beans are used as a forage crop (straw of the plant) for livestock. Biotin is usually made industrially, which is expensive and time-consuming, therefore, beans reduce the production energy and costs (Broughton et al. 2003).

1.4 Population Structure of Common Beans

Common beans have a complex population structure. There are different gene pools, races, intermediate species and admixed accessions due to genetic isolation, fragmentation and artificial selection for different morphological traits. Common beans are mainly autogamous, but outcrossing rates can be ~60-70%, permitting gene flow and introgressions (Santalla et al. 2004). Also, ecogeographic conditions have disrupted the gene flow between wild and domesticated common beans and, additionally between the different gene pools (Mesoamerican and Andean). Understanding the genetic diversity, population structure, phylogeny, origin and evolution of common beans is important for informing future breeding programmes and to elucidate patterns of adaptations (Beebe et al. 2012).

1.4.1 Wild common beans

There are two ecogeographic gene pools of wild common beans, the Mesoamerican and Andean. The wild Mesoamerican gene pool is distributed from Mexico and Central America, while the wild Andean gene pool is dispersed from southern Peru to northern Argentina, including Bolivia (Blair et al. 2013a; Kami et al. 1995; Mamidi et al. 2013; Rossi et al. 2009; Santalla et al. 2004). Research suggests a Mesoamerican origin of wild common beans. The Mesoamerican origin is supported by loci (Mamidi et al. 2013), the linkage disequilibrium (Rossi et al. 2009) and nucleotide diversity (Bitocchi et al. 2012). This theory suggests that the wild ancestor of common beans evolved in Mesoamerica, likely Mexico (Bitocchi et al. 2012). Then, the wild ancestral population split asymmetrically causing two bottlenecks (Figure 2, (Schmutz et al. 2014). The Mesoamerican gene pool had a smaller pre-domestication bottleneck, compared to the wild Andean, explaining how the wild Mesoamerican gene pool is more diverse. Similarities between the gene pools may be due to ancestral introgressions or homoplasy (Rendon-Anaya et al. 2017b).

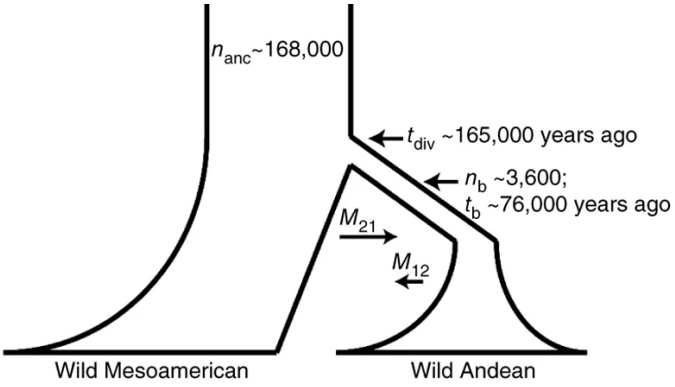


Figure 2; The uneven split of the wild Mesoamerican and Andean common bean gene pools. Highlighting the asymmetric split and bottlenecks of the wild populations from Schmutz et al. (2014). n_{anc}; the estimated size of the ancestral population, n_b; the estimated size of the bottleneck population, t_{div}; the estimated timing of the start of the bottleneck, t_b; the estimated length of the bottleneck, M_{21} and M_{12}; represent the migration rates between the two wild populations.

Other wild common bean gene pools or subpopulations have been suggested including a Guatemalan, Colombian and Ecuadorian-northern Peruvian (Blair et al. 2012). The Ecuadorian-northern Peruvian population has no domesticated species and is thought to be a relic ancestral population that diverged early from the Mesoamerican, carrying only a portion of the ancestral genetic diversity (Frascarelli et al. 2025). Then due to geographic and reproductive isolation, there was no gene flow between the gene pools (Bitocchi et al. 2012; Kami et al. 1995). The subpopulation in some analyses is considered a sister species to *Phaseolus vulgaris* named *Phaseolus debouckii*, located in Ecuador and Peru (Aguilar et al. 2022; Rendon-Anaya et al. 2017a). However, further wild accessions need to be analysed from this group to investigate this relationship conclusively (Cortinovis et al. 2020; Lynch and Conery 2000).

1.4.2 Domestication

Crop domestication causes a genetic bottleneck because of artificial selection (selective breeding) (Papa and Gepts 2003). Resulting in landraces having lower levels of genetic diversity compared to their wild counterparts (Cortes et al. 2011; Kwak and Gepts 2009). In common beans there were two separate, parallel domestications, one for each ecogeographic gene pool, ~8,000 years ago (Castro-Guerrero et al. 2016). Consequently, the domestications took place after the separation of the Andean and Mesoamerican wild gene pools. Research estimates that the origins of common bean domestication are the Oaxaca valley for the Mesoamerican gene pool and southern Bolivia-northern Argentina for the Andean one (Bitocchi et al. 2013; Kwak et al. 2009; Rossi et al. 2009).

The Mesoamerican domesticated population had a reduction in diversity of 72%, compared to the Mesoamerican wild gene pool, suggesting a large genetic bottleneck during Mesoamerican domestication. However, overall research suggests that the wild Mesoamerican gene pool has more genetic diversity than the Andean (Pinon et al. 2021). Multiple domestications were suggested for the Mesoamerican gene pool due to the large amount of diversity observed. However, this has been resolved concluding that there was one domestication event in the Mesoamerican gene pool, as the populations were not sufficiently isolated to undergo genetic isolation and independent evolution (Chacon et al. 2005; Papa and Gepts 2003; Rossi et al. 2009).

The Andean domesticated gene pool has a lower loss in genetic diversity compared to the corresponding wild population, suggesting there was already reduced genetic variability in the Andean wild population (Bitocchi et al. 2013; Mamidi et al. 2013; Schmutz et al. 2014). Some studies have found that Andean landraces have a higher genetic diversity than their wild relatives, potentially due to admixture with Mesoamerican populations and an accumulation of mutations (Schmutz et al. 2014). Consequently, the Andean gene pool has maintained genetic diversity during domestication (Trucchi et al. 2021). Some studies have found that the Andean gene pool has a complex structure and many polymorphisms (Cortes et al. 2011). However, since these studies focus on Andean diversity alone, the Andean gene pool may be over-represented (Becerra et al. 2010; Blair et al. 2007).

The morphological, biochemical, agronomic and genetic differences between the two gene pools are well documented (Bitocchi et al. 2012; Gepts and Bliss 1985; Gepts and Debouck 1991; Sajgalik et al. 2019; Singh et al. 1991b). The gene pools can be distinguished based on their morphological differences from artificial selection and ecogeographic conditions. For

example, the Mesoamerican gene pool is small seeded while the Andean is more commonly large seeded (Blair et al. 2009; Zhang et al. 2008). There are differences between seed colour, seed pattern, genome size (Castro-Guerrero et al. 2016), cooking time, flowering time (Sadohara et al. 2022), nutritional content (Mesoamerican landraces have higher levels of iron seed content (Caproni et al. 2020)), linkage disequilibrium (Ambachew et al. 2024) and phaseolin proteins. Also, within both gene pools there are high levels of genetic and morphological diversity, such as in seed colour, seed size, growth habit and agro-ecological adaptation (Beebe et al. 2001; Blair et al. 2007; Debouck et al. 1993; Plestenjak et al. 2024).

1.4.3 Domestication syndrome traits

Domestication has caused morphological differences to their wild counterparts due to artificial selection by humans, outcrossing and divergence. Some traits are commonly selected for in crops; these are called domestication syndrome traits. Examples of these traits include photoperiod insensitivity (PI), determinacy (D), gigantism (seeds and pods), growth habit (GH) (bush types), seed dormancy (lost in domesticated), shape and plant colour, dissemination mechanisms (shattering), seed colour, pod colour (human selection for variation) (Di Vittori et al. 2017). While domestication also affected genes in pathways related to flowering time, reproduction, nitrogen metabolism, and hormone production (Santalla et al. 2004; Schmutz et al. 2014).

Photoperiod insensitivity and determinacy arose separately in both gene pools during the domestication of common beans, likely co-selected by growers (Bhakta et al. 2017; Kwak et al. 2012; Repinski et al. 2012; Weller et al. 2019). Wild common beans tend to be indeterminate and photoperiod sensitive (PS), requiring a particular day length to flower. Indeterminate growth is advantageous in the wild due to competition with surrounding vegetation, while photoperiod sensitivity was likely reinforced by divergent natural selection and local adaptation. On the other hand, photoperiod insensitivity was selected (likely unconsciously) as cultivated common beans were spread along a greater range of latitudes and environments.

Determinacy, a developmental feature that causes common beans to have a terminal inflorescence when switching to a reproductive state (Cavalcante et al. 2020), optimised agricultural management and harvesting efficiency. Determinate common beans tend to have a bush growth habit with reduced branching and vining abilities compared to the indeterminate varieties (Kwak et al. 2012), therefore translocating biomass resources into an increased fitness

output (Figure 5). While indeterminate and photoperiod sensitive landraces are common, the combined selection for photoperiod insensitivity and determinacy resulted in common bean varieties with shorter flowering periods, earlier maturation and easier management during harvesting (Daba et al. 2016; González et al. 2016). Photoperiod insensitivity and determinacy are advantageous traits from an agronomical point of view due to earlier harvesting and shorter exposure to unfavourable weather patterns under climate change, consequently providing better food security for communities (Botero and Barnes 2022; Perez et al. 2020). Many studies have identified that common bean flowering is closely associated with the environment (altitude), gene pool, growth habit, location, or seed size (Sadohara et al. 2022; White and Laing 1989).

1.4.4 Races within gene pools

Classifying the races and subgroups of common beans is complex due to admixture populations and introgressions, but also due to overlapping morphological traits. At first, the identification of races was based on agro-morphology. This identified three races within the Mesoamerican gene pool (Durango, Jalisco, Mesoamerica) and three races within the Andean gene pool (Chile, Nueva Grande and Peru) (Singh et al. 1991a). However, more recent studies associate morphological data with genetic diversity to understand the race structure and subgroups of common beans. An early genetic analysis based on allozyme analysis suggested five groups in the Mesoamerican gene pool and four within the Andean (Singh et al. 1991b). However, later studies identified three Andean races (Chile, Nueva Grande and Peru) and four Mesoamerican races (Jalisco, Durango, Guatemala and Mesoamerica) (Becerra et al. 2010; Blair et al. 2013a; Chacon et al. 2005; Cortes et al. 2011).

In certain studies, there has not been enough morphological and genetic markers to distinguish sub-groups; for example, the Mesoamerican races Durango and Jalisco can be grouped together, and the Andean, Chilean group cannot always be distinguished (Blair et al. 2009; Cortes et al. 2011; Pinon et al. 2021). Also, there can be overlaps between the Nueva Grande and Peruvian races, likely because of a large overlap in morphology, or because the limited number of markers could not separate admixture between the races (Blair et al. 2007). Also, subgroups within the races have been found. For example, two genetic subgroups for the Andean races Nueva-Grande and Peru and for the Mesoamerica race (Beebe et al. 2000; Blair et al. 2009). These limitations emphasise the issues linked to morphological-based classification into races.

1.4.5 Dissemination of Common Beans

Common beans have secondary centres of domestication resulting from random genetic drift, human selection, and plant material movement by humans. This affects the genetic diversity due to founder effects, genetic bottlenecks, artificial selection and ecogeographic conditions. Common beans have now been introduced and are cultivated worldwide.

Both the Andean and Mesoamerican common bean gene pools were disseminated first to the Caribbean before European colonisation of the Americas (Duran et al. 2005). Then, during the Post-Columbian exchange of genetic resources, the gene pools reached Europe and Africa (Myers et al. 2022). The dissemination of common beans in Europe was through the Iberian Peninsula, causing a genetic bottleneck (Santalla et al. 2002). This bottleneck was mitigated by later gene flow from the Americas and hybridisation within Europe (Angioi et al. 2010; Catarcione et al. 2023; Sajgalik et al. 2019; Savic et al. 2021). Africa, Brazil and China are other secondary centres of domestication with both Andean and Mesoamerican diversity and phenotypic differences (Blair et al. 2010b; Burle et al. 2010; Maciel et al. 2003; Wu et al. 2020; Zhang et al. 2008).

1.4.6 Gene flow

The population structure of common beans is made more complex due to the presence of introgressive hybridisation and admixture driven by their dynamic nature and breeding mechanisms. Within the gene pools there are different races, intermediate species and admixed accessions due to genetic isolation, fragmentation and artificial selection for various morphological traits. The ecogeographic conditions, together with isolation by distance, have disrupted the gene flow between wild and domesticated common beans, as well as between the different gene pools (Beebe et al. 2012; Santalla et al. 2004). Consequently, there are large differences in their life history traits, morphology and genetics (Beebe et al. 2012; Bitocchi et al. 2017; Broughton et al. 2003; Gepts and Debouck 1991). However, this earlier research was based on morphological traits, phaseolin seed proteins or low-density molecular markers, that provide a lower resolution than whole genome approaches.

In common beans, there is gene flow and hybridisations between the wild and domesticated pool, forming competitive weedy plants with cultivated traits (Chacon-Sanchez et al. 2021; Papa and Gepts 2003; Santalla et al. 2004; Singh et al. 1991b). However, these hybridisations between the domesticated and wild gene pools can affect the primary gene pool, either

increasing or decreasing the diversity of the wild relative (Harlan and Wet 1971). The wild gene pools may have partial reproductive isolation, an important consideration for bean breeding (Koinange and Gepts 1992).

Introgressive hybridisation and putative sources of novel diversity can be found in domesticated common beans due to frequent but uneven gene flow between the Andean and Mesoamerican gene pools (Beebe et al. 2001; Bitocchi et al. 2017; Blair et al. 2013a; Gepts and Bliss 1985). The Mesoamerican and Andean gene pools meet in the north-west of South America. The region covering Colombia, therefore, Colombia may contain a higher proportion of admixture diversity and introgressions and potentially a novel or useful source of diversity (Blair et al. 2013a; Blair et al. 2007; Debouck et al. 1993; Leitao et al. 2021a; Tohme et al. 1996). Although these studies have a relatively small number of molecular markers, therefore impacting the accuracy when estimating introgression patterns. Introgressions between the gene pools can also be found in secondary centres of domestication, such as in Europe (Angioi et al. 2010; Catarcione et al. 2023).

Finally, hybridisations are common among races within the domesticated gene pools, adding to the difficulty of disseminating gene flow owing to a lack of genetic isolation (Blair et al. 2007; Blair et al. 2009; Debouck et al. 1993). Due to the large amount of gene flow within common beans accessions that do not fit into a certain race or gene pool are common (Blair et al. 2007).

Natural intraspecific variation provides important resources for discovering the genetic and molecular basis of phenotypes. For example, life history, plant development, adaptation and productivity traits (Raman et al. 2019). By understanding the gene flow between and within Common Bean gene pools and races, the ancestry and relatedness of the accessions can be hypothesised. Also, interspecific and intraspecific breeding in *Phaseolus species* could provide novel genetic diversity to improve the fitness of cultivars under biotic and abiotic stress. However, further work using a higher number of molecular markers or whole genome sequencing will help to resolve the relationships between gene pools within common beans.

1.5 Exploring genetic diversity to respond to climate change

Climate change related impacts, such as increasing temperatures and precipitation irregularity, are the main threats to common bean production and future sustainable yields (IPCC 2021). Areas including South America are predicted to get the highest increase in ecological and agricultural drought. Consequently, food security will be increasingly under risk, with predictions suggesting there could be a 10% decrease in crop yields by 2050. Prioritising heat and drought tolerant cultivars is one way to reduce the effects of climate change on food security (Tai et al. 2014). For this, we need to understand the impacts climate change has on plant responses and environmental interactions and to investigate the genetic diversity available for bean breeding programmes.

1.6 Global water scarcity

Water stress has been identified as one of the major global environmental global risks (WEF 2024). The effects of water shortages are recorded worldwide across continents (Bista et al. 2024; IPCC 2023; Majumder 2015; World Resource Institute 2023). Estimates suggest at least 50% of the world experiences water-deficit stress yearly (World Resource Institute 2023). Regions in Chile, Venezuela and Brazil are predicted to have the biggest increase in drought length (Naumann et al. 2018). This will disproportionately affect the poorest regions, without irrigation infrastructure, and that rely on rain-fed systems (Magrin et al. 2014; Rawal and Navarro 2019).

Globally, water resources are overexploited and poorly managed; other challenges include water pollution and unsuitable infrastructure (Majumder 2015). Climate change is linked to recurring droughts and erratic global rainfall, exacerbating water stress impacts (FAO 2024). Alongside this, the El Niño-Southern Oscillation (ENSO) climatic pattern is linked to higher temperatures, leading to a greater number of heatwaves, droughts and wildfires (WEF 2024). Water scarcity reduces water and food security through increasing food prices, malnutrition and poverty (IPCC 2023).

1.6.1 Drought definitions

Drought is usually characterised as an unusually long period of sustained dry weather (Fioravanti et al. 2025). However, there are multiple ways to define drought stress, dependent on the region, impacts, duration and frequency, making drought stress complex to investigate. The definitions based on impacts include; meteorological drought (insufficient precipitation), agricultural drought (when the precipitation does not meet the needs of crops, affecting yields) and hydrological drought (when the lack of precipitation impacts water reserves such as lakes) (Tuberosa 2012).

Drought stress can also be classified by the plant developmental stage. Terminal drought occurs during flowering and seed-filling but before maturity while vegetative drought occurs during vegetative growth and before flowering (Farooq et al. 2017; Habus Jercic et al. 2018; Shavrukov et al. 2017; Tabassum et al. 2018).

1.6.2 Drought stress in agricultural crops

Agriculture is responsible for ~70% of annual global freshwater use (Barezzi et al. 2024). Research has shown the severity of impacts of drought stress throughout the plant kingdom (Alza et al. 2024; Bohra et al. 2024; El Bey et al. 2024; Polania et al. 2020). Consequently, prioritising drought tolerant cultivars which maintain yields under drought stress, is a priority.

Drought is a complex stress and can activate different mechanisms in plants depending on duration, intensity, frequency, plant genotype, agricultural practises and plant development stage (Munoz-Perea et al. 2006). Plant adaptations to abiotic stresses, such as drought, involve a complex interaction among genes and pathways (Araujo et al. 2015). The main two drought tolerance strategies in common beans are drought avoidance and escaping (Polania et al. 2022). Drought escaping accessions rapidly complete the lifecycle before the onset of drought, therefore are adapted to seasonal, predictable long-term droughts where precipitation may not return (Shavrukov et al. 2017). Drought avoidance strategies include ‘water spenders’ (anisohydric) that accelerate development and 'water savers' (isohydric) that delay growth to conserve resources.

Water spenders maintain photosynthesis during drought stress by keeping their stomata open at the expense of water loss. If this strategy enables reproduction and seed set, it can be a useful crop trait that prioritises yield stability under long-term droughts. However, if spenders

cannot complete their lifecycle before the return of rains, the strategy is more suited to milder and shorter droughts after which development can resume (Nesporová et al. 2024). Water savers are plants which reduce photosynthesis by closing stomata rapidly upon the detection of water deficit, and are most effective during milder, shorter droughts when precipitation returns, stomata reopen, and growth and development continue (Nesporová et al. 2024). Another strategy, the functional 'stay-green' (SG) response, delays leaf senescence, thereby maintaining chlorophyll levels, and prolonging photosynthetic assimilation. In this way, prolonged carbon fixation can reduce drought-associated yield losses (Labastida et al. 2023), a strategy that has been identified in crop species including sorghum (Borrell et al. 2022), wheat (Kumar et al. 2022), cowpea (Nunes et al. 2022), and common beans (Sofi et al. 2021).

The mechanisms that allow plants to survive water-deficit can affect carbon fixation, transpiration and photosynthate mobilisation to seeds and pods in the plants (including seed abortion) (Polania et al. 2020; Tardieu 2012). Also, flowering time and leaf, shoot and root growth (Beebe et al. 2013; Neumann 2008), as well as differences in hormone pathways (e.g. ABA concentration), stomatal conductance and chlorophyll concentration (Hageman et al. 2020; Teran and Singh 2002b; Tuberosa 2012; Wang et al. 2024). While stomatal conductance and leaf temperature can act as proxies for drought stress in plants (Marchin et al. 2020; Smith et al. 2019)

There are many approaches to combating the yield impacts of drought in commercial varieties. Consequently prioritising a drought resilience strategy in breeding programmes, is an important consideration for farmers to maintain yields under different drought stresses (Polania et al. 2016). The process of domestication, and modern breeding for high yielding varieties, have both limited the genetic diversity of our crops and inadvertently produced varieties with high stomatal conductance and low water use efficiency (Huang and Zeng 2024; Lei et al. 2023). Utilising and mobilising natural diversity during breeding programmes could address these imbalances to improve commercial cultivars. Landraces and crop wild relatives are a rich and underexplored source of novel diversity which could help to combat abiotic and biotic stresses (Renard et al. 2023).

1.6.3 Common beans and drought stress

Common beans are particularly susceptible to drought stress, with predictions suggesting ~60% of common bean production is impacted by drought stress (Beebe et al. 2008; Hageman et al. 2020; Villordo-Pineda et al. 2015). In the field, yields can be reduced by 53% along with seed weight, harvest index and days to maturity (Smith et al. 2019; Teran and Singh 2002a). Drought stress is likely to become more common in many regions where common bean is produced due to climate change and ENSO impacts. Common beans are also subjected to seasonal drought in certain regions of Central and South America, including Colombia (Hernández-López et al. 2024; Sánchez-Reinoso et al. 2020). However, many farmers in these regions rely exclusively on rain-fed agriculture, and without access to water for irrigation, these droughts strongly impact yields and food security.

Prior research into drought tolerance in common beans has investigated the Mesoamerican (Polania et al. 2017a; Villordo-Pineda et al. 2015) and Andean gene pools separately (Dramadri et al. 2019). As well as investigating and comparing the drought tolerance strategies of different races within the gene pools (Beebe et al. 2008; Munoz-Perea et al. 2006; Polania et al. 2020; Teran and Singh 2002a). Other research has investigated both gene pools together (Labastida et al. 2023), or included secondary centres for diversification such as European common beans (Papathanasiou et al. 2022), African varieties (Darkwa et al. 2016) or admixed accessions (Leitao et al. 2021a). Wild Colombian accessions have been investigated under drought stress (Cortés et al. 2013) and Colombian landraces have been included in other panels to investigate responses under drought stress. Also, Colombian accessions susceptible to drought stress have been crossed with more distant relatives (López-Hernández et al. 2023). However, few panels have focused on drought tolerance in a panel focusing on accessions from Colombian and neighbouring countries utilising a GWAS approach at whole-genome resolution.

This research suggests that *Phaseolus* species contain novel genetic diversity which could be used in breeding programs to improve drought tolerance (Buitrago-Bitar et al. 2021; Hoyos-Villegas et al. 2017). Breeding drought tolerant cultivars could reduce the need for irrigation systems, consequently reducing production costs whilst improving yields under drought environments (Teran and Singh 2002b)

1.7 Breeding programmes

Modern breeding programmes are moving beyond a yield-centred paradigm to target resistance to biotic and abiotic stress, such as drought stress, and also nutritional quality (Assefa et al. 2019; Caproni et al. 2020; Kachinski et al. 2022; Singh and Schwartz 2010). Landraces and crop wild relatives offer a promising reservoir of genetic diversity for these traits by trait introgression from landraces into the elite genetic background (Hu et al. 2021a; Suarez et al. 2021a; Suarez et al. 2021b; Tai et al. 2014).

Understanding the genetic diversity, population structure, patterns of adaptations, and how these correlate with determinacy and photoperiod insensitivity is required to guarantee the retention of these key domesticated traits within future breeding cycles, given their association with crop management and production (Beebe et al. 2012). Also, mobilising diversity from drought tolerant cultivars into modern commercial varieties is one way to maintain yields under climate change. However, a multi-faceted approach would be preferable, incorporating government policies (Li et al. 2020), new irrigation methods (Barezzi et al. 2024), soil supplements (Lui and Mihara 2024) and breeding programmes.

1.8 Colombian and Admixed Common Beans

Colombia is the northernmost part of the Andean gene pool and south of the Mesoamerican and may act as a region of confluence between them. Consequently, it has been proposed that the region has a large amount of admixture and introgressive hybridisation (Blair et al. 2013a; Blair et al. 2007; Leitao et al. 2021a; Tohme et al. 1996). Admixture and hybridisation lead to introgressions from differential parental origins, introducing new alleles and novel epistatic interaction into a population. This allows for new trait combinations that could merge exotic variation from diverse germplasm with more agronomic-desirable traits such as determinacy, photoperiod insensitivity and resistance to water deficit. Also, a large proportion of common bean cultivation in Colombia takes place on low-input, small-scale farms in mountainous areas, producing yields without access to fertilisers or irrigation. Consequently, the cultivars are adapted to the biotic and abiotic stresses from the ecogeographic region and may provide novel diversity (Beebe et al. 2012). Evaluating the genetic diversity in Colombian common beans is important for the conservation, management and utilisation of the genetic resources.

1.9 Aims and Objectives of the Thesis

Common beans are important for food security, nutritional security and ecological services. They are genetically diverse, which is suggested by the two domestications and gene pools, but also the various species, races, and genetic groups. In my PhD I plan on investigating the hypothesis that Colombian Common Beans are genetically diverse and contain admixed diversity from the two gene pools. I characterise the gene pools and subpopulations and their relationships. Linking the population structure to domestication traits such as determinacy and photoperiod sensitivity. This understanding is important for the conservation, management and utilisation of the genetic resources through breeding programs.

Further to this, research suggests that common beans show natural tolerance to biotic and abiotic stresses (such as drought) in both gene pools and admixed populations. In my PhD, I plan on investigating the hypothesis that Colombian Common Beans could provide novel diversity for drought responsive traits. This is because Colombia contains diversity from both gene-pools, therefore, there is opportunity for introgressions and admixture. The number of drought events are expected to increase worldwide it is important to identify key candidate genes to safeguard common beans yields for large farms as well as for small-holder, subsistence farmers through future breeding programs. Furthermore, the research will expand knowledge into how common beans, and the different races respond to different types of drought stress. Drought is a complex trait; therefore, exploring and mobilising the adaptations (genetic diversity) in landraces could improve drought tolerance in elite cultivars.

Objectives:

- To investigate the hypothesis that Colombia has a large amount of genetic admixture diversity and introgressed lines from the Common bean Mesoamerican and Andean gene pools and, therefore, holds a large amount of genetic diversity for the study with fewer geographic effects.

- To understand the population structure and phylogeny of the diversity panel of Colombian common beans in the context of the Mesoamerican and Andean gene pools, and the gene flow and admixed structural variation across them.

- To understand the range of natural phenotypic diversity in the panel for domestication traits (photoperiod insensitivity and determinacy) and tolerance to water-deficit.

- To identify QTLs in the diversity panel for the domestication traits and tolerance to water-deficit, undertaking GWAS.

- To identify novel genetic diversity in the common bean diversity panel and identify putative candidate genes for future breeding programs and to safeguard yields against different climate change scenarios.

Chapter 2

Selective breeding for determinacy and photoperiod sensitivity in common bean (*Phaseolus vulgaris* L.)

All sections except 2.3.5, 2.3.8, 2.4.3, 2.4.5, 2.4.8, 2.4.9, 2.5.4 and 2.5.8 published at:

Kate E Denning-James, Caspar Chater, Andrés J Cortés, Matthew W Blair, Diana Peláez, Anthony Hall, José J De Vega, Genome-wide association mapping dissects the selective breeding of determinacy and photoperiod sensitivity in common bean (Phaseolus vulgaris L.), G3 Genes|Genomes|Genetics, Volume 15, Issue 6, June 2025, jkaf090, https://doi.org/10.1093/g3journal/jkaf090

Contributions: Whole Genome Sequencing completed by Genomic services at Earlham Institute (Norwich, UK). The selection of the diversity panel was completed by supervisory team. CIAT's Genebank and IPK's Genebank provided the germplasm. All other work completed during the PhD by Kate E Denning-James and other authors supervised, reviewed and edited.

2.1 Abstract

Common bean (*Phaseolus vulgaris* L.) is a legume pulse crop that provides significant dietary, and ecosystem benefits globally. This study investigated two key traits, determinacy and photoperiod sensitivity, that are integral to its management and crop production, and that were early selected during the domestication of both Mesoamerican and Andean gene pools. Still, significant variation exists among common bean landraces for these traits. Since landraces form the basis for trait introgression in pre-breeding, understanding these traits' genetic underpinnings and relation with population structure is vital for guiding breeding and genetic studies.

This study explored genetic admixture, principal component, and phylogenetic analyses, using whole-genome sequencing to define subpopulations and gene pools. Genome-wide association mapping (GWAS) identified marker-trait associations, in a diversity panel of common bean landraces. A clear correlation was observed among these traits, gene pools and subpopulation structure. Extensive admixture was found between the Andean and Mesoamerican gene pools in some regions. Thirteen QTLs were identified for determinacy, and ten QTLs for photoperiod sensitivity, along with underlying putative causative genes. This study identified known and novel causative genes, and a high proportion of pleiotropic effects for these traits in common bean, and likely translatable to other legume species. The mean allele dosage of the QTLs was examined, to further explore the population structure and phenotypes at a SNP level.

2.2 Introduction

2.2.1 Importance of Common Beans

The common bean is a global staple that provides significant dietary and economic services by improving health and nutrition, while helping to reduce poverty, specifically in developing countries (1.3.2 and 1.3.3). Common beans have also been labelled as one of the essential crops to mediate climate change due to their lower environmental impact and protection of food and nutritional security (Foyer et al. 2016). There are hundreds of varieties, and the prevailing type grown in a country depends on market preferences (Rawal and Navarro 2019).

Common beans are rich in essential dietary components, such as protein, minerals, fibre, and micronutrients (Blair et al. 2013b; Castro-Guerrero et al. 2016; Patto et al. 2015) and protect against some forms of malnutrition (Bernardi et al. 2023; Ganesan and Xu 2017; Jha et al. 2015; Suarez-Martinez et al. 2016). Also, common beans have a symbiotic relationship with nitrogen-fixing bacteria, allowing them to fix atmospheric nitrogen and enhance nitrogen levels in the soil, thereby reducing the need for expensive chemical fertilisers while improving yields (Cusworth et al. 2021; Mupangwa et al. 2021; Mylona et al. 1995; Phiri and Njira 2022).

2.2.2 Common Bean population structure

The common bean underwent two separate domestications resulting in two gene pools, the Andean and Mesoamerican. In addition, there are different races, intermediate species, and admixed accessions. The gene pools of common beans grow in a large variety of environments in the neotropics. These ecogeographic conditions, together with isolation by distance, have disrupted the gene flow between wild and domesticated common beans, and between the different gene pools (Beebe et al. 2012; Santalla et al. 2004). Consequently, there are large differences in their life history traits, morphology, and genetics (Bitocchi et al. 2017; Broughton et al. 2003; Gepts and Debouck 1991).

2.2.3 The Diversity Panel

The diversity panel was comprised of 144 genotypes, mainly from Colombia and surrounding countries in Central and South America (Figure 3, Appendix 1). The panel contained accessions from elite backgrounds, landraces, heirlooms, weedy and wild materials. The material was sourced from the International Centre for Tropical Agriculture (CIAT)'s genebank (90 accessions), the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)'s genebank (37 accessions), and heirlooms bought from the catalogues from 'Jungle Seeds' (JungleSeeds 2020) and 'Beans and Herbs' (Herbs 2020) in 2020 (17 accessions). The panel was chosen to maximise genetic diversity across Central and South America. Accessions were prioritised to represent the common bean gene pools and their races (Mesoamerican: Durango, Jalisco, Mesoamerica; Andean: Chile, Nueva Grande and Peru). Within each gene pool, the accessions focused on Colombia and neighbouring countries, to capture regional genetic diversity and potential gene flow between gene pools. Including those putatively considered admixed varieties, for further investigation. Accessions from Brazil were included as representatives of the secondary centre of diversity for common beans (Burle et al. 2010) and from Chile to represent the full geographic diversity of the Andean common bean gene pool (Becerra et al. 2010). Finally, to reduce possible redundancy in the panel, accessions were selected to represent a diverse range of seed coat colours (Figure 4).

Figure 3; Distribution of the 127 common beans with location data from the diversity panel. The coordinates of the capital city were used for those without coordinate data. Produced with QGIS. Figure from Denning-James et al. (2025).

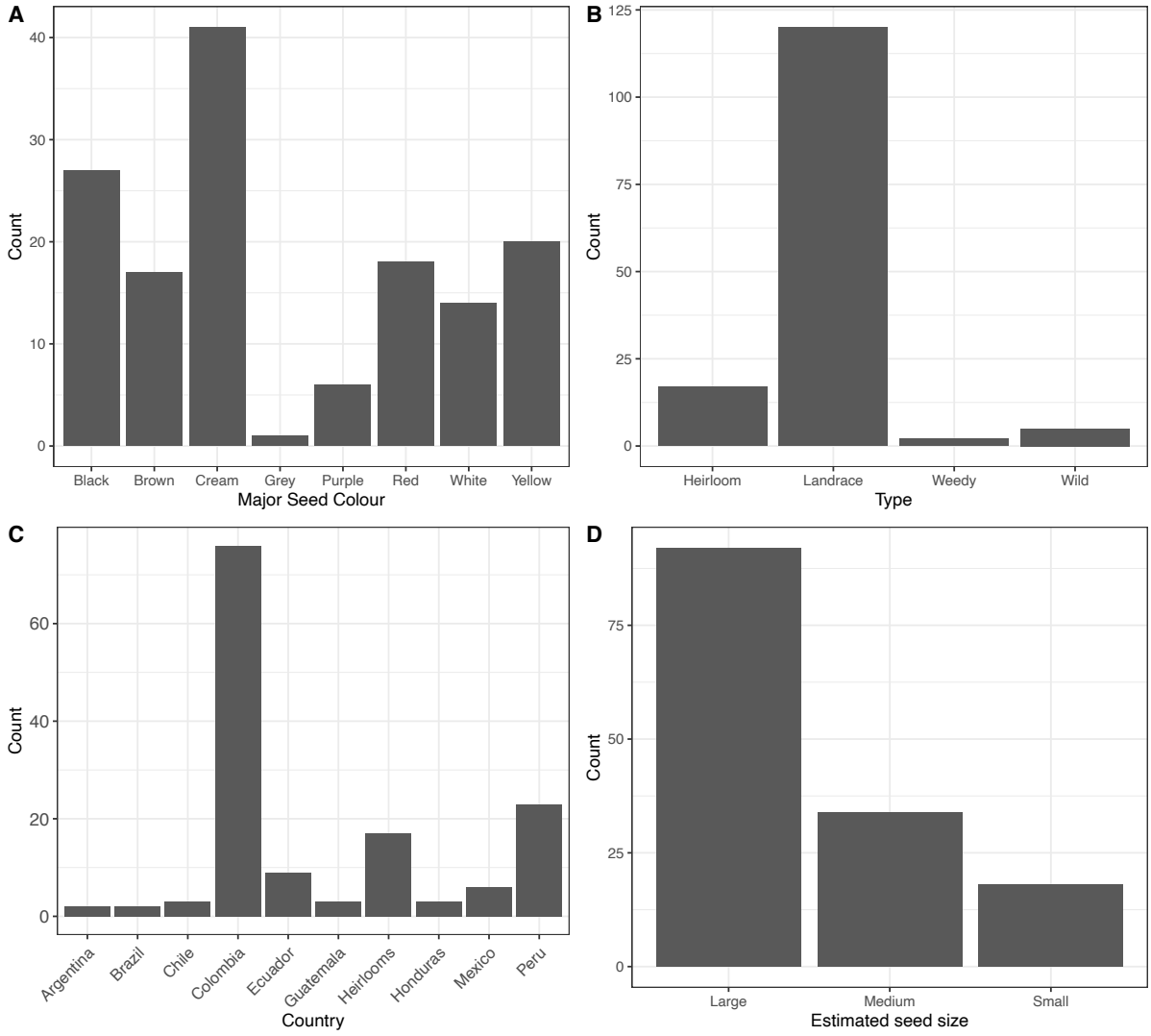


Figure 4; Frequency distribution of phenotypic traits for the diversity panel (144 accessions). (A) Major seed colour, (B) type of accession, (C) country of origin or collection and (D) estimated seed size. Data collated from corresponding genebank websites and phenotypic studies during 2022 (Appendix 1).

2.2.4 Determinacy and Photoperiod Insensitivity

Photoperiod insensitivity and determinacy arose separately in both gene pools during the domestication of common beans, likely co-selected by growers as both traits are advantageous (Repinski et al. 2012; Weller et al. 2019). Photoperiod insensitivity was selected (likely unconsciously) as cultivated common beans were spread along a greater range of latitudes and environments, allowing the crop to maintain yields across diverse conditions. Determinacy, a developmental feature that causes common beans to have a terminal inflorescence when

switching to a reproductive state (Cavalcante et al. 2020)(Figure 5), optimised agricultural management and harvesting efficiency by promoting more uniform flowering and maturation. Photoperiod insensitivity and determinacy are advantageous traits from an agronomical point of view due to earlier harvesting and shorter exposure to unfavourable weather patterns under climate change, consequently providing better food security for communities (Botero and Barnes 2022; Perez et al. 2020).

Determinacy and photoperiod insensitivity are distinct traits, but have been indirectly linked during domestication, as they have allowed the dissemination of common beans into new environments while improving traits associated with harvesting. Their co-selection is likely due to agricultural practices and environmental conditions.

Figure 5; Examples of the three different growth habits for common beans. (A) A determinate bush (PHA13862), (B) an indeterminate bush (G5910) and (C) an indeterminate climbing growth habit (G50516K). Photos taken at harvesting during September 2023. Black arrows are equivalent to 20cm.

2.2.5 Admixed diversity in Colombia

Colombian common beans include a range of growth habits and photoperiod insensitivity (PI). Colombia may act as a region of confluence between the two common bean gene pools consequently there may be a large amount of admixture and introgressive hybridisation (Blair et al. 2013a; Blair et al. 2007; Leitao et al. 2021a; Tohme et al. 1996). Several studies have

identified admixture and introgressions between the Andean and Mesoamerican gene pools of common beans, introducing new alleles and novel epistatic interaction into a population.

Tohme et al. (1996) and Blair et al. (2007) identified genetic differences between the two gene pools, but also reported shared variation and admixed genotypes, particularly in regions such as Colombia. Blair et al. (2013a) identified genetic diversity with admixed ancestry, contributing to the increased genetic diversity in common bean. Leitao et al. (2021a) also reported introgression and gene flow between gene pools.

These studies suggest that admixture and introgressions have contributed to the dissemination of common beans and the possible incorporation of beneficial variation. However, these studies mainly focused on population structure and genetic diversity, rather than linking introgression to specific traits. This chapter adds to the prior work by exploring how admixture impacts traits associated with photoperiod insensitivity and determinacy. Determining how introgressions may contribute to trait variation in common beans.

2.2.6 Project aims

Landraces and crop wild relatives offer a promising reservoir of genetic diversity for domestication traits by introgression from the landraces into the elite genetic background (Hu et al. 2021a; Suarez et al. 2021a; Suarez et al. 2021b; Tai et al. 2014). Understanding the genetic diversity, population structure, patterns of adaptations, and how these correlate with determinacy and photoperiod insensitivity is required to guarantee the retention of these key domesticated traits within future breeding cycles, given their association with crop management and production (Beebe et al. 2012).

A genome-wide association mapping (GWAS) was utilised to identify significant SNPs for photoperiod insensitivity and determinacy in this diversity panel. The novelty of this work lies in that prior research commonly focused on the Mesoamerican diversity rather than the Andean, due to the greater genetic diversity in the former, and had ignored admixed materials as an essential source of variation. Furthermore, research had rarely utilised whole genome sequencing of common bean accessions to undertake a GWAS on determinacy and photoperiod insensitivity phenotypes. Instead, previous work has mostly used QTL mapping and low-density marker panels, resulting in poor resolution (García-Fernández et al. 2021; González et al. 2016; Kwak et al. 2008).

2.3 Methods

2.3.1 Genotyping

163 *Phaseolus vulgaris* accessions were selected and requested in 2020. However, due to poor DNA concentrations (<10ng/uL), 147 were sent for sequencing. Then after sequencing one sample was removed as further research suggested this was a cowpea (*Vigna*), while two samples were removed after sequencing but before further processing due the low raw DNA yield. The resulting diversity panel is as described in section 2.2.3.

The genotypes were whole genome re-sequenced using Illumina short reads. The accessions were grown at the Norwich Research Park (Norwich, UK) in 2021 until the expansion of the first true leaf, after which they were snap-frozen (~50-100 mg). The genomic DNA extraction for short read sequencing from each accession was completed with a Qiagen DNAeasy kit (Qiagen, Germany). The DNA concentration of the samples was quantified for quality control using the Tecan Plate Read Infinite F200 Pro for a fluorometry based assay.

The sequencing of the samples was completed by Genomic services at Earlham Institute (Norwich, UK). Low Input Transposase Enabled (LITE) libraries, a cost-effective low volume variant of the standard Illumina TruSeq DNA protocol, were constructed for the 144 accessions using a protocol based on the Illumina Nextera kit (Illumina, California USA). They were sequenced with two NovaSeq 6000 S4 v 1.5 flow cells with 150bp paired-end reads, following the protocol in (Kirkwood et al. 2021). All sequence data is available in the SRA database under the BioProject number PRJEB81566.

2.3.2 Phenotyping

All 144 common bean accessions were evaluated at the Norwich Research Park (Norwich, UK) in temperature-controlled glasshouses (set from 17-21°C). The experiments were conducted in two seasons; summer 2022 (March: July) with long daylength (16:8) and winter 2022 (November: March) with short daylength (12:12). The accessions were organised in a randomised block design with 3 or 2 replications, respectively (Figure 6). Management was conducted according to recommendations for common bean cultivation. Whereby they were irrigated, grown in 5 litre pots with JIC cereal mix soil (65% peat, 25% loam, 10% grit, 3kg/M3 dolomitic limestone, 1.3kg/M3 PG mix and 3kg/M3 osmocote exacte) and stakes were used for those with indeterminate growth habits.

Within the glasshouse, there were sensors recording temperature and humidity, that were averaged over 2 hours for plotting in R with ggplot (Gemini Tinytag, Ultra 2) (Wickham 2016). Daily data on time of sunrise and time of sunset which calculated daylength was collected from the CW4326, Norwich, UK station (52.643°, 1.237°) (Visual Crossing Corporation 2025).

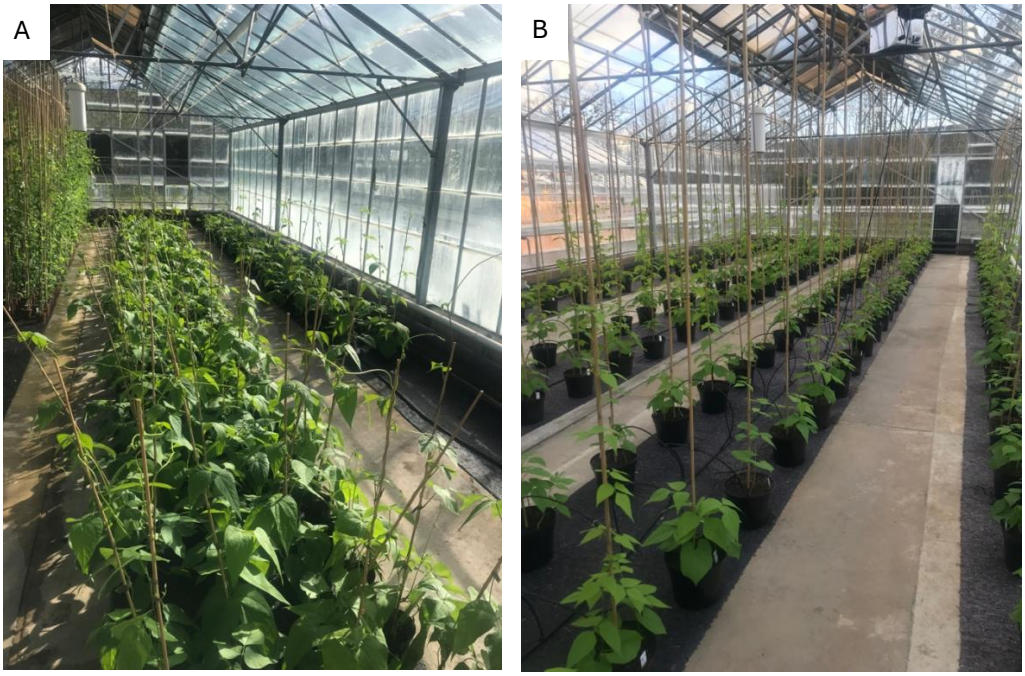


Figure 6; Layout of the glasshouse (144 accessions) from phenotyping trials in summer 2022 (A) and winter 2022 (B).

The diversity panel was characterised for days to flowering (DTF), seed size (SS), weight of 100 seeds (E100_SW; estimated based on the weights of seeds harvested and projected to 100 seeds), determinacy (D; terminal flower bud presence) (Cavalcante et al. 2020) and photoperiod sensitivity (PS; flowering in none, one or both seasons). DTF was split into the two seasons due to photoperiod sensitivity in certain accessions and PS was characterised in three ways for the GWAS (Appendix 1).

The phenotypic data were tested for normal distribution using the Shapiro-Wilk test from the rstatix package (p-value > 0.05) (Kassambara 2023). The statistical analysis of variance (one-way ANOVA) of the phenotypic data was done in R and plotted with ggplot2 v 3.5.1 (Wickham 2016). Correlations and associations were calculated between discrete and continuous phenotypes with a one-way ANOVA, and the correlation ratio was calculated as eta-squared.

Cramer Vs calculated correlation between discrete datasets and significance was calculated with chi-squared, then Pearson correlation coefficients calculated correlations between continuous datasets with normal distribution. This was visualised using the R package 'corrplot' v 0.95 (Wei and Simko 2021).

2.3.3 Pre-processing genotype data

The raw sequences were analysed with Kraken 2 (v 2.0.7) (Wood et al. 2019). The parameter paired was used to classify the sequence data by taxa, and to confirm if there was a large amount of contamination in the sequence data as the kraken database classifies species with 16s RNA. This was used as a proxy for contamination. The percentage unclassified, as common bean was not present in the database, ranged from 92.3% to 96.8%. The raw sequence reads were processed with TrimGalore (v. 0.5.0) (Krueger et al. 2023) to remove adapters and poor-quality reads, and then quality checked using fastqc (v 0.11.5) (Andrews 2010) and multiqc (Ewels et al. 2016).

The trimmed reads were aligned to the Andean reference genome, *Phaseolus vulgaris* G19833, v2.1 (Schmutz et al. 2014) and to a Mesoamerican reference genome, *Phaseolus vulgaris,* Labor Ovalle v1.1, downloaded from Phytozome (Goodstein et al. 2012). These references were initially selected to represent genetic diversity from both common bean gene pools and to minimise reference bias during alignment. Both assemblies are reference grade, with chromosome level scaffolding and high contiguity, making them suitable for downstream analyses.

Alignment took place to both reference genomes with BWA-MEM (v 0.7.13) (Li and Durbin 2009) and '-M -R' to add read group information and allow compatibility with GATK. SAMtools (v 1.7) combined, compressed and sorted the aligned files (Danecek et al. 2021). Picardtools (https://broadinstitute.github.io/picard/) (v 2.1.1) marked duplicates and BamTools (v 2.5.1) indexed the alignments (Barnett et al. 2011). The percentage of alignments and proportion of heterozygous sites were calculated at this stage and analysis continued with the Andean reference genome. The genotype data was divided in 10 Mbp regions with FreeBayes v 1.0.2 (Garrison and Marth 2012) to run Genome Analysis ToolKit (GATK v 4.2) haplotype caller with default parameters (Van der Auwera and O'Connor 2020). This identified 20.2 million variant loci (~17.1M SNPs and ~3.4M indels). At this stage, the Andean reference genome (G19833 v2.1) was selected for further analysis because it produced higher alignment percentages across the

diversity panel. Using the Andean reference genome therefore reduced the risk of missingness or uneven coverage compared to using a more divergent reference genome (Section 2.4.1; Figure 7).

2.3.4 Population structure analysis

The resulting VCF file, from running GATK with the Andean reference ('Andean VCF'), was filtered further with BCFtools (v. 1.12) to retain calls with a minimum depth of 5 reads per variant call (FMT/DP >= 5), a locus call quality over 30, maximum missing calls per locus of 5%, to keep only biallelic SNP locus, and for a minor allele frequency over 2%. The resulting VCF had ~9 million SNP loci. Then, the VCF was filtered for a maximum heterozygosity of 20% per locus using TASSEL 5 (v. 20230314) (Bradbury et al. 2007). This was then filtered for linkage disequilibrium (LD) (based on LD decay) and thinned with a window size of 10 bps with BCFtools prune (v. 1.12).

The population structure of the panel was analysed using ADMIXTURE (v 1.3.0) (Alexander and Lange 2011) on a subset of 88,786 SNP loci. ADMIXTURE was run for K=2 to K=10 and the ideal number of K was determined from the cross-validation error. Accessions were allocated a group when their membership coefficient (q) was greater than 0.7. Plotting was completed in R with the package 'ggplot2' (Ginestet 2011).

The population structure was also analysed with STRUCTURE v 2.3.5 (Pritchard et al. 2000) for K=2-10 with 10 iterations per K. However, for this analysis, the VCF needed to be further thinned using BCFtools with a window size of 30 to reduce the number of SNPs to 29,599. This analysis was not taken further because the fewer SNPs is a less reliable overview of the population structure of the diversity panel.

2.3.5 Introgression analysis

The 'Andean VCF' from GATK was filtered with BCFtools (v. 1.12) (Danecek et al. 2021) for biallelic loci, a minor allele frequency of 1% and thinned with a window size of 5 bp, forming a VCF with a subset of 2,572,124 loci.

Accessions were identified with a q greater than 0.999 at K=2 from ADMIXTURE analysis. This identified 56 Andean, 22 Mesoamerican and 56 admixed. Using these subpopulations and the Andean VCF with 2,572,124 loci, private alleles were identified to represent the two groups (Martin et al. 2023) with VcfHunter v.3. Then, the private alleles identified in the accessions of each subpopulation were used to calculate the average expected allele ratio for the subpopulation the allele belongs to. After, for each allele from a subpopulation, the observed allele ratio is calculated for each accession. In the next step, the parameters were set to a window size of 12 with no overlap and a ploidy of 2. The observed average ratio of alleles is calculated for each subpopulation and normalised by the expected allele ratio. This infers the haplotypes from the subpopulations in all the accessions. The file is then reformatted for plotting. Plotting took place in R using the package 'ggplot2' (Ginestet 2011).

Private alleles were also identified for K=6 and K=4 from ADMIXTURE. However, this was unsuccessful due to the low proportion of alleles which were private to each subpopulation. Possibly because the subpopulations did not have enough different private alleles; therefore, were too closely related. Normality of distribution was tested with the Kruskal Wallis test (p-value > 0.05) (Kruskal and Wallis 1987).Then, significant associations between genomic proportions from the Mesoamerican or Andean backgrounds were completed with the Dunn test and Bonferroni correction for multiple groups with non-normal distributions from the package FSA v 0.10.0 (Derek H et al. 2025; Dunn 1964). Dunn test selected instead of ANOVA as the data is non-normally distributed, therefore, violates ANOVA assumptions (using the mean).

2.3.6 Genome wide association study

As in section 2.3.5 the 'Andean VCF' from GATK was filtered with BCFtools (v. 1.12) (Danecek et al. 2021) for biallelic loci, a minor allele frequency of 1% and thinned with a window size of 5 bp. A principal component analysis (PCA) generated with GAPIT v.3 (Wang and Zhang 2021) was used to understand the genetic relationship between accessions, on a subset of 2,572,124 loci.

A genome-wide association study investigated marker-trait association for determinacy and photoperiod insensitivity phenotypes using GAPIT v.3 (Wang and Zhang 2021) with three principal components. Three models were used Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) (Huang et al. 2019), Fixed and random model Circulating Probability Unification (FarmCPU) (Liu et al. 2016) and Mixed Linear Model (MLM) (Zhang et al. 2010). BLINK and FarmCPU were identified as the best multi-locus models for different heritability levels, improving statistical power (Cebeci et al. 2023; Huang et al. 2019; Merrick et al. 2022). MLM was chosen for single-locus analysis as a baseline for comparison to BLINK and FarmCPU.

GAPIT was run on the whole panel (144 accessions) and on the Andean subpanel (as defined at K2 ADMIXTURE; 108 accessions). To run BLINK, GAPIT completed the analysis with the option 'Random.model=TRUE' as to not calculate R^2 for phenotypic variance explained (PVE) values after GWAS. The quantile-quantile (QQ) plots were used to understand the suitability of the models to the data. Plotting was completed in R using the packages 'ggplot2' (Ginestet 2011), and 'qqman' (Turner 2018).

2.3.7 Selecting significant loci, candidate gene mining and functional annotation

Significant marker-trait associations (MTAs) were investigated further when they had a $-\log_{10}$(p-value) over 7 and were confirmed by two models from GAPIT. QTLs were defined as ±100 kbp from the MTA based on the estimated LD (linkage disequilibrium) decay distances in published common bean diversity panels, using a $r^2 = 0.25$ cutoff (estimated decay as 114Kb) (Campa et al. 2018; Moghaddam et al. 2016; Raggi et al. 2019; Reinprecht et al. 2024; Ugwuanyi et al. 2022; Valdisser et al. 2017; Wu et al. 2020; Wu et al. 2024). This is shorter than an earlier calculated recombination rate in common bean of 3.72 cM/Mb (Bhakta et al. 2015). LD decay was estimated for the diversity panel (mean $R^2 = 0.27$) and each subpopulation at K=2 (Andean mean $R^2 = 0.21$, Mesoamerican mean $R^2 = 0.2$) using PopLDdecay software following Wu et al. (2020) (Zhang et al. 2019).

Identified loci were compared to the Andean reference genome, *Phaseolus vulgaris* G19833 v2.1 in JBrowse (Diesh et al. 2023; Schmutz et al. 2014) whilst considering 'high impact' mutations identified by SnpEff (Cingolani et al. 2012). Once genes were identified, their putative function was explored using PhytoMine (Goodstein et al. 2012) (*Phaseolus vulgaris v.2*), BLAST (Camacho et al. 2009) against the non-redundant (nr) protein database at NCBI, and finally against the TAIR database, if no gene function could be identified in close relatives (Huala et al. 2001). The loci were compared to previous studies and literature. PulseDB was used for comparison, particularly for QTLs and markers related to developmental and flowering phenotypes (Humann et al. 2019). QTLs and markers were mapped to the reference genome to estimate the conversion from cM to Mb in JBrowse.

2.3.8 Mean allele dosage

To visualise SNPs within the ~200kb QTLs, the VCF from the GWAS analysis was loaded using VCFR in R (Knaus and Grünwald 2017) along with the Andean reference genome and annotation (Schmutz et al. 2014). Using the annotation, only SNPs within the coding region of genes were selected and the SNP calls were transformed into dosages, so that homozygous reference were 0, heterozygous were 1, homozygous alternative were 2 and NAs were 3. NAs were included as the VCF included insertions and deletions.

The SNPs were plotted with pheatmap v 1.0.13 in R (Kolde 2025) with the columns ordered by SNP location and the rows clustered using Euclidean distance. Multiple methods of plotting were tested. The SNPs were plotted by individual accessions, however, due to the large variation among individuals it was hard to identify trends. Then the mean for each position was calculated based on determinacy. This often resulted in losing information on the differences in SNP calls between the Mesoamerican and Andean gene pools. Consequently, the SNP mean was calculated based on determinacy and the population structure at K2, to identify trends with a higher resolution. To identify SNPs that were different between the groups, only the SNPs that had a difference greater than or equal to 0.5 were included for plotting, while those below were removed. This helped remove SNPs which did not have differences and to improve clarity. Finally, selected candidate genes with 'high' SNPeff mutations were investigated further (Cingolani et al. 2012).

2.4 Results

2.4.1 Alignment to reference genomes

The diversity panel was aligned to the Mesoamerican and Andean reference genomes (Figure 7). All accessions aligned with a percentage greater than 89% or 91% against the Mesoamerican and Andean genomes, respectively. Except for the wild accession from Ecuador, G23723, which aligned to both genomes with ~87%. Figure 7 suggests the population structure of the diversity panel, with two groups appearing, those that align better to either the Andean or Mesoamerican reference genome. This population structure incorporates Andean and Mesoamerican, heirlooms, landraces, wild and weedy accessions. The analysis continued with the Andean reference genome, as the alignment was higher across the panel. If there had been a large difference in alignments, I would have continued with both.

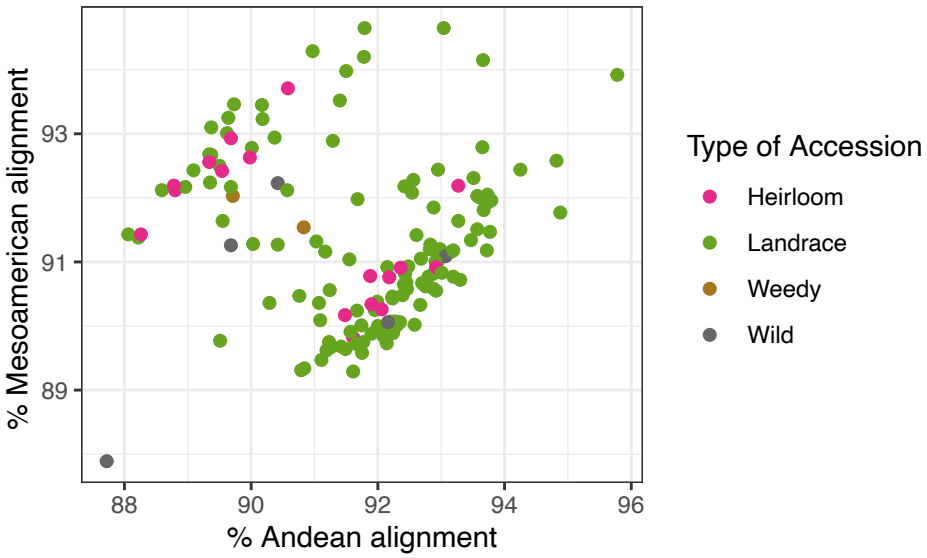


Figure 7; A scatter plot representing the alignment of the diversity panel (144 accessions) to the Mesoamerican and Andean reference genomes. The colours represent the type of accession from gene bank passport data (Appendix 1).

2.4.2 Population structure

The diversity panel split into the two gene pools, the Andean and Mesoamerican at K=2 (Figure 8A, Figure 9A). Analysis of the cross-validation (CV) error (Figure 8D) suggested that K=4 and K=6 were the better representations of the panel with K values from 2 to 10 (Alexander and Lange 2011). At ADMIXTURE K4, the Andean gene pool from K2 is divided into subpopulations, the Mesoamerican gene pool maintains the same structure as K2 (Figure 8B). While the Andean gene pool is split into three subpopulations, the Colombian (C1 and C2, Figure 8), the Andean (A1) and the C-EP. At K6 (Figure 8C), the Mesoamerican group split into two subpopulations (M1 and M2), while the Andean subgroup split into 4 subpopulations. Two of these subpopulations included only accessions from Colombia and were named C1 and C2. A subpopulation containing accessions from Colombia, Ecuador and Peru was named C-EP. The remaining subpopulation was named A1.

At K=2 STRUCTURE, the population structure of the diversity panel matches the results from ADMIXTURE (Pritchard et al. 2000) (Figure 8A). It splits the panel into the Andean and Mesoamerican gene pools. Analysis of the delta K suggests that K=6 was the better representation of the panel using STRUCTURE (Pritchard et al. 2000) (Figure 9C). At K=6 (Figure 9B) the Andean gene pool was split into C-EP, C1, C2 and A1, however, the Mesoamerican gene pool remained the same.

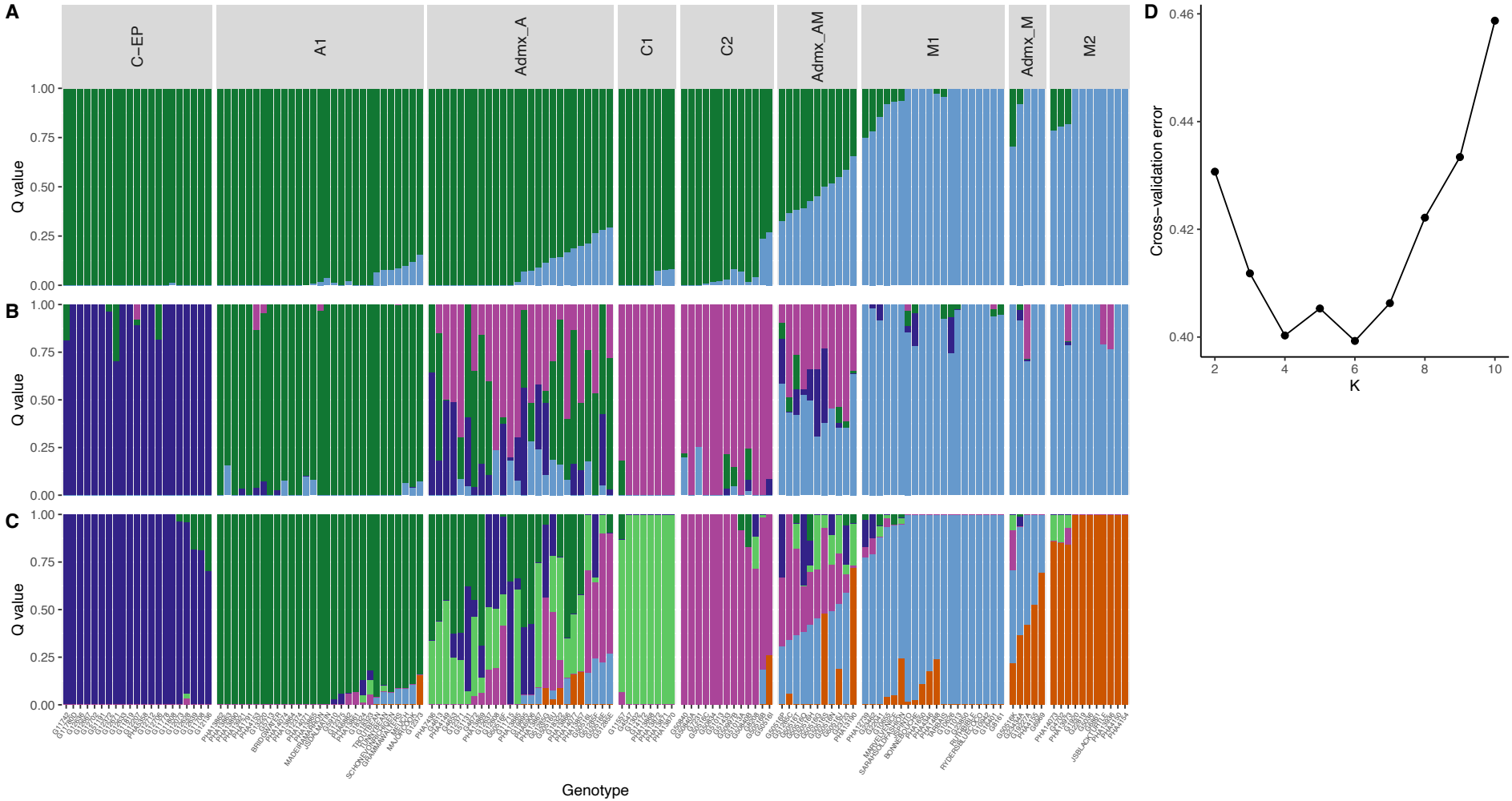


Figure 8; Analysis of the population structure of 144 accessions belonging to our diversity panel focusing on Colombia at K=2, Andean or Mesoamerican groups (A), K=4 (B) and K=6 (C) and the cross-validation error for different K values (D) ran using ADMIXTURE (v 1.3.0)(Alexander and Lange 2011). (C-EP) accessions mainly from Peru, then Ecuador and Colombia; (A1) Andean accessions from a variety of South American countries;

(C1) mostly determinate Colombian landraces; (C2) indeterminate Colombian landraces; (M1) mainly medium seeded** from Central America and Colombia; (M2) mainly small seeded** from Central America and Colombia. (Admx_AM) Andean X Mesoamerican hybrids; (Admx_A) and (Admx_M) admixed accessions between subpopulations (ancestry composition q <0.7 at K=6). ** $p < 0.01$ using a two tailed student t-test with unequal variance. Plots A and C from Denning-James et al. (2025).

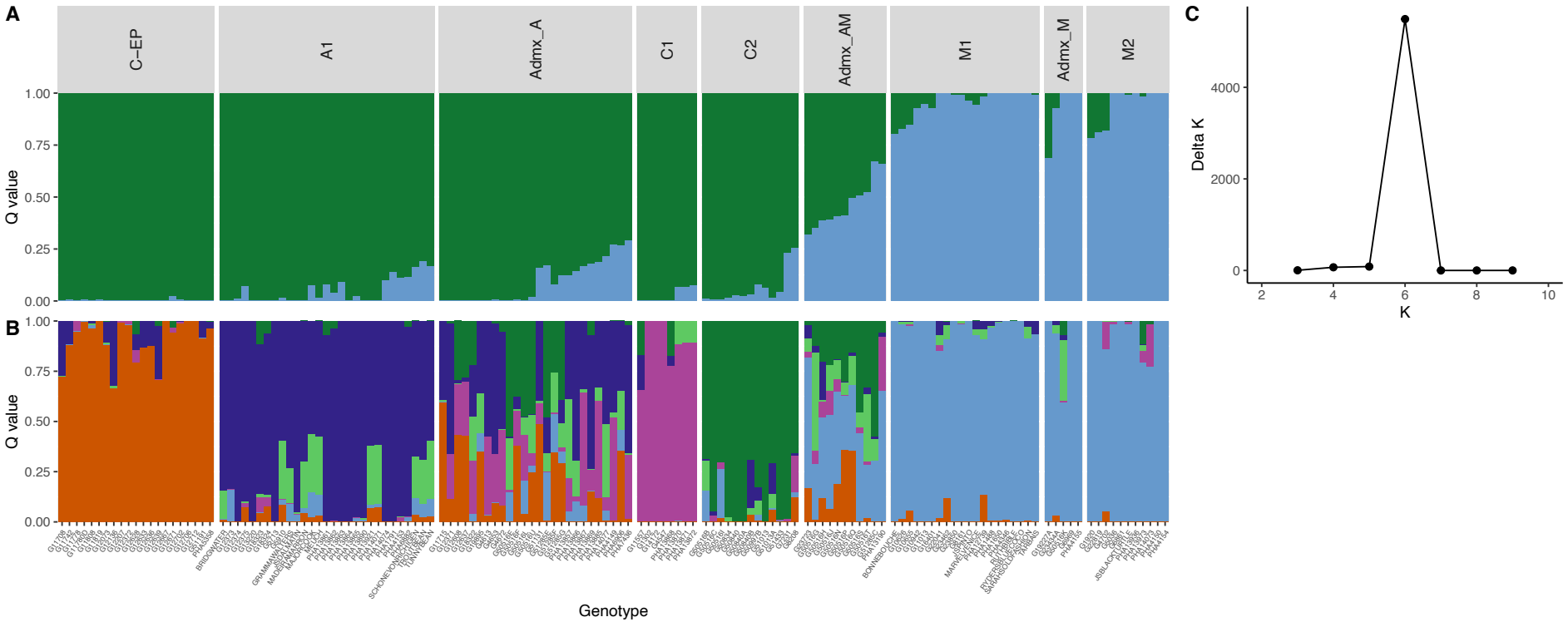


Figure 9; Analysis of the population structure of 144 accessions belonging to the diversity panel at K=2 (A) and K=6 (B) and the Delta K for K values 2-10 (C) ran using STRUCTURE v 2.3.5 (Pritchard et al. 2000). (A, B) Structured by ADMIXTURE K=6 results as in Figure 8 (Alexander and Lange 2011).

In the PCA (Figure 10A), PC 1 explained 38.8% of the variation in the diversity, splitting the two gene pools, while PC2 accounted for 5.06% of the variation, splitting the Mesoamerican subgroups (M1 and M2) and separating the Andean subpopulations (C-EP, A1 and Colombian). PC3 accounted for 3.67% variation (Figure 10B), separating C-EP from the other Andean subgroups and separating C2 from C1 and A1. When analysing the 108 Andean accessions from the diversity panel PC1 explained 10.3% of the variation, splitting the Colombian (C1 and C2) from C-EP and A1. PC2 explains 6.1% of the Andean variation, separating C-EP and A1 (Figure 11B).

A total of 11 accessions were classified as admixed between the Andean and Mesoamerican gene pools (Admx_AM), as they had an ancestry composition lower than 70% from either of the origins ($q<0.7$). The Admx_AM accessions were all indeterminate and produced a variety of seed sizes. Seven were landraces and two were wild. There was also a mix of photoperiod sensitive and insensitive accessions (Figure 10).

The Colombian subgroups (C1 and C2; Figure 10E) contained medium and large seeded landraces. However, the subpopulations distinguished by determinacy; C1 contained mainly insensitive determinate accessions while C2 contained sensitive indeterminate accessions (Figure 10C and Figure **10**D). The A1 group contained large and medium seeded landraces that were mainly photoperiod insensitive.

The C-EP population contained accessions from Ecuador, Peru and Colombia (Figure 10F). This group contained large-seeded indeterminate landraces and included accessions from races previously identified to be from the Andean gene pool. The Mesoamerican subgroups (M1 and M2; Figure 10) were also distinguished by phenotypic data. They both contained indeterminate and determinate accessions; however, M1 were mainly medium seeded while M2 were mainly small seeded. This is summarised in

Table **1** and Appendix 1. Colombian accessions can be found within all the subgroups and admixed groups at K=6 (Figure 10F). While the admixture accessions are mainly from Colombia, while one sample is a wild 'Ecuador' accession.

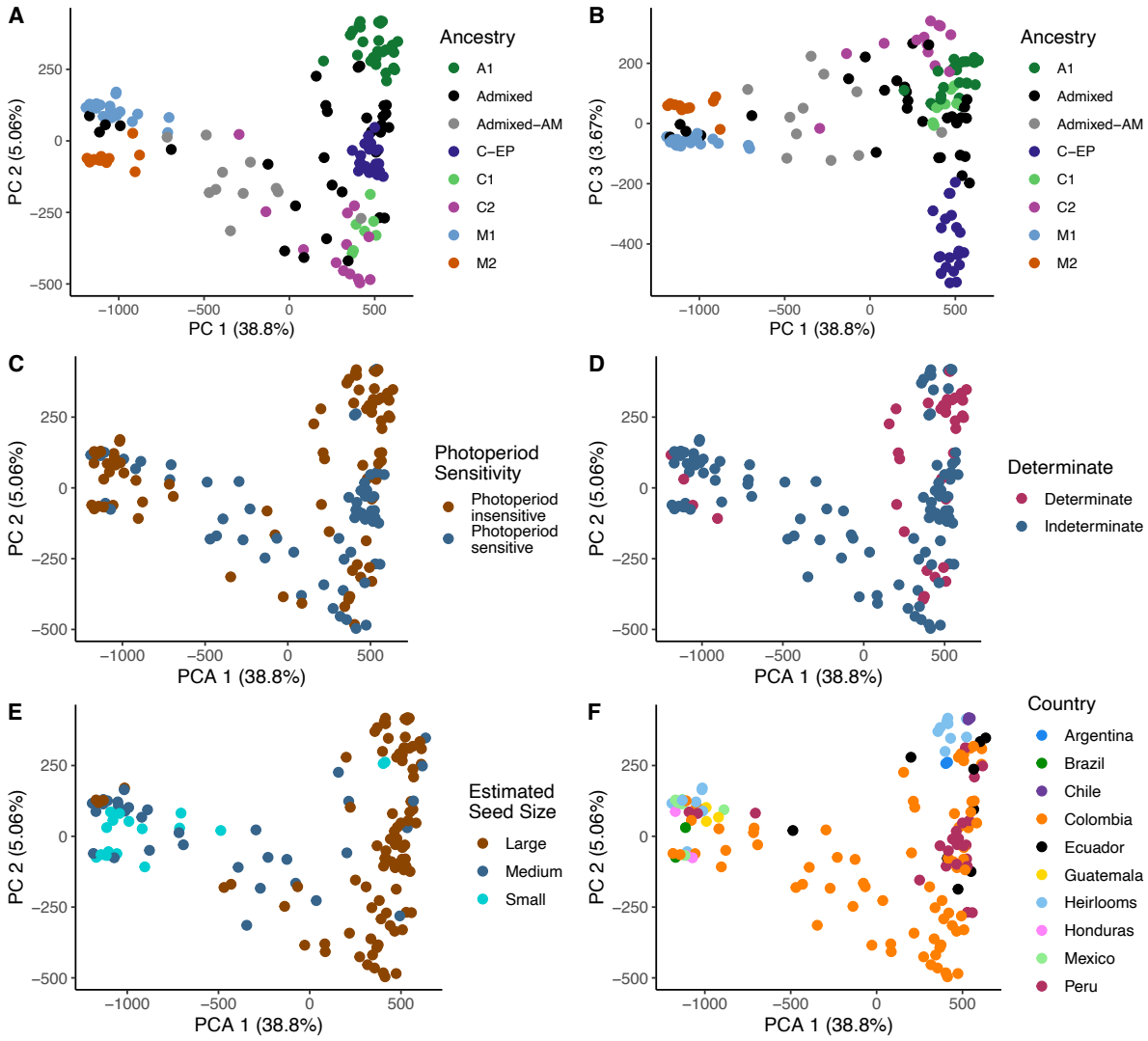


Figure 10; Principal component results on all 144 accessions from the diversity panel from GAPIT v.3 (Wang and Zhang 2021). (A) Principal component analysis (PCA) plot of PC1 against PC2 and (B) plot of PC1 against PC3. The colours illustrate the population structure of the diversity panel. PC1 against PC2 coloured by (C) photoperiod insensitivity, (D) determinacy, (E) estimated seed size and (F) country of sample collection (Appendix 1).

The population structure of the diversity panel can also be seen when analysing the proportion of heterozygous sites (Figure 11A). The Andean accessions had a lower proportion of heterozygous sites (<0.1) than the Mesoamerican accessions, which were more heterozygous. The six highly heterozygous accessions (>25% of the loci) were found within the Andean X Mesoamerican hybrid (Admixed-AM) subpopulation and were from Colombia, the region *Liborina*. Also, the two weedy accessions have a high proportion of heterozygous sites (>0.19) and the outlier accession with the lowest alignment to the Andean reference genome and low proportion of heterozygous sites was a wild accession from Ecuador.

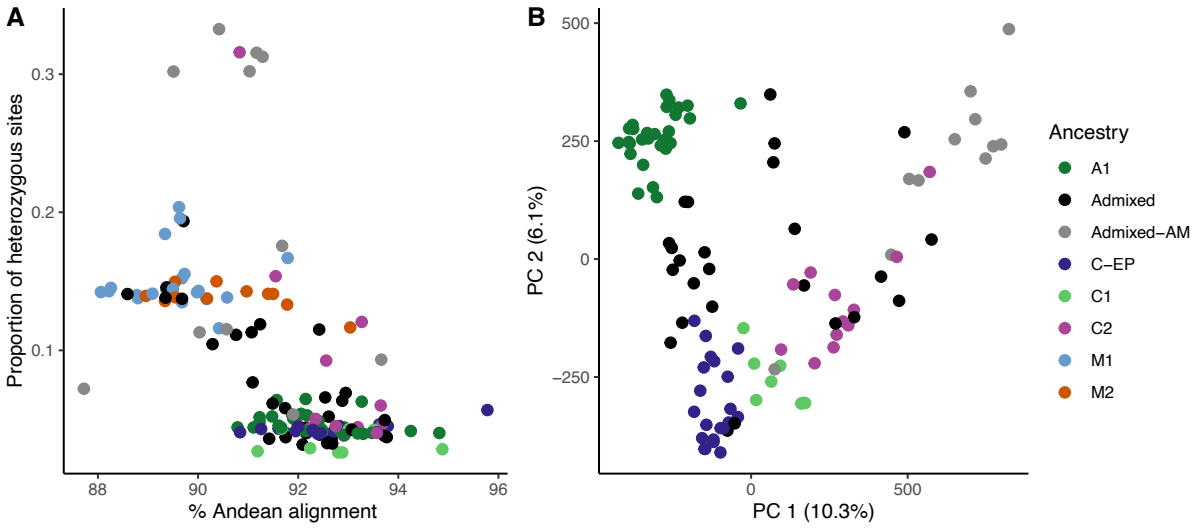
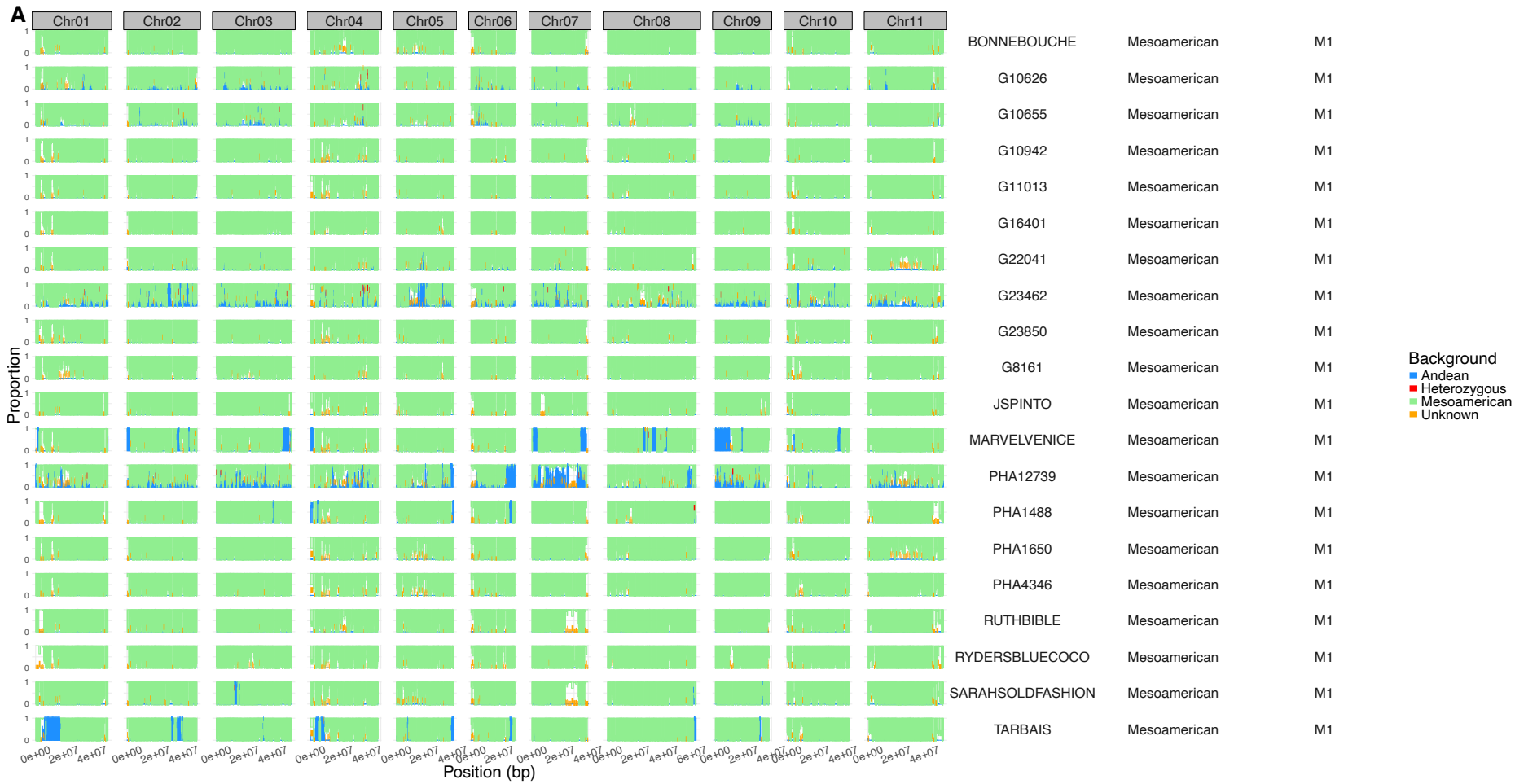


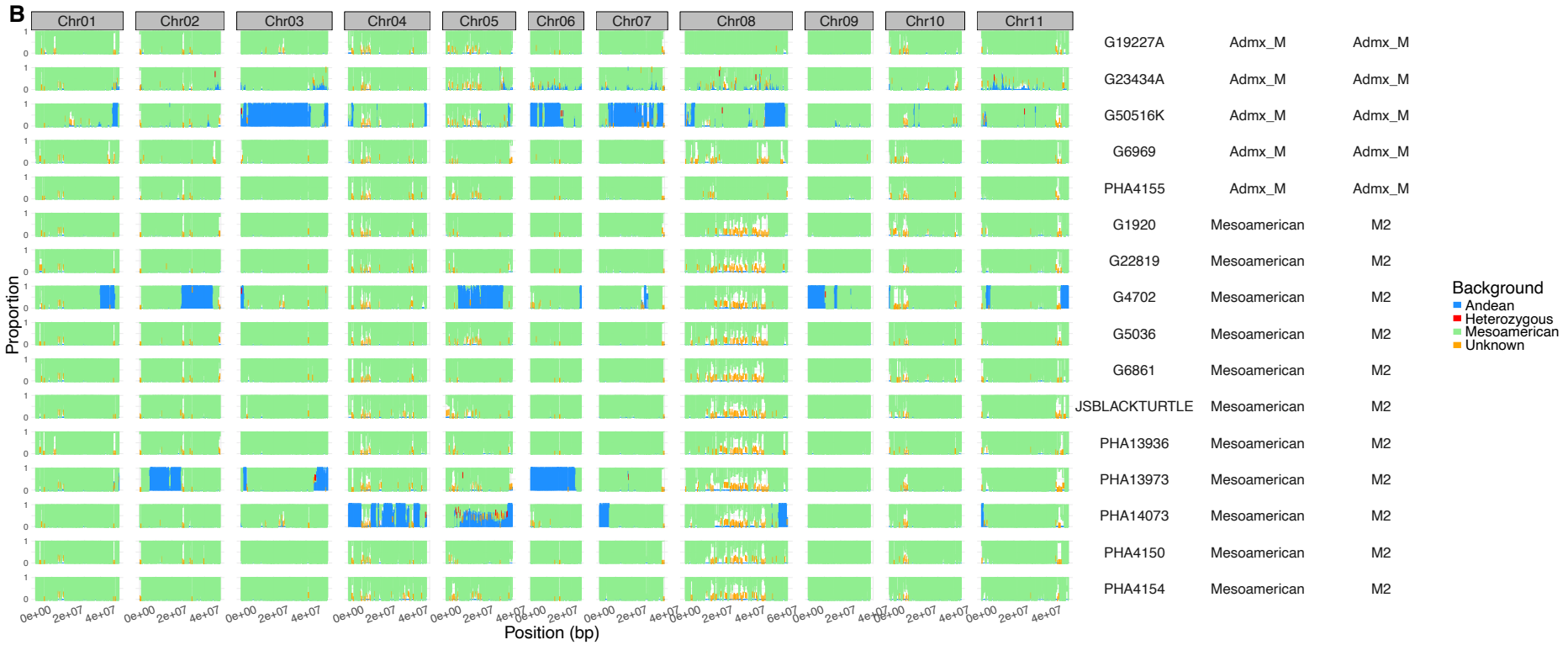
Figure 11; (A) Proportion of heterozygous sites against the percentage of read-pair alignment to the Andean reference genome G19833 for 144 accessions (Schmutz et al., 2014). (B) Principal component results on the 108 Andean accessions from the diversity panel ran using GAPIT v.3 (Wang and Zhang 2021). The colours illustrate the population structure of the diversity panel from ADMIXTURE K=6 (Alexander and Lange 2011).

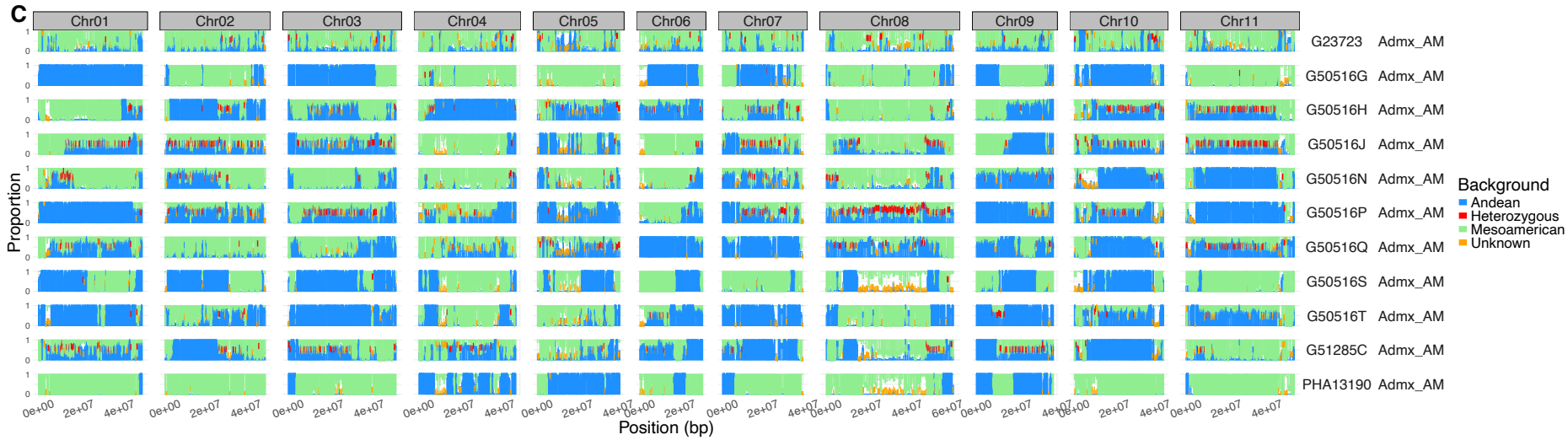
Table 1; Phenotypic characteristics associated with each subpopulation for the diversity panel (144 accessions). Table from Denning-James et al. (2025).

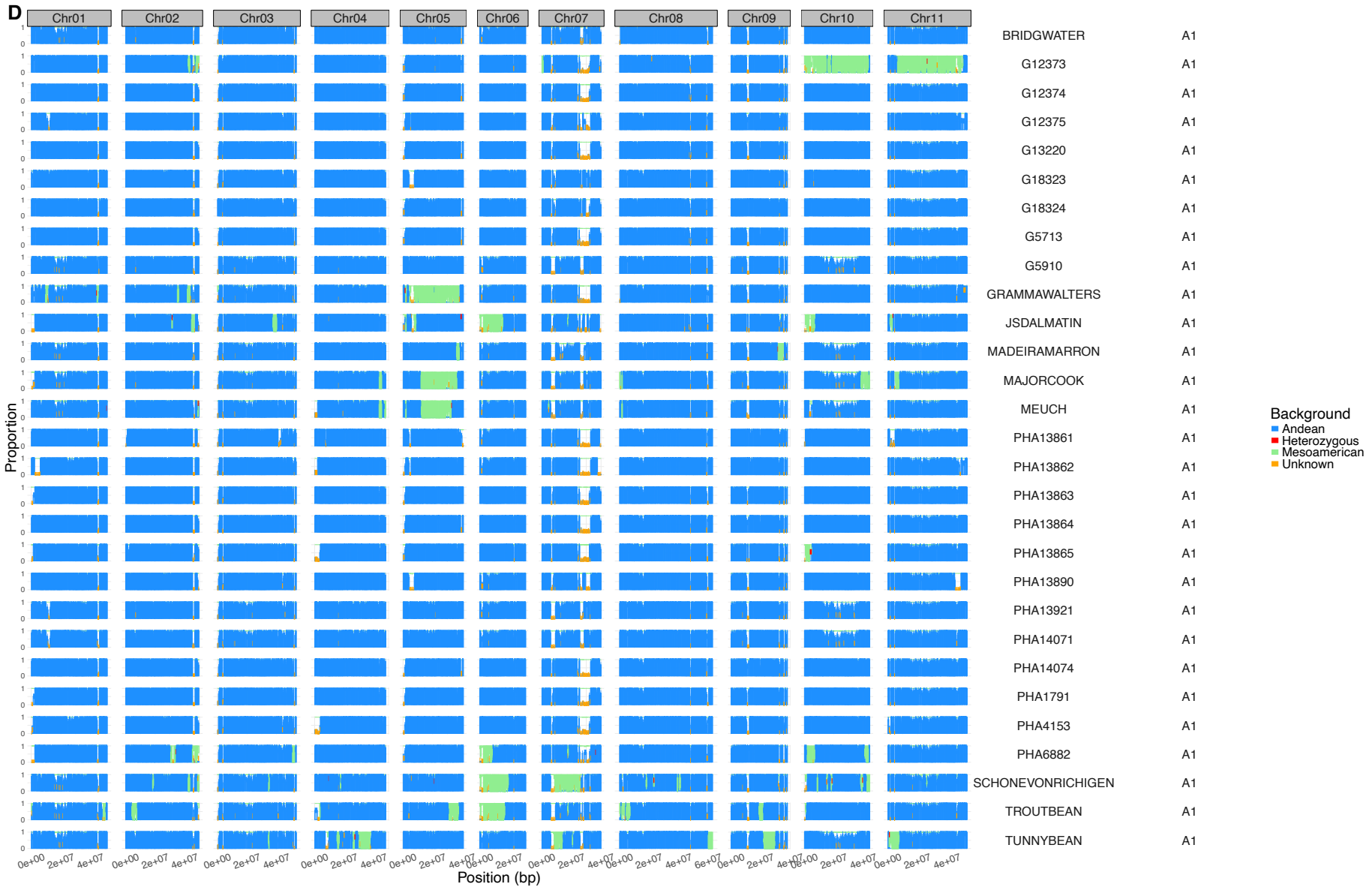
Subpopulation	Gene pool	Determinate	Photo. Sen.	Seed size	Origin
C1	Andean	Mainly determinate	Insensitive	Mainly large	Colombia and Ecuador
C2	Andean	Indeterminate	Mainly sensitive	Mainly large	Colombia
A1	Andean	Both	Mainly insensitive	Mainly large	South America, Heirlooms, Colombia
C-EP	Andean	Indeterminate	Sensitive	Large	Colombia, Ecuador, Peru
Admix_A	Andean	Mainly indeterminate	Both	Mainly large	Colombia and South America
M1	Mesoamerican	Mainly indeterminate	Both	Mainly medium**	Central America, Colombia, Heirlooms, Peru
M2	Mesoamerican	Mainly indeterminate	Mainly insensitive	Mainly small**	Central America, Colombia
Admix_M	Mesoamerican	Mainly indeterminate	Insensitive	Small and medium	Colombia, Brazil, Heirlooms, Central America
Admix_AM	AxM hybrids	Indeterminate	Mainly sensitive	Mainly medium	Colombia and Ecuador

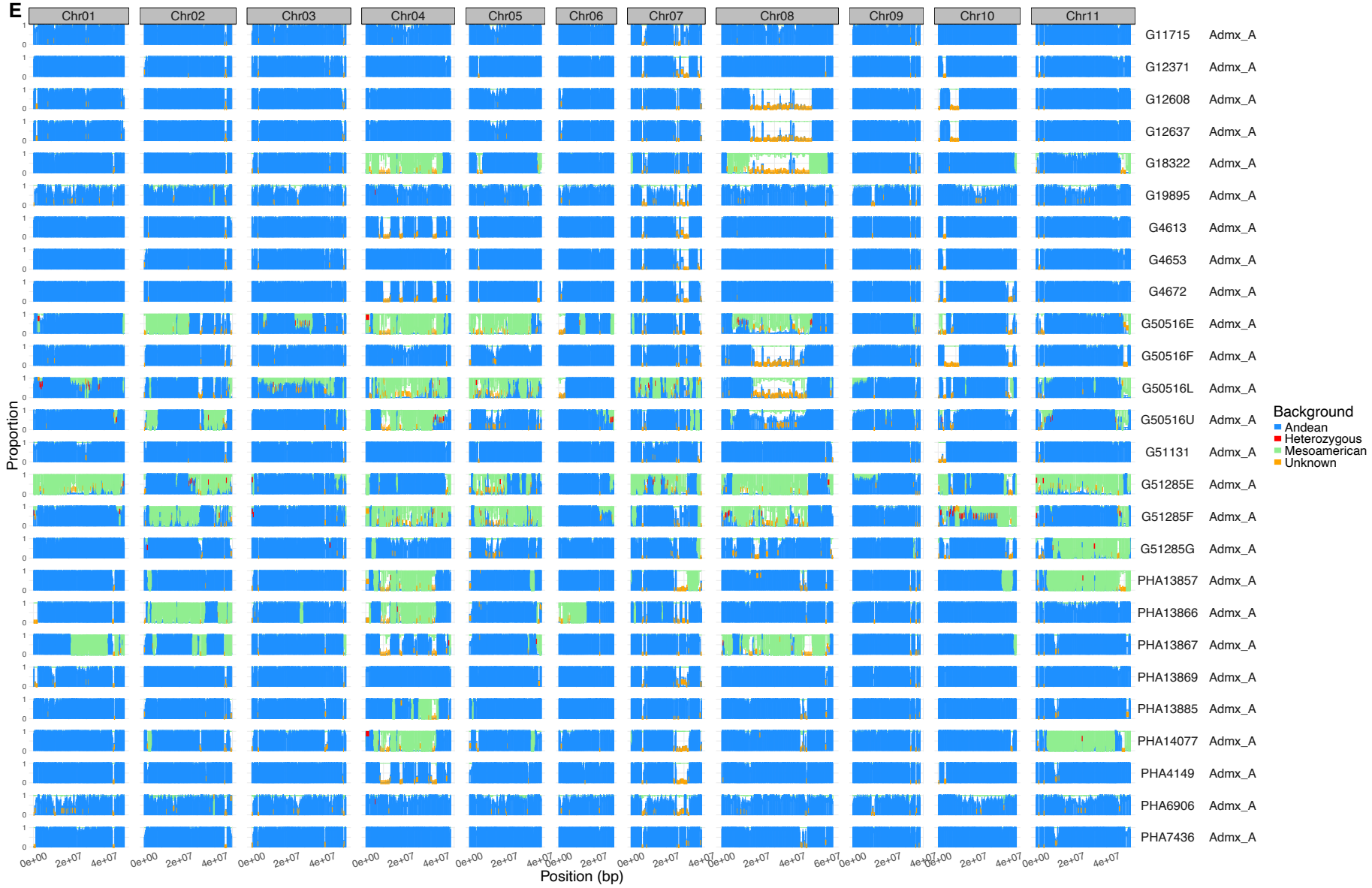
2.4.3 Introgression analysis




B
Chr01
Chr02
Chr03
Chr04
Chr05
Chr06
Chr07
Chr08
Chr09
Chr10
Chr11
Proportion
Position (bp)
G19227A Admx_M Admx_M
G23434A Admx_M Admx_M
G50516K Admx_M Admx_M
G6969 Admx_M Admx_M
PHA4155 Admx_M Admx_M
G1920 Mesoamerican M2
G22819 Mesoamerican M2
G4702 Mesoamerican M2
G5036 Mesoamerican M2
G6861 Mesoamerican M2
JSBLACKTURTLE Mesoamerican M2
PHA13936 Mesoamerican M2
PHA13973 Mesoamerican M2
PHA14073 Mesoamerican M2
PHA4150 Mesoamerican M2
PHA4154 Mesoamerican M2
Background
Andean
Heterozygous
Mesoamerican
Unknown


C
Chr01
Chr02
Chr03
Chr04
Chr05
Chr06
Chr07
Chr08
Chr09
Chr10
Chr11
G23723 Admx_AM
G50516G Admx_AM
G50516H Admx_AM
G50516J Admx_AM
G50516N Admx_AM
G50516P Admx_AM
G50516Q Admx_AM
G50516S Admx_AM
G50516T Admx_AM
G51285C Admx_AM
PHA13190 Admx_AM
Proportion
Position (bp)
Background
Andean
Heterozygous
Mesoamerican
Unknown


D
Chr01
Chr02
Chr03
Chr04
Chr05
Chr06
Chr07
Chr08
Chr09
Chr10
Chr11
BRIDGWATER A1
G12373 A1
G12374 A1
G12375 A1
G13220 A1
G18323 A1
G18324 A1
G5713 A1
G5910 A1
GRAMMAWALTERS A1
JSDALMATIN A1
MADEIRAMARRON A1
MAJORCOOK A1
MEUCH A1
PHA13861 A1
PHA13862 A1
PHA13863 A1
PHA13864 A1
PHA13865 A1
PHA13890 A1
PHA13921 A1
PHA14071 A1
PHA14074 A1
PHA1791 A1
PHA4153 A1
PHA6882 A1
SCHONEVONRICHIGEN A1
TROUTBEAN A1
TUNNYBEAN A1
Proportion
Position (bp)
Background
Andean
Heterozygous
Mesoamerican
Unknown


E
Chr01
Chr02
Chr03
Chr04
Chr05
Chr06
Chr07
Chr08
Chr09
Chr10
Chr11
G11715 Admx_A
G12371 Admx_A
G12608 Admx_A
G12637 Admx_A
G18322 Admx_A
G19895 Admx_A
G4613 Admx_A
G4653 Admx_A
G4672 Admx_A
G50516E Admx_A
G50516F Admx_A
G50516L Admx_A
G50516U Admx_A
G51131 Admx_A
G51285E Admx_A
G51285F Admx_A
G51285G Admx_A
PHA13857 Admx_A
PHA13866 Admx_A
PHA13867 Admx_A
PHA13869 Admx_A
PHA13885 Admx_A
PHA14077 Admx_A
PHA4149 Admx_A
PHA6906 Admx_A
PHA7436 Admx_A
Proportion
Position (bp)
Background
Andean
Heterozygous
Mesoamerican
Unknown


F
Chr01
Chr02
Chr03
Chr04
Chr05
Chr06
Chr07
Chr08
Chr09
Chr10
Chr11
Proportion
Position (bp)
G11557 Colombian C1
G1302 Colombian C1
G14172 Colombian C1
G4547 Colombian C1
PHA13868 Colombian C1
PHA13870 Colombian C1
PHA13871 Colombian C1
PHA13872 Colombian C1
G50516B Colombian C2
G50516C Colombian C2
G50516I Colombian C2
G50834 Colombian C2
G50840 Colombian C2
G50840A Colombian C2
G50840B Colombian C2
G50997B Colombian C2
G51013 Colombian C2
G51013A Colombian C2
G51433 Colombian C2
G7231 Colombian C2
G8208 Colombian C2
Background
Andean
Heterozygous
Mesoamerican
Unknown

Figure 12; Introgression analysis on all 144 accessions from the diversity panel using VcfHunter (Martin et al. 2023) split by ADMIXTURE K=6 population structure. (A) M1, (B), M2 and admixed M, (C) admixed between Andean and Mesoamerican, (D) A1, (E) admixed A, (F) C1 and C2 and (G) C-EP. The colours represent the 'ancestral' background, and the staggered columns (y-axis) reflects the proportion (0-1) of SNPs in that window across the genome (x-axis) from each background, calculated for each accession. Accession name and subpopulation names at, K=4 and K=6 population structure from ADMIXTURE, respectively, were labelled.

The subpopulations M1, M2 and admixed M were classified as Mesoamerican (Figure 8, Figure 12A and B), because they had a proportion of diversity from the Mesoamerican background over 0.7. Within these three Mesoamerican subpopulations, 39% of accessions had a proportion of diversity from the Andean background (<0.1, <0.15 and <0.25 respectively). There were no significant differences among the M1, M2 and admixed M subpopulations regarding the proportion of diversity from the Andean and Mesoamerican backgrounds (Figure 13A). However, three accessions from M2 (G4702, PHA13973 and PHA14073) have a higher proportion of Andean diversity in chromosomes 2 and 5.

There was a significant difference ($p<0.01$) in the proportion of diversity from an unknown origin between M1 and M2 in chromosome 8 from ~20Mbp to ~50Mb (Figure 13B). This region of unknown diversity on chromosome 8 in M2 can also be found in other subpopulations such as 'Admixed AM' (G50516S and PHA13190), 'Admixed A' (G12608, G12637, G18322, G50516F and G50516L), all C1 accessions and all except 3 (G50840B, G50997B and G51013A) of the C2 accessions. There was no significant difference between M1 and M2 on any other chromosome.

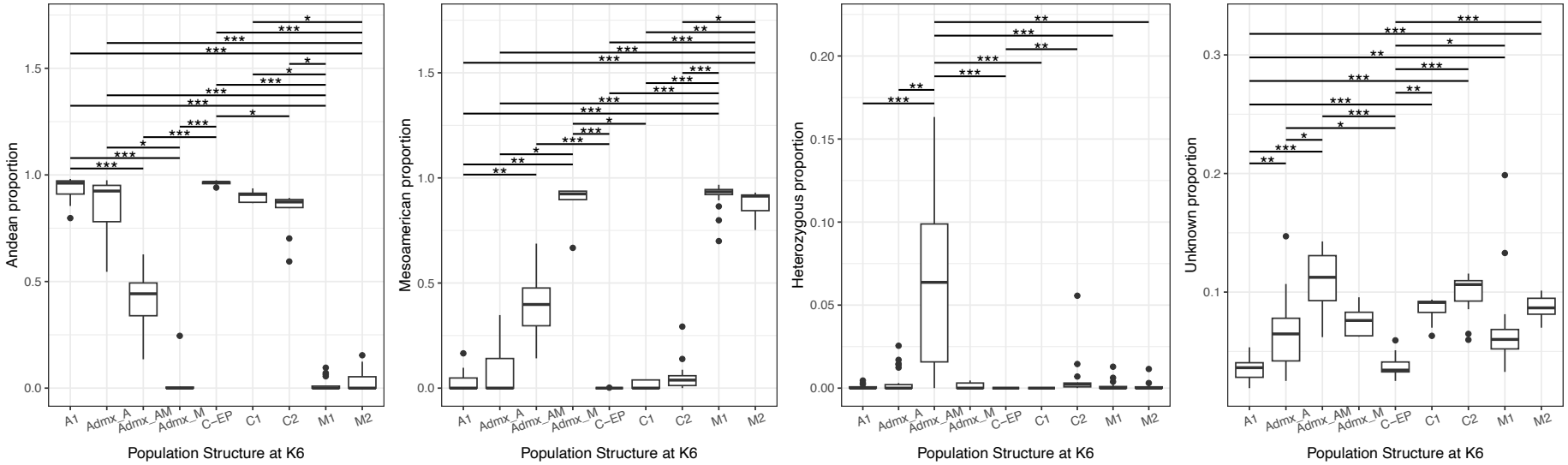
Among the 'admixed AM' accessions the proportion of diversity from the Mesoamerican background varied from 0.68 to 0.14 while the Andean background varied from 0.63 to 0.14 (<0.7)(Figure 12C, Figure 13A). The 'Admixed AM' accessions have a higher proportion of heterozygous diversity across their genomes (chromosomes 1, 2, 4, 9 and 10, $p<0.05$) compared to all other subpopulations at K=6 (Figure 13C).

The A1 and C-EP subpopulation have a proportion of Andean diversity >0.8, >0.94, respectively. As well as a proportion from the Mesoamerican background <0.17 and <0.002 and low levels of heterozygosity (<0.005, 0, respectively) (Figure 13). A1 was not significantly different to the other Andean or Colombian subpopulations (admixed A, C-EP, C1 or C2) when analysing the proportions of Andean and Mesoamerican diversity (Figure 13A). However, A1 and C-EP have significantly lower levels of unknown diversity (<0.05, <0.06) compared to C1, C2 and admixed A (Figure 13B).

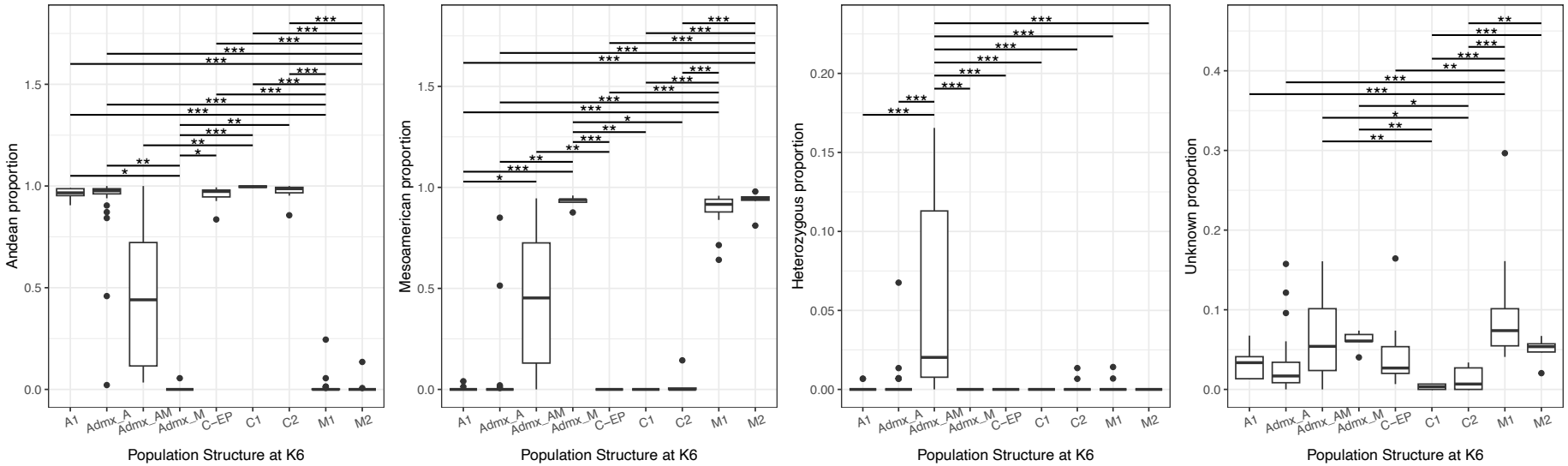
Four A1 accessions (Troutbean, Schonevonrichigen, PHA6882 and JSDalmatin) have a Mesoamerican introgression at the start of chromosome 6 from ~0Mbp to ~20Mbp (Figure 12D). This can also be found in one 'admixed A' accession (PHA13866, Figure 12E) and one C2 (G50516I, Figure 12F). One A1 accession (G12373) had a Mesoamerican introgression, this can also be found in one C2 (G50516I) and in four 'admixed A' accessions (G51285E, G51285G, PHA13857 and PHA14077).

The C1 and C2 accessions have no significance difference when analysing proportions across the genomes (Figure 13A). Further to this there were also no differences between C1 and C2 for their proportions of Andean, or heterozygosity by chromosomes (Figure 13B). However, in chromosome 4, three C1 accessions (PHA1870, PHA13871 and PHA13872) have a Mesoamerican introgression, and another three C1 accessions (G1302, G14172 and G4547) have a large section of unknown diversity. Also, 77% of C2 accessions have a section of unknown diversity at the start of chromosome 6.

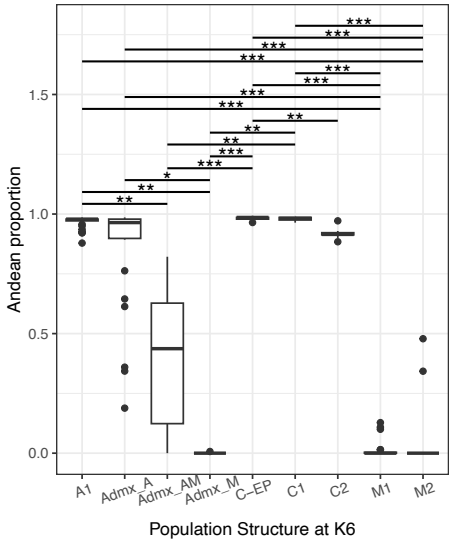
A Whole Genome

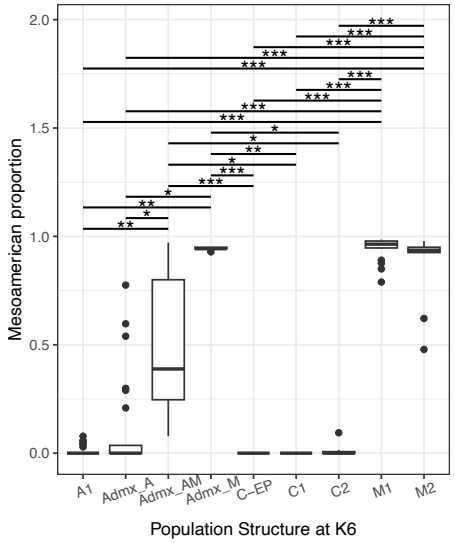


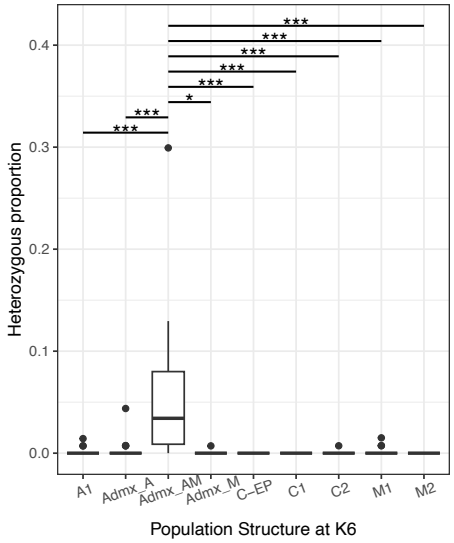
B Chr01

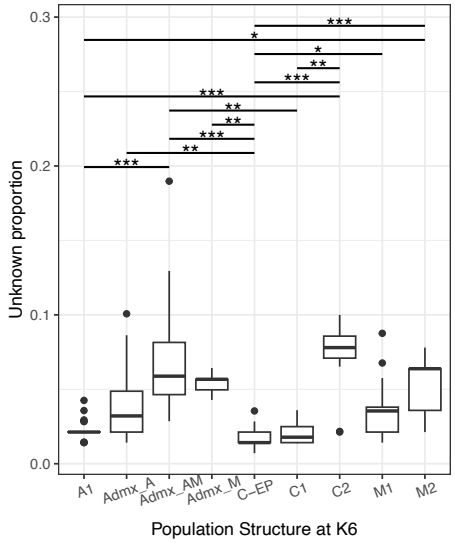


Chr02

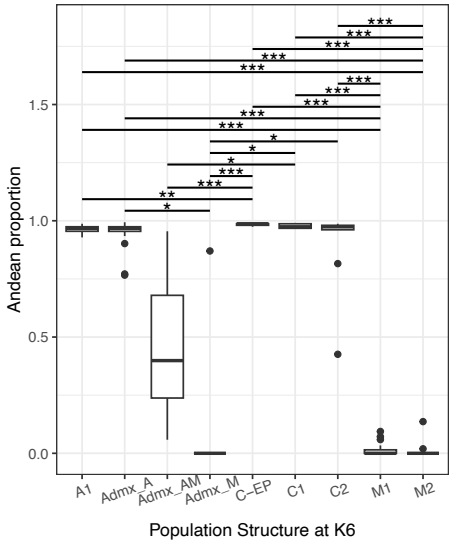

Andean proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

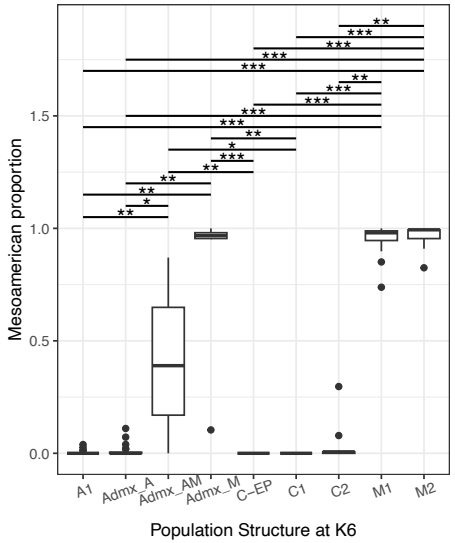

Mesoamerican proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

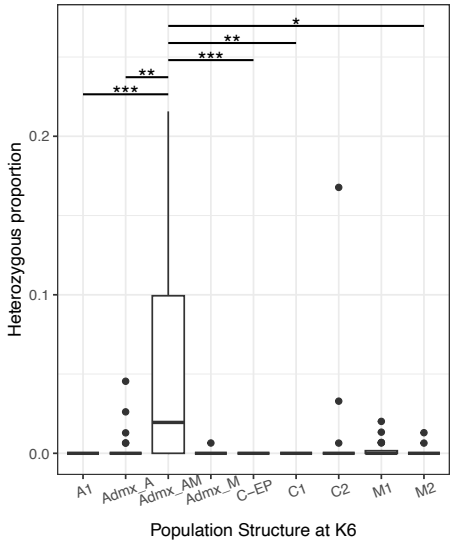

Heterozygous proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

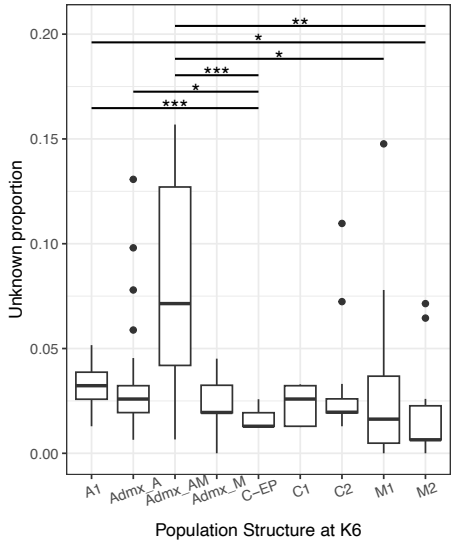

Unknown proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

Chr03

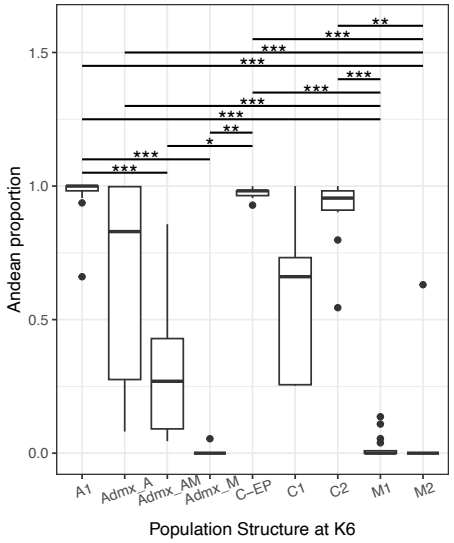

Andean proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

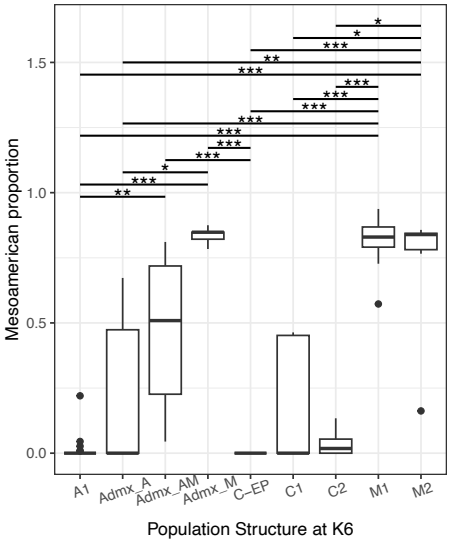

Mesoamerican proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

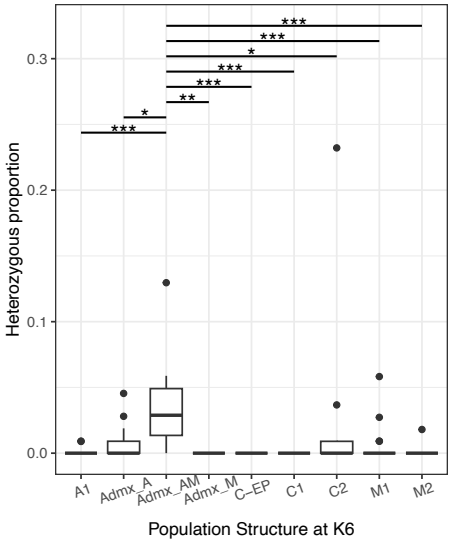

Heterozygous proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

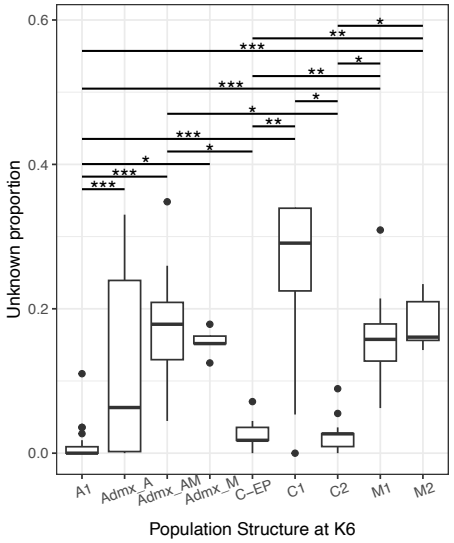

Unknown proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

Chr04

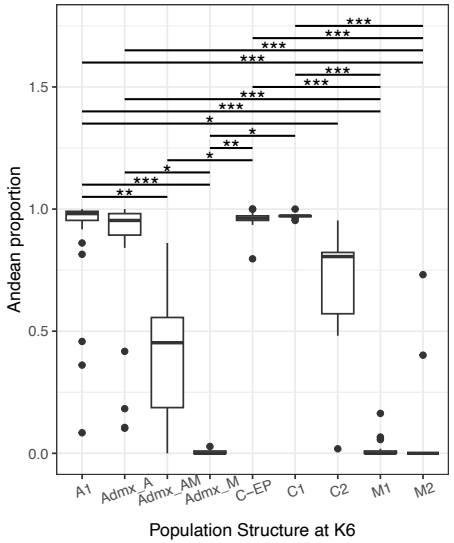

Andean proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

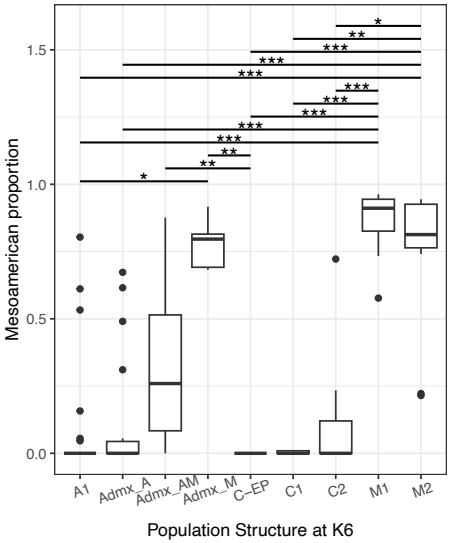

Mesoamerican proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

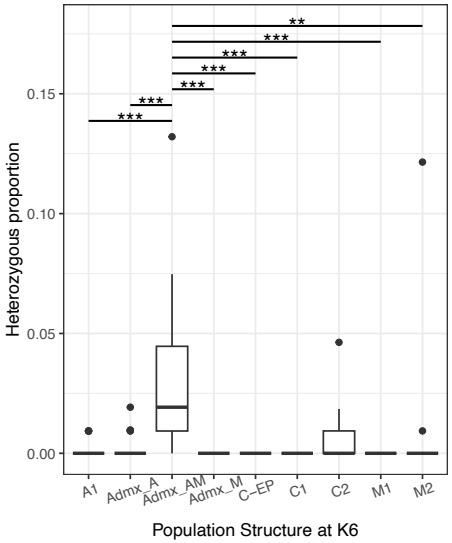

Heterozygous proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

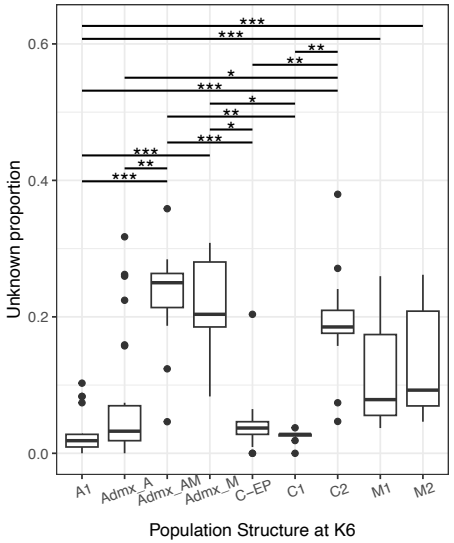

Unknown proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

Chr05

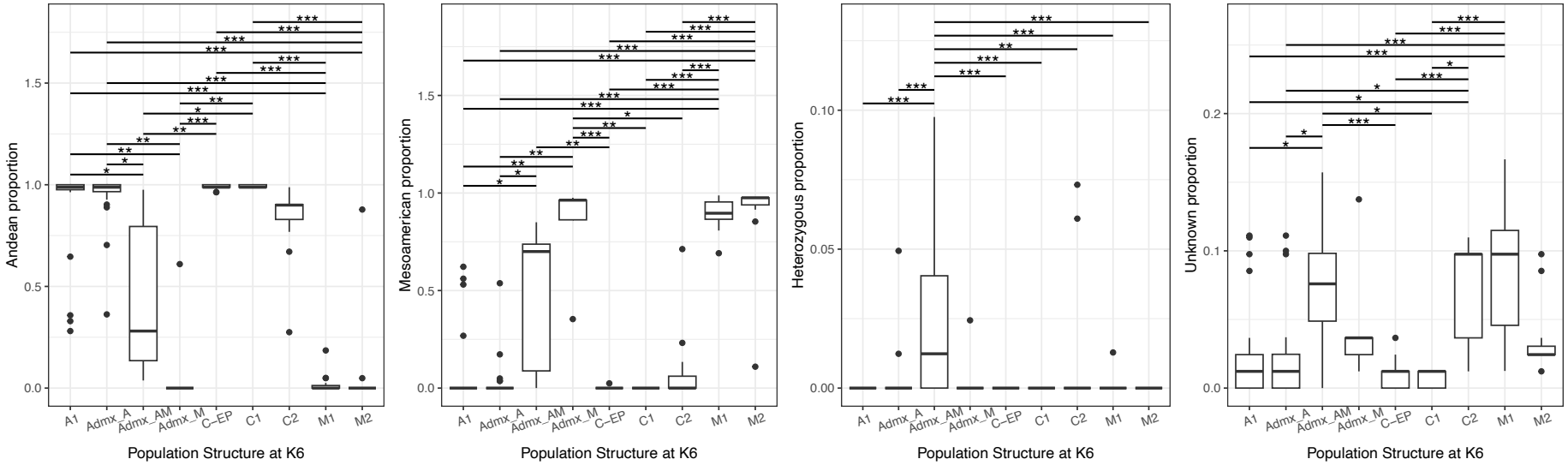

Andean proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2


Mesoamerican proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

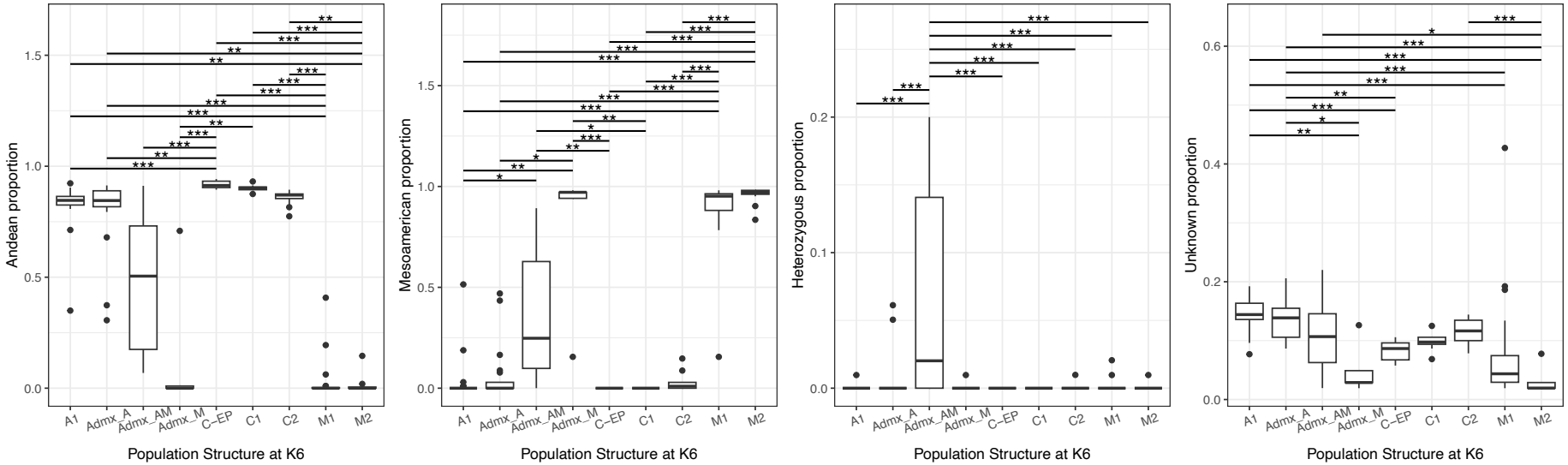

Heterozygous proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2


Unknown proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

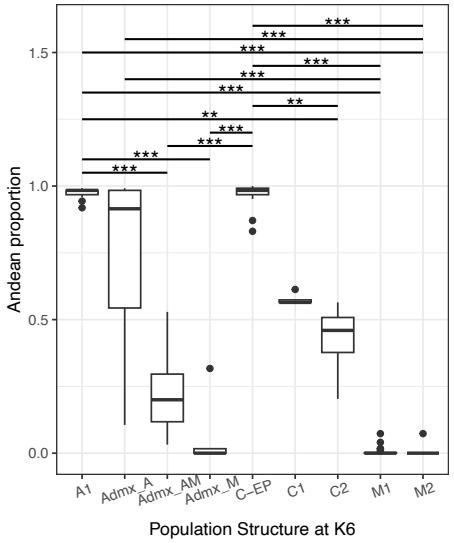
Chr06

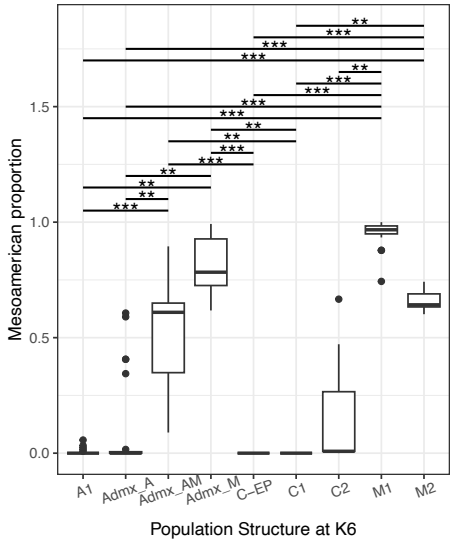

Andean proportion
Mesoamerican proportion
Heterozygous proportion
Unknown proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

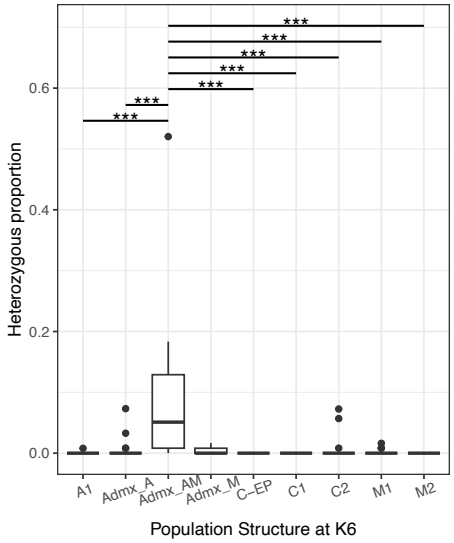
Chr07

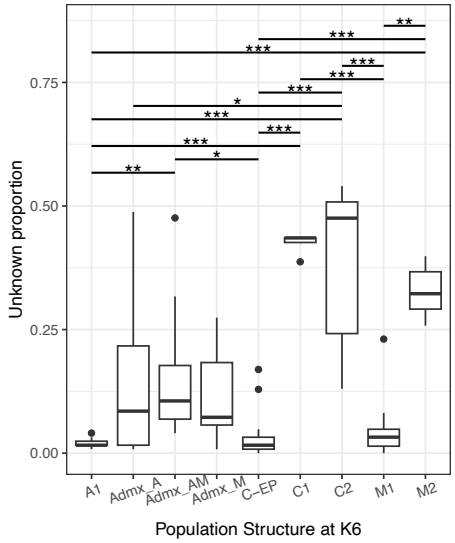

Andean proportion
Mesoamerican proportion
Heterozygous proportion
Unknown proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

Chr08

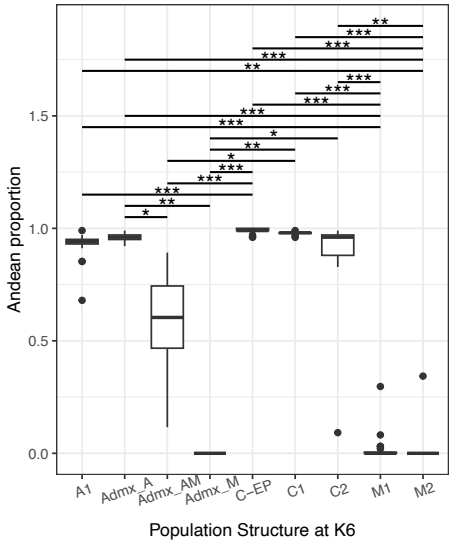

Andean proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

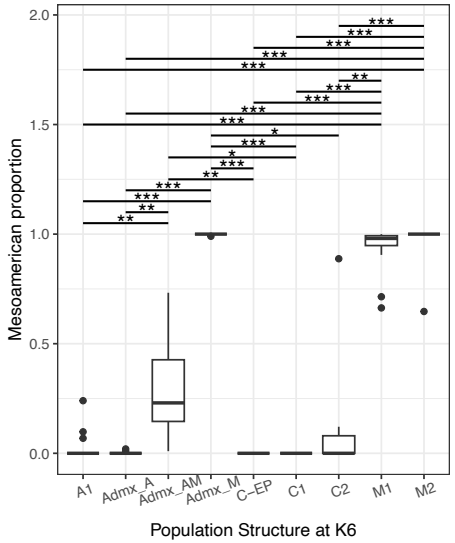

Mesoamerican proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

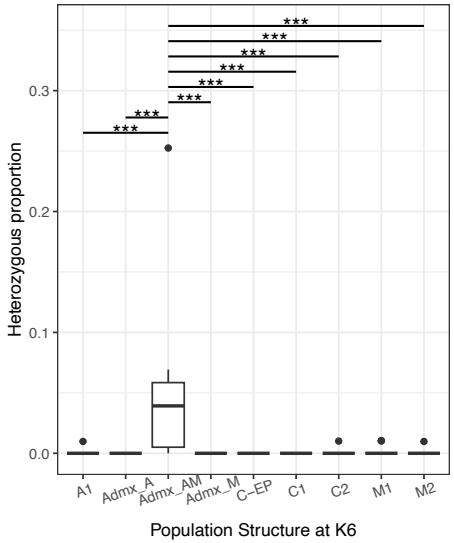

Heterozygous proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

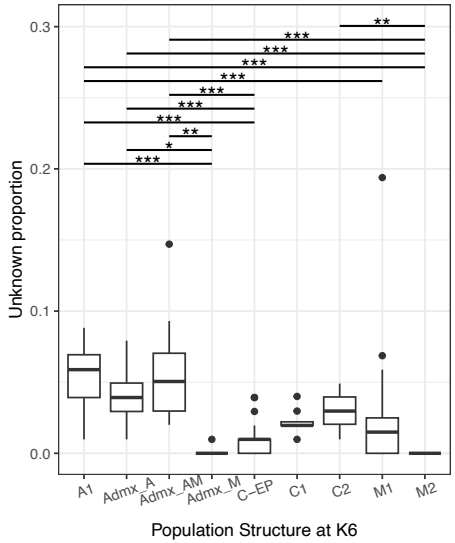

Unknown proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

Chr09


Andean proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2


Mesoamerican proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2


Heterozygous proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2


Unknown proportion
Population Structure at K6
A1
Admx_A
Admx_AM
Admx_M
C–EP
C1
C2
M1
M2

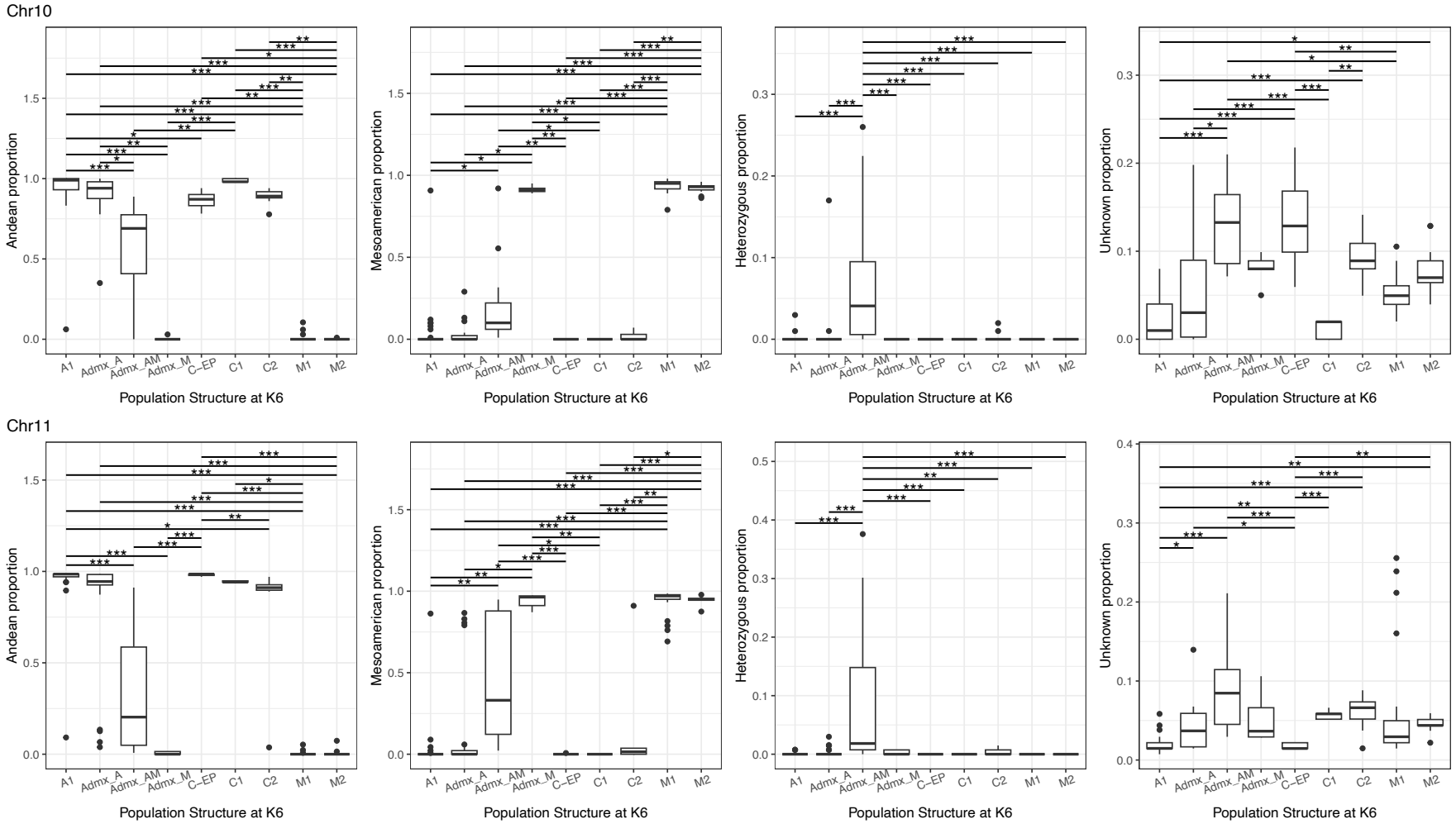


Figure 13; Comparison of genetic landscape metrics, namely, proportion of Andean, Mesoamerican, heterozygous or unknown background content, among subpopulations for the diversity panel (144 accessions). (A) Comparison across the whole genomes, (B) comparisons calculated per chromosome. Grouping by the admixture K=6 population structure. Significance was calculated with the Dunn test and Bonferroni correction (* p <0.05, ** p<0.01, *** p<0.001).

2.4.4 Phenotypic variation and correlations

The correlation coefficient was estimated for each pair of traits (Figure 14), averaged over two seasons or studied in both years. There was a positive correlation between DTF from winter and summer (r=0.57) (Figure 14C). Both DTF were associated with PS (η^2= 0.64 (Win_DTF), η^2 = 0.51 (Sum_DTF)) and D (η^2 = 0.18 (Win_DTF), η^2 = 0.13 (Sum_DTF)) (Figure 14A). Population structure at either two or six ancestries (K2, K6) were associated with D (*V* = 0.64 (K6), *V* = 0.29 (K2)), SS (*V*= 0.61 (K6), *V* =0.52(K2)) (Figure 14B), Est_SW (η^2= 0.51 (K6), η^2= 0.44 (K2) and major seed colour (*V*=0.31 (K6), *V*=0.35(K2)). SS was not correlated with Sum_DTF (η^2= 0.02), Win_DTF (η^2= 0.01), D (η^2= 0) or PS (η^2= 0) (Figure 14A). Est_SW was associated with SS (η^2= 0.76), major seed colour (η^2= 0.26), country (η^2= 0.21) and type (η^2= 0.19) but not with Win_DTF (η^2= 0) or Sum_DTF (η^2=0) (Figure 14A). Then D and PS were associated (*V* = 0.47) (Figure 14B). Major seed colour associated with D (*V*=0.47), PS (*V*=0.41), SS (*V*=0.4) and type (*V*=0.37)

Figure 15A, B and C showed the distributions of the phenotyping for traits E100_SW, S22_DTF and W23_DTF, respectively. The seed weights (Figure 15A) were normally distributed, while the DTF in summer and winter (Figure 15B and C) were binomial distributions; the peaks were around 42- and 54-days post-sowing in summer, and around 70- and 90- days in winter. When analysing the phenotypes by subpopulation, it can be seen that C-EP (Figure 15F) did not flower during winter in the UK, W23_DTF, as they were mainly photoperiod sensitive. This is further supported by the correlation plot (Figure 14). Furthermore, determinacy, photoperiod insensitivity and days to flowering are correlated. The determinate accessions flower earlier than the indeterminate, supporting the binomial distribution.

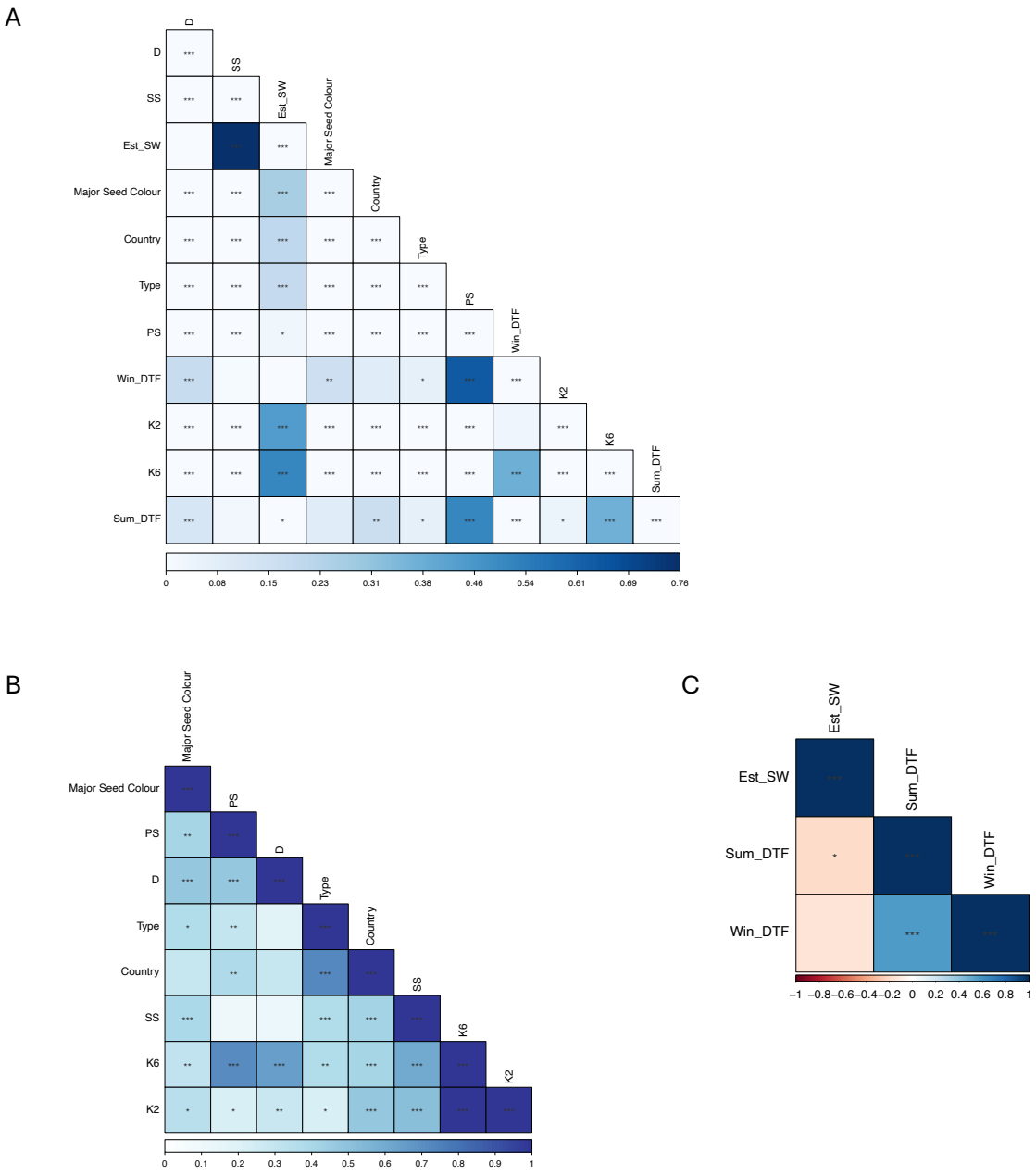


Figure 14; Correlation plots produced in R using the 'corrplot' package (Wei and Simko 2021). Phenotypes are agronomic traits measured in 144 common bean genotypes grown at the Norwich Research Park, Norwich, UK in 2022 and 2023, results of the population structure analysis and data from gene bank passport data (IPK and CIAT) (Appendix 1). (A) Associations between discrete and continuous phenotypes assessed with a one-way ANOVA and measured with eta-squared. (B) Correlation between discrete datasets, the association was calculated with Cramer Vs and significance with Chi-squared (Mangiafico. 2025). (C) Pearson correlation coefficients among continuous datasets with normal distribution. (K6) K6 subgroups from ADMIXTURE; (K2) K2 subgroups from ADMIXTURE; (D) Determinacy; (SS) Seed size; (Est_SW) Estimated weight of 100 seeds; (PS) photoperiod sensitivity; (Win_DTF) days to flowering from winter 2023; (Sum_DTF) days to flowering from summer 2022. *Indicates $p < 0.05$; **indicates $p < 0.01$; ***indicates $p < 0.001$.

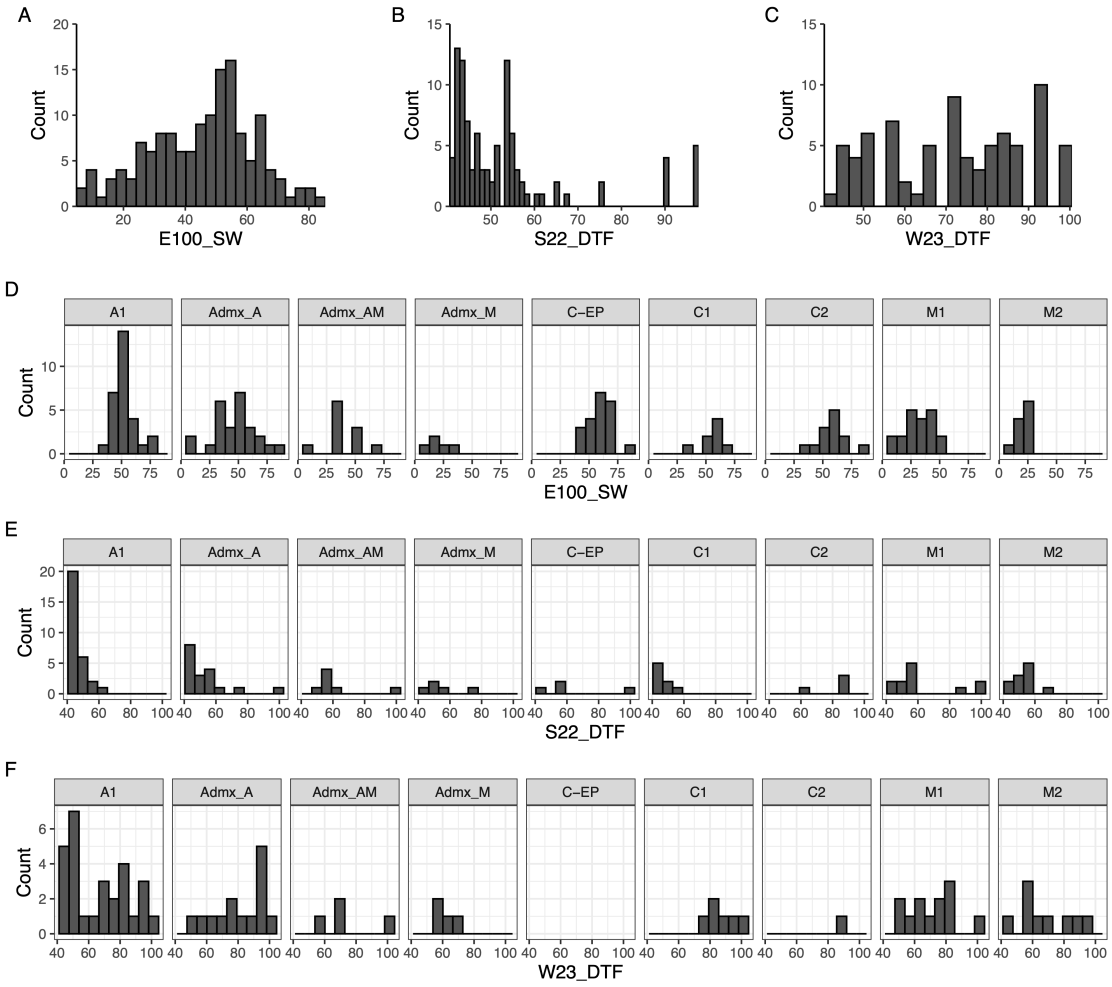


Figure 15; Frequency distribution of seed weight and days to flower traits evaluated in two seasons in a common bean diversity panel (144 accessions). (A) E100_SW, estimated weight of 100 seeds; (B) phenological days to flowering in the summer 2022 (S22_DTF) (and (C) in the winter 2023 (W23_DTF) at the Norwich Research Park, excluding those which did not flower. The distributions were split into the subpopulations from K6 ADMIXTURE. (D) E100_SW***; (E) S22_DTF***; (F) W23_DTF*. Completed a one-way ANOVA for E100_SW, S22_DTF and W23_DTF. *Indicates $p < 0.05$; **indicates $p < 0.01$; ***indicates $p < 0.001$. Figure from Denning-James et al. (2025).

2.4.5 Environmental analysis

The daylength during the trials varied from 11.4-16.9 hours a day (Figure 16A) and from 7.6-11.8 hours a day (Figure 16B) for the long daylength and short daylength trials, respectively. The maximum temperature for the short daylength trial was 26°C, and the minimum was 14.7°C (Figure 16B). The maximum temperature for the long daylength trial was 37°C and the minimum was 15.3°C, both trials took place over 3 months.

The first accession to flower in the long daylength trial was on 19/04/2022 (42 days after sowing (DAS)), the average date for flowering was 30/04/2022 (53 DAS) and the final accession to flower was the 14/06/2022 (98 DAS, Figure 16A). During this period the maximum temperature was 28.6°C and the minimum was 15.3°C. For the short daylength trial the first accession to flower was on the 09/01/2023 (43 DAS), the average flowering was on the 08/02/2023 (73 DAS) and the final accession to flower was on the 07/03/2023 (100 DAS). During this period the maximum temperature was 25.1°C and the minimum was 14.7°C (Figure 16).

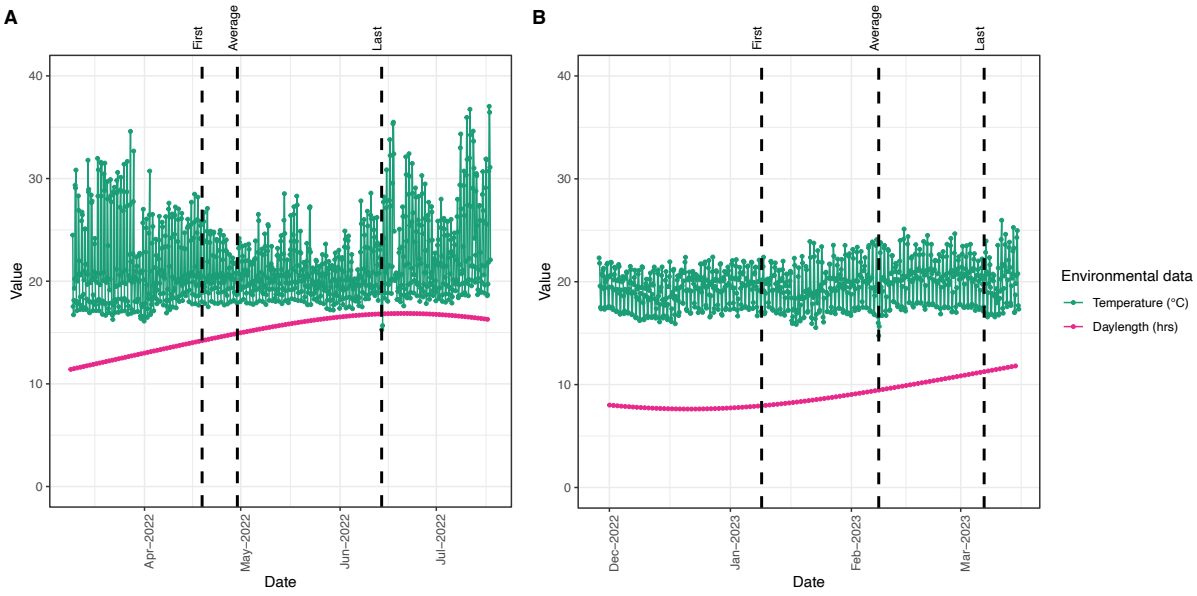


Figure 16; Plot represent (A) the long daylength trial from April 2022 to July 2022 and (B) the short daylength trial from Dec 2022 to March 2023. Data for daylength collected from CW4326, Norwich, UK station (52.643°, 1.237°) (Visual Crossing Corporation 2025) and data for temperature collected from within the glasshouse with the Gemini Tinytag, Ultra 2. The vertical dotted lines represent the first accession to flower, the average time for flowering (removing those which did not) and the last plant to flower.

2.4.6 GWAS for Determinacy

The GWAS was performed using the models BLINK, FarmCPU and MLM with GAPIT (Figure 17A and B) for the determinacy phenotype. The QQ plots (Figure 17C and D) provided evidence that the selected models were well fitted to identify significant marker trait associations (MTAs) for the dataset. While Figure 17E provided evidence of a poor model fit with the BLINK model, which was consequently not used for further analysis on the Andean genome. Thirteen MTAs were identified with a significant p-value, corresponding to thirteen QTLs. Seven significant MTAs were the focus of the analysis, that were identified for the whole panel based on the criteria laid out in the methods (vertical lines in Figure 17). The seven QTLs were found on chromosomes Pv01, Pv07, Pv08, Pv09 and Pv10 (Table 2). Five of the seven QTLS were also identified for the Andean subset. Putative candidate genes were identified for determinacy based on the significant MTAs and corresponding QTL windows. The identified genes and QTLs are listed in Appendix 2.

Table 2; 23 QTLs associated with determinacy and photoperiod sensitivity. Table from Denning-James et al. (2025).

Name	Chrom	Start	End	Trait	Panel
D1.1	Chr01	6,512,000	6,521,000	Determinacy	Andean + Whole
D1.2	Chr01	11,363,000	11,372,000	Determinacy	Andean
D1.3	Chr01	42,404,000	42,413,000	Determinacy	Andean + Whole
D1.4	Chr01	44,856,000	44,847,000	Determinacy	Whole
D1.5	Chr01	44,932,000	44,941,000	Determinacy	Andean + Whole
D1.6	Chr01	45,098,000	45,107,000	Determinacy	Whole
D2.1	Chr02	24,821,000	24,830,000	Determinacy	Andean
D3.1	Chr03	25,608,000	25,617,000	Determinacy	Andean
PS4.1	Chr04	38,316,000	38,325,000	Photo sensitivity	Whole
PS5.1	Chr05	16,423,000	16,432,000	Photo sensitivity	Whole
PS5.2	Chr05	18,321,000	18,330,000	Photo sensitivity	Andean
PS7.1	Chr07	16,829,000	16,838,000	Photo sensitivity	Andean + Whole
PS7.2	Chr07	26,485,000	26,494,000	Photo sensitivity	Andean + Whole
D7.1	Chr07	36,860,000	36,869,000	Determinacy	Andean + Whole
PS8.1	Chr08	4,234,000	4,243,000	Photo sensitivity	Whole
D8.1	Chr08	7,440,000	7,449,000	Determinacy	Andean
PS8.2	Chr08	8,320,000	8,329,000	Photo sensitivity	Andean
D8.2	Chr08	47,582,000	47,591,000	Determinacy	Whole
D9.1	Chr09	20,814,000	20,823,000	Determinacy	Andean + Whole
PS9.1	Chr09	21,640,000	21,649,000	Photo sensitivity	Whole
PS9.2	Chr09	34,445,000	34,454,000	Photo sensitivity	Andean
D10.1	Chr10	43,762,000	43,771,000	Determinacy	Andean + Whole
PS11.1	Chr11	204,000	213,000	Photo sensitivity	Andean

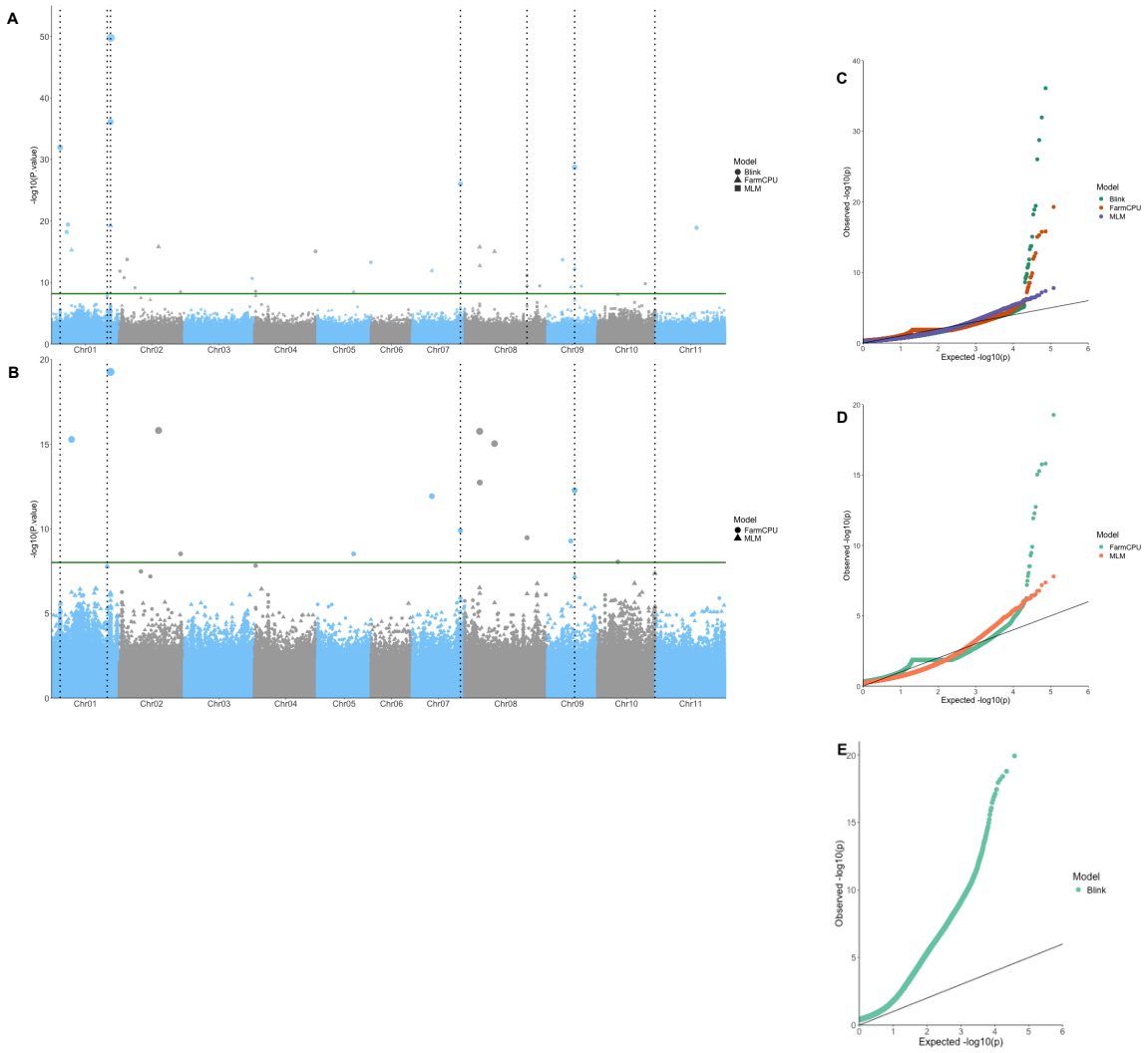


Figure 17; Manhattan plots highlighting markers significantly associated with determinacy on (A) the whole panel (n=144) and (B) the Andean subpanel (n=108). The analyses were completed with GAPIT and the models FarmCPU, BLINK or MLM (Huang et al. 2019; Liu et al. 2016; Wang and Zhang 2021; Zhang et al. 2010). The X-axis represents the genomic position of markers, and the Y-axis is the -log 10 of the P-values for association with the phenotype. The vertical lines correspond to QTLs found by at least two models. Point size correlates to -log10(P-value). Quantile-quantile (QQ) plots are provided for (C) the whole panel, (D) the Andean panel and (E) models not selected for the Andean panel. Plots A-D from Denning-James et al. (2025).

2.4.7 GWAS for Photoperiod Sensitivity

The GWAS was performed using the BLINK and FarmCPU models with GAPIT for the PS phenotype (Figure 18A and B). The QQ plots (Figure 18C and D) provided evidence that the selected models are fitted to identify significant marker trait associations (MTAs) for the dataset. Figure 18E and Figure **18**F showed the models and phenotypes not selected for further analysis based on the fit as evidenced by the QQ plots. Ten QTLs were identified. Six QTLs were the focus of the analysis, for the whole panel based on criteria laid out in the methods. The MTAs were found on chromosomes Pv04, Pv05, Pv07, Pv08 and Pv09 (vertical lines in Figure 18). Six QTLs were identified for the Andean subset panel in Chromosomes Pv05, Pv07, Pv08, Pv09 and Pv11. The QTL in Pv04 and Pv09 were found in the full dataset only. The QTL in Pv9 and Pv11 were found in the Andean subset only. Candidate genes were identified for the significant MTAs and their corresponding QTLs. The identified genes and QTLs are listed in Appendix 3.

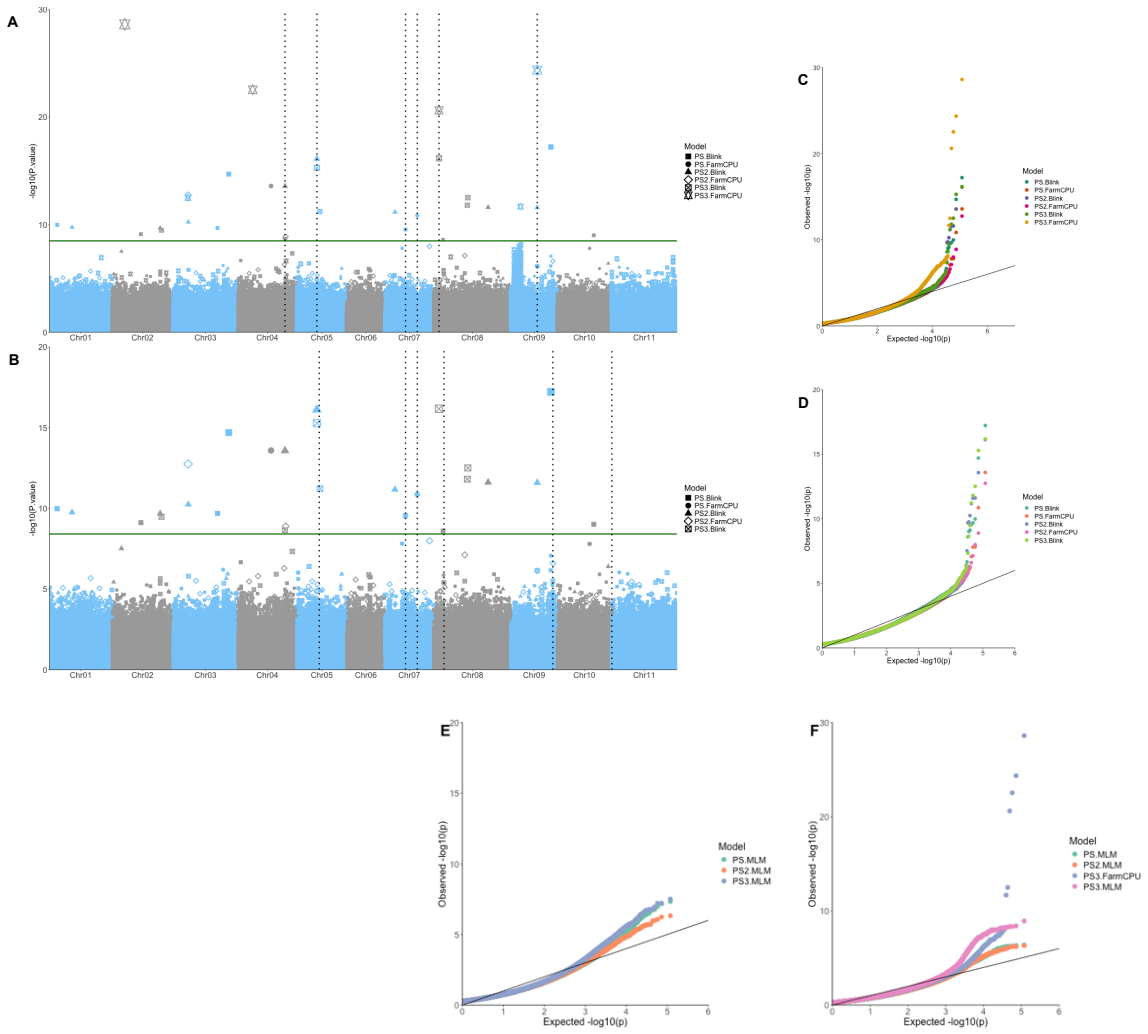


Figure 18; Manhattan plots highlighting markers significantly associated with photoperiod insensitivity on (A) the whole panel (n=144) and (B) the Andean subpanel (n=108). The analyses were completed with GAPIT and the models FarmCPU, BLINK or MLM (Huang et al. 2019; Liu et al. 2016; Wang and Zhang 2021; Zhang et al. 2010). The X-axis represents the genomic position of markers, and the Y-axis is the -log 10 of the P-values for association with the phenotype. The vertical lines correspond to QTLs found by at least two models. Point size correlates to -log10(P-value). Quantile-quantile (QQ) plots are provided for (C) the whole panel, (D) the Andean panel, (E) models not selected for the whole panel and (F) models not selected for the Andean panel. Plots A-D from Denning-James et al. (2025).

2.4.8 Mean allele dosage from QTLs for determinacy

The mean SNP dosage for polymorphisms in CDS within QTLs (SNP mean) was calculated for groups of accessions. These were grouped based on determinacy and population structure at K=2, the Mesoamerican (M) and Andean (A) gene pools, and their admixed (Admx). These allowed for the visual identification of differential allelic haplotypes among these groups (such as 'Determinate M' vs 'Indeterminate M').

Homozygous reference

The 'determinate A' group had the highest proportion of the homozygous reference allele and haplotype (HOMR) in multiple genes, such as Phvul.001G057600, Phvul.001G057500, Phvul.001G056500, Phvul.001G057700 (D1.1, Figure 19A), Phvul.007G245800 (D7.1, Figure 19I), Phvul.008G077000, Phvul.008G077100, Phvul.008G076500, Phvul.008G077500, Phvul.008G077600 (D8.1, Figure 19J), Phvul.008G170100 (D8.2, Figure 19K), Phvul.009G138100, Phvul.009G138300 (D9.1, Figure 19L) and Phvul.010G158300 (D10.1, Figure 19M). The 'indeterminate A' accessions had a higher proportion of HOMR in other genes, such as Phvul.002G116100 (D2.1, Figure 19G) and Phvul.003G099100 (D3.1, Figure 19H).

The 'indeterminate M' group had a higher proportion of HOMR in some genes compared to the other groups, such as Phvul.001G057000 (D1.1, Figure 19A). While in the genes Phvul.007G244900, Phvul.007G244700 (D7.1, Figure 19I) and Phvul.009G138300 (D9.1, Figure 19L), 'indeterminate M' have a higher proportion of HOMR compared to 'determinate M'.

In some situations, both indeterminate groups had a higher proportion of HOMR compared to the corresponding determinate, such as in Phvul.001G168400 (D1.3, Figure 19C), Phvul.001G189200 (D1.4, D1.5, Figure 19D) (indeterminate A is 100% HOMR) and Phvul.007G246200 (D7.1, Figure 19I). On the other hand, both determinate groups have a higher proportion of HOMR compared to the corresponding indeterminate groups in the gene Phvul.010G158400 (D10.1, Figure 19M).

Homozygous alternative

In the genes Phvul.001G057600 and Phvul.001G057500 (D1.1, Figure 19A), the ‘indeterminate A’ had a higher proportion of homozygous alternative (HOMA) alleles compared to the ‘determinate A’. However, the ‘determinate A’ group have a higher proportion of HOMA in the gene Phvul.010G158500, compared to the ‘indeterminate A’ (D10.1, Figure 19M). On the other hand, the ‘determinate M’ have a higher proportion of HOMA alleles in Phvul.001G168100 (D1.3, Figure 19C), Phvul.001G190500 (D1.5, Figure 19E) and Phvul.001G191500 (D1.5, Figure 19F).

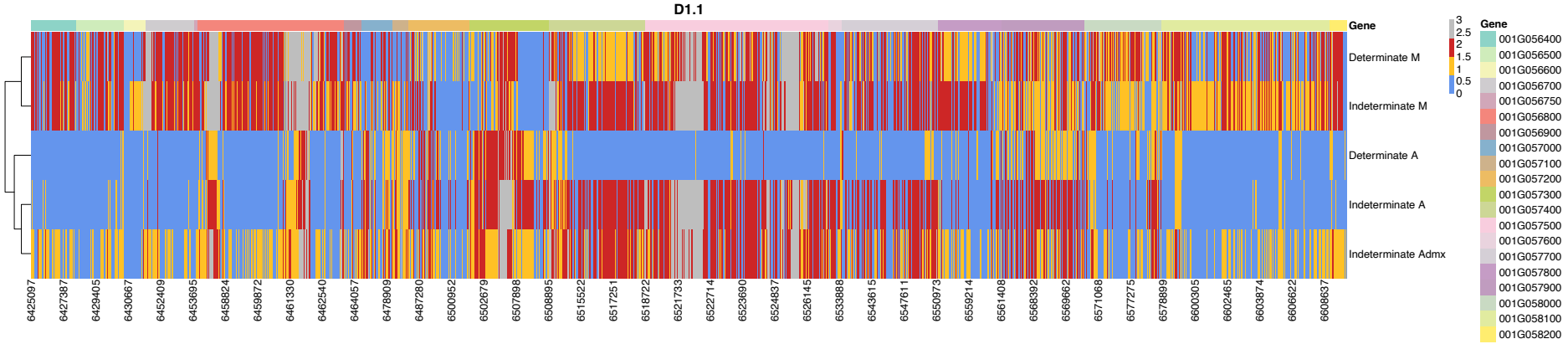
Uncalled diversity

Other genes had differences in the proportion of uncalled loci. For example, in Phvul.001G057500 (D1.1, Figure 19A) and Phvul.002G116550 (D2.1, Figure 19G) the ‘determinate A’ group had a lower proportion of uncalled SNP loci when compared to the ‘determinate M’, ‘indeterminate M’ and ‘indeterminate A’. This supports PAV, or large allelic variation in this gene within those groups. In Phvul.001G057800 ‘determinate M’ is the only group with uncalled loci, suggesting the absence of this gene or large allelic differences in this gene with the reference genome (Andean background) (D1.1, Figure 19A). Similarly, the ‘determinate M’ group had higher allelic variation (or missing gene) in Phvul.001G192200 (D1.6, Figure 19F) and Phvul.007G245800 (D7.1, Figure 19I). In Phvul.001G077200 (D1.2, Figure 19B) ‘indeterminate A’ had more uncalled positions than the other groups. Finally, the ‘indeterminate M' group has a large proportion of uncalled regions in Phvul.001G168200 (D1.3, Figure 19C).

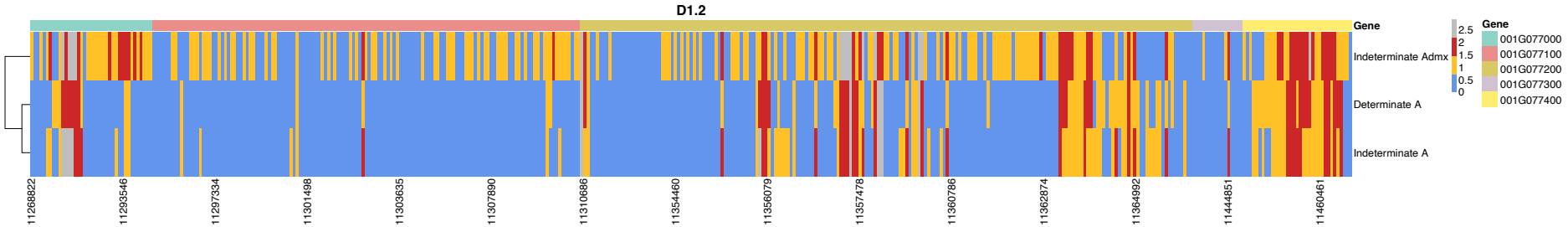
No clear differences

Certain genes such as Phvul.001G168300, Phvul.001G167900 (D1.3, Figure 19C), Phvul.002G115800 (D2.1, Figure 19G), Phvul.003G099300 (D3.1, Figure 19H), Phvul.008G170000 (D8.2, Figure 19K) have no clear differences among subgroups. While in the gene Phvul.002G116500 (D2.1, Figure 19G) all the diversity is uncalled in the Andean gene pools, suggesting it may be a private Mesoamerican gene or allele in the diversity panel.

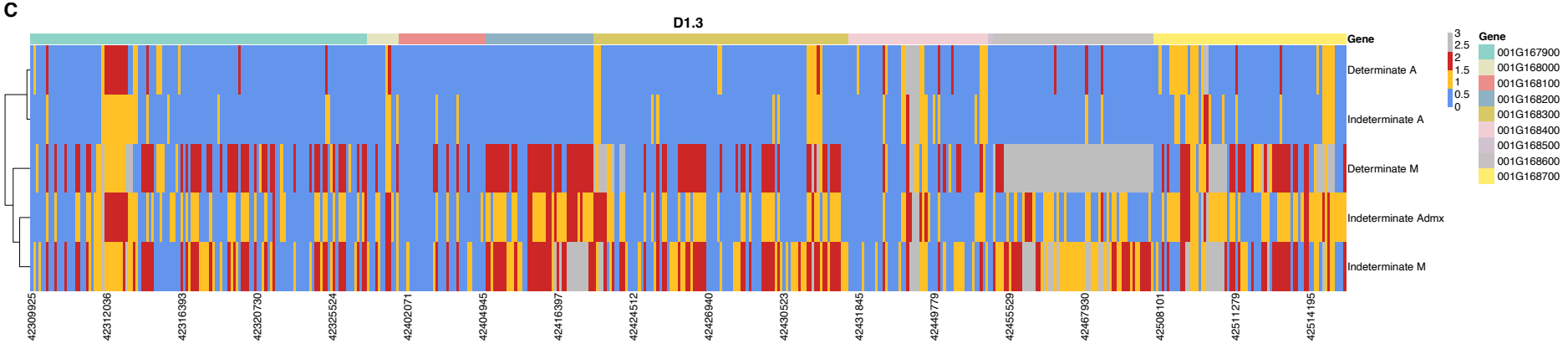
A


D1.1
Gene
Determinate M
Indeterminate M
Determinate A
Indeterminate A
Indeterminate Admx
3
2.5
2
1.5
1
0.5
0
Gene
001G056400
001G056500
001G056600
001G056700
001G056750
001G056800
001G056900
001G057000
001G057100
001G057200
001G057300
001G057400
001G057500
001G057600
001G057700
001G057800
001G057900
001G058000
001G058100
001G058200
6425097
6427387
6429405
6430667
6452409
6453695
6458824
6459872
6461330
6462540
6464057
6478909
6487280
6500952
6502679
6507898
6508895
6515522
6517251
6518722
6521733
6522714
6523690
6524837
6526145
6533888
6543615
6547611
6550973
6559214
6561408
6568392
6569962
6571068
6577275
6578899
6600305
6602465
6603874
6606622
6608637

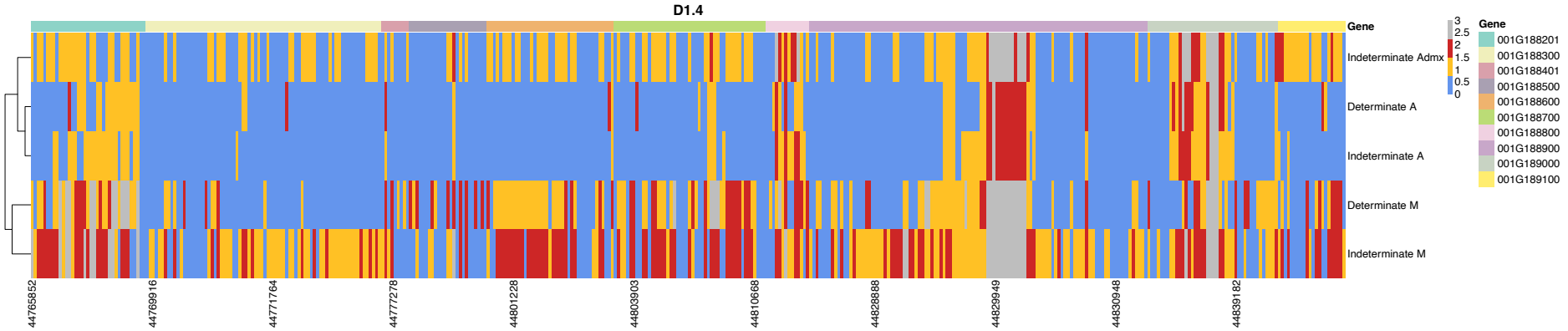
B


D1.2
Gene
Indeterminate Admx
Determinate A
Indeterminate A
2.5
2
1.5
1
0.5
0
Gene
001G077000
001G077100
001G077200
001G077300
001G077400
11268822
11293546
11297334
11301498
11303635
11307890
11310686
11354460
11356079
11357478
11360786
11362874
11364992
11444851
11460461

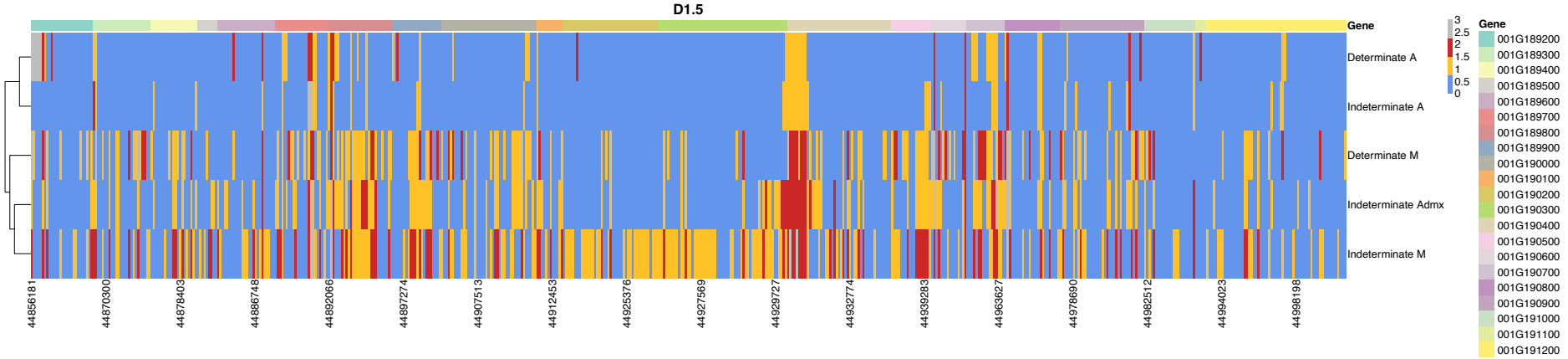
C


D1.3
Gene
Determinate A
Indeterminate A
Determinate M
Indeterminate Admx
Indeterminate M
42309925
42312036
42316393
42320730
42325524
42402071
42404945
42416397
42424512
42426940
42430523
42431845
42449779
42455529
42467930
42508101
42511279
42514195
Gene
001G167900
001G168000
001G168100
001G168200
001G168300
001G168400
001G168500
001G168600
001G168700

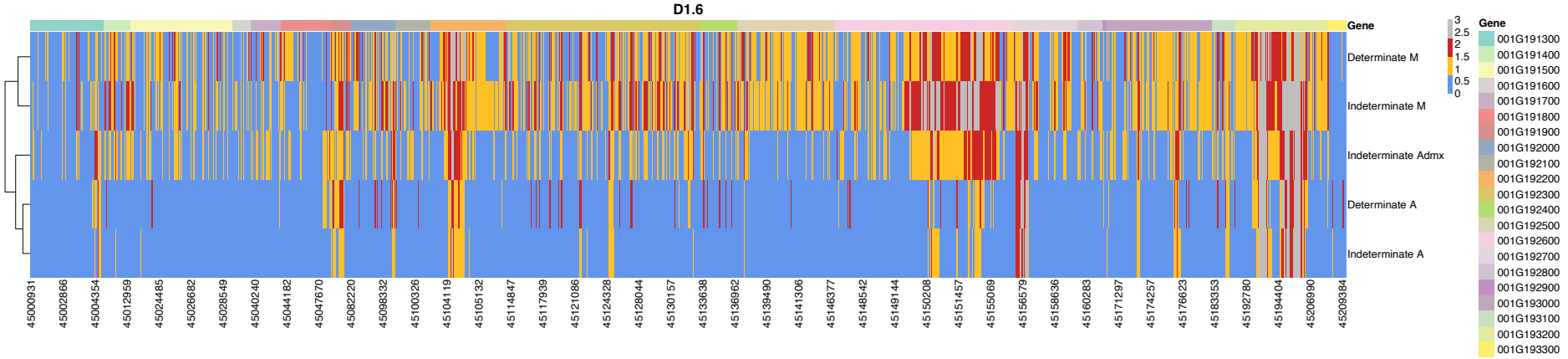
D


D1.4
Gene
Indeterminate Admx
Determinate A
Indeterminate A
Determinate M
Indeterminate M
44765852
44769916
44771764
44777278
44801228
44803903
44810668
44828888
44829949
44830948
44839182
Gene
001G188201
001G188300
001G188401
001G188500
001G188600
001G188700
001G188800
001G188900
001G189000
001G189100

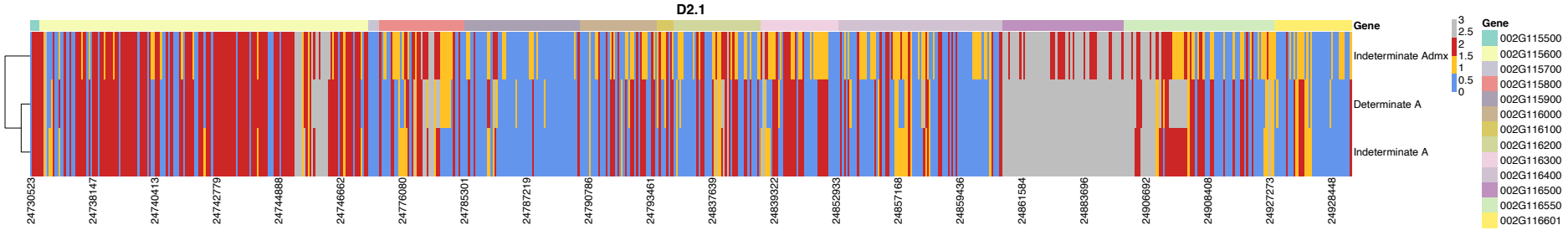
E


D1.5
Gene
Determinate A
Indeterminate A
Determinate M
Indeterminate Admx
Indeterminate M
44856181
44870300
44878403
44886748
44892066
44897274
44907513
44912453
44925376
44927569
44929727
44932774
44939283
44963627
44978690
44982512
44994023
44998198
3
2.5
2
1.5
1
0.5
0
Gene
001G189200
001G189300
001G189400
001G189500
001G189600
001G189700
001G189800
001G189900
001G190000
001G190100
001G190200
001G190300
001G190400
001G190500
001G190600
001G190700
001G190800
001G190900
001G191000
001G191100
001G191200

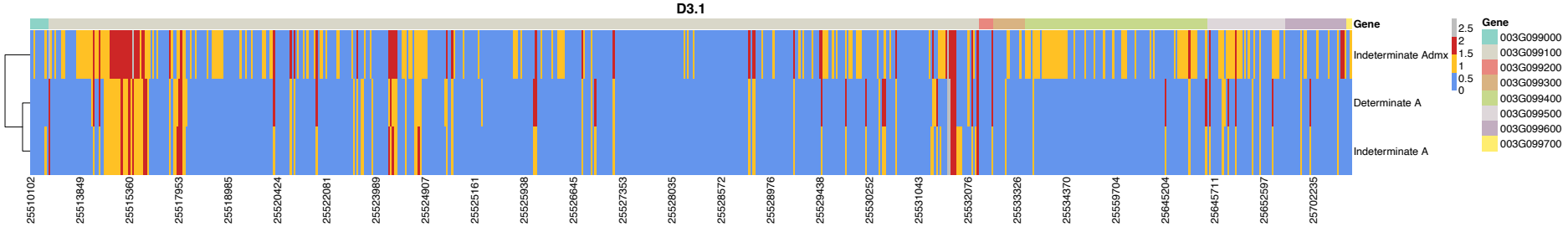
F


D1.6
Gene
Determinate M
Indeterminate M
Indeterminate Admx
Determinate A
Indeterminate A
45000931
45002866
45004354
45012959
45024485
45026682
45028549
45040240
45044182
45047670
45082220
45098332
45100326
45104119
45105132
45114847
45117939
45121086
45124328
45128044
45130157
45133638
45136962
45139490
45141306
45146377
45148542
45149144
45150208
45151457
45155069
45156579
45158636
45160283
45171297
45174257
45176623
45183353
45192780
45194404
45206990
45209384
3
2.5
2
1.5
1
0.5
0
Gene
001G191300
001G191400
001G191500
001G191600
001G191700
001G191800
001G191900
001G192000
001G192100
001G192200
001G192300
001G192400
001G192500
001G192600
001G192700
001G192800
001G192900
001G193000
001G193100
001G193200
001G193300

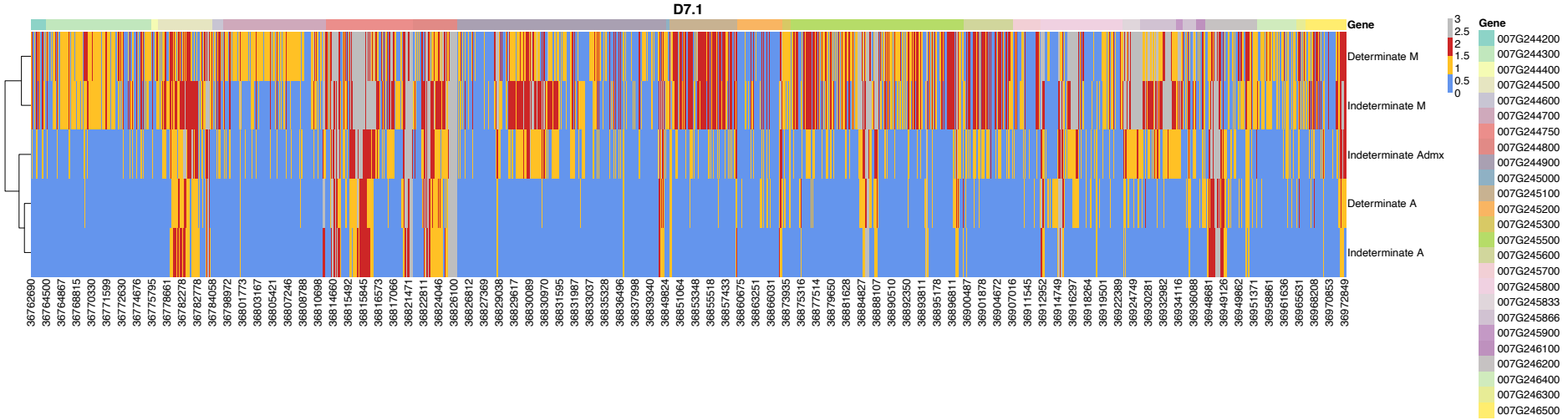
G


D2.1
Gene
Indeterminate Admx
Determinate A
Indeterminate A
3
2.5
2
1.5
1
0.5
0
Gene
002G115500
002G115600
002G115700
002G115800
002G115900
002G116000
002G116100
002G116200
002G116300
002G116400
002G116500
002G116550
002G116601
24730523
24738147
24740413
24742779
24744888
24746662
24776080
24785301
24787219
24790786
24793461
24837639
24839322
24852933
24857168
24859436
24861584
24883696
24906692
24908408
24927273
24928448

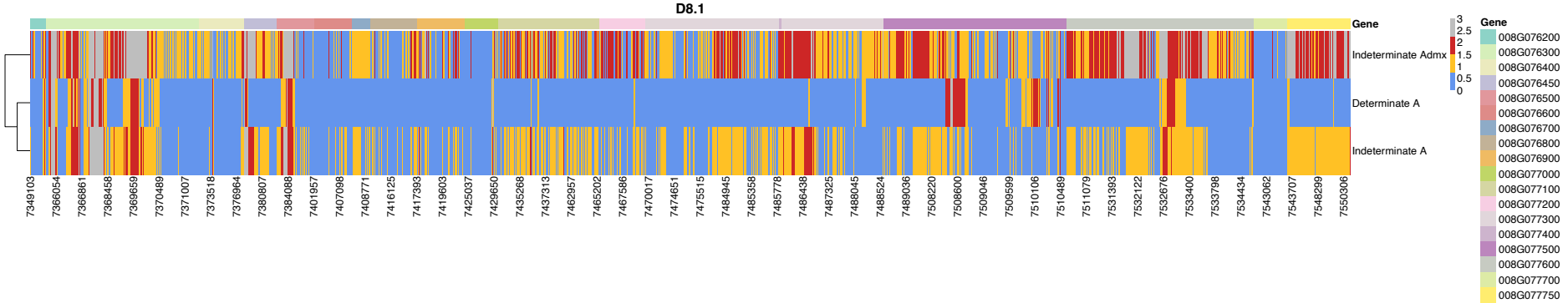
H


D3.1
Gene
Indeterminate Admx
Determinate A
Indeterminate A
2.5
2
1.5
1
0.5
0
Gene
003G099000
003G099100
003G099200
003G099300
003G099400
003G099500
003G099600
003G099700
25510102
25513849
25515360
25517953
25518985
25520424
25522081
25523989
25524907
25525161
25525938
25526645
25527353
25528035
25528572
25528976
25529438
25530252
25531043
25532076
25533326
25534370
25559704
25645204
25645711
25652597
25702235

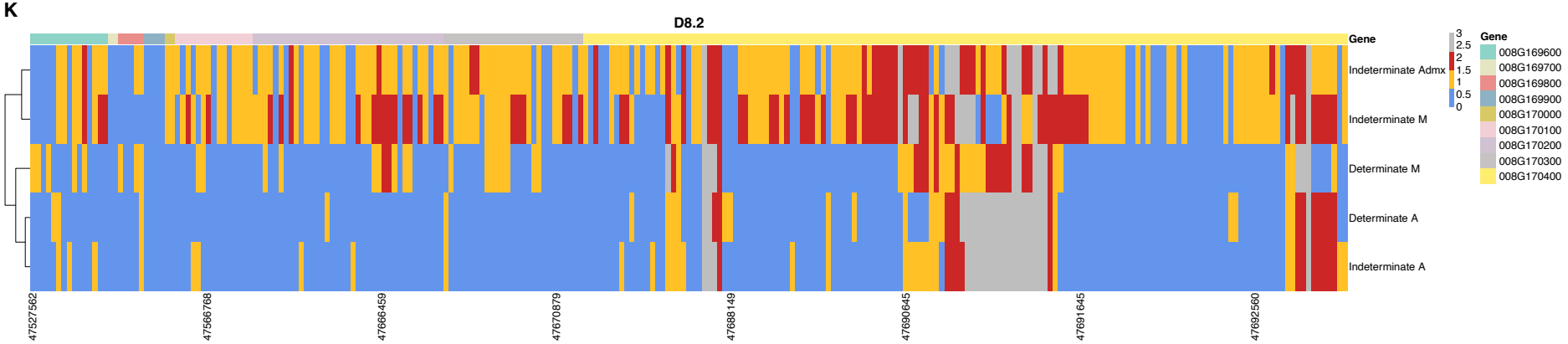
I


D7.1
Gene
Determinate M
Indeterminate M
Indeterminate Admx
Determinate A
Indeterminate A
3
2.5
2
1.5
1
0.5
0
Gene
007G244200
007G244300
007G244400
007G244500
007G244600
007G244700
007G244750
007G244800
007G244900
007G245000
007G245100
007G245200
007G245300
007G245500
007G245600
007G245700
007G245800
007G245833
007G245866
007G245900
007G246100
007G246200
007G246400
007G246300
007G246500
36762690
36764500
36764867
36768815
36770330
36771599
36772630
36774676
36775795
36778661
36782278
36782778
36784058
36798972
36801773
36803167
36805421
36807246
36808788
36810698
36814660
36815492
36815845
36816573
36817066
36821471
36822811
36824046
36826100
36826812
36827369
36829038
36829617
36830089
36830970
36831595
36831987
36833037
36835328
36836496
36837998
36839340
36849824
36851064
36853348
36855518
36857433
36860675
36863251
36866031
36873935
36875316
36877514
36879650
36881628
36884827
36888107
36890510
36892350
36893811
36895178
36896811
36900487
36901878
36904672
36907016
36911545
36912952
36914749
36916297
36918264
36919501
36922389
36924749
36930281
36932982
36934116
36936088
36948681
36949126
36949862
36951371
36958861
36961636
36965631
36968208
36970853
36972849

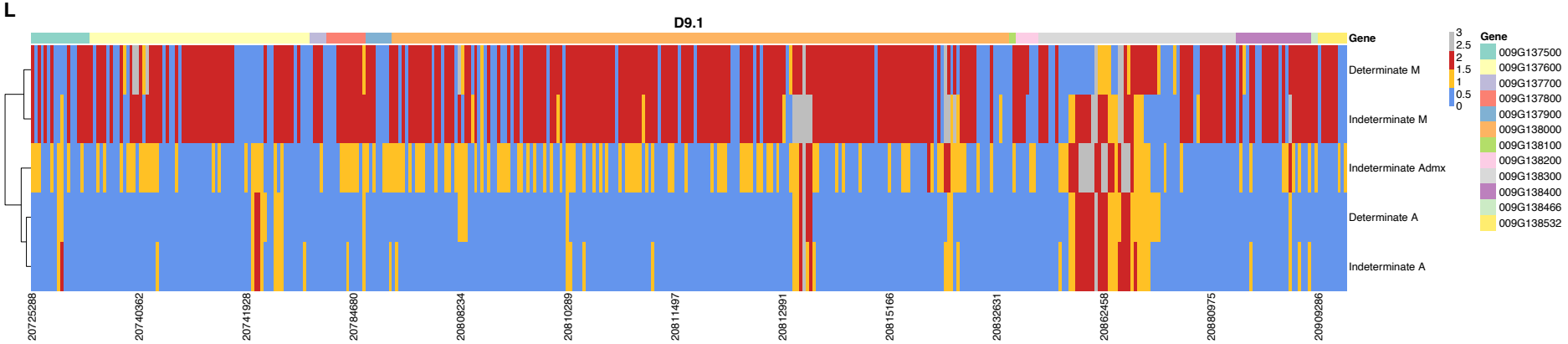
J


D8.1
Gene
Indeterminate Admx
Determinate A
Indeterminate A
3
2.5
2
1.5
1
0.5
0
Gene
008G076200
008G076300
008G076400
008G076450
008G076500
008G076600
008G076700
008G076800
008G076900
008G077000
008G077100
008G077200
008G077300
008G077400
008G077500
008G077600
008G077700
008G077750
7349103
7366054
7366861
7368458
7369659
7370489
7371007
7373518
7376964
7380807
7384088
7401957
7407098
7408771
7416125
7417393
7419603
7425037
7429650
7435266
7437313
7462957
7465202
7467586
7470017
7474651
7475515
7484945
7485358
7485778
7486438
7487325
7488045
7488524
7489036
7508220
7508600
7509046
7509599
7510106
7510489
7511079
7531393
7532122
7532676
7533400
7533798
7534434
7543062
7543707
7548299
7550306

K


D8.2
Gene
Indeterminate Admx
Indeterminate M
Determinate M
Determinate A
Indeterminate A
3
2.5
2
1.5
1
0.5
0
Gene
008G169600
008G169700
008G169800
008G169900
008G170000
008G170100
008G170200
008G170300
008G170400
47527562
47566768
47666459
47670879
47688149
47690645
47691645
47692560

L


D9.1
Gene
Determinate M
Indeterminate M
Indeterminate Admx
Determinate A
Indeterminate A
3
2.5
2
1.5
1
0.5
0
Gene
009G137500
009G137600
009G137700
009G137800
009G137900
009G138000
009G138100
009G138200
009G138300
009G138400
009G138466
009G138532
20725288
20740362
20741928
20784680
20808234
20810289
20811497
20812991
20815166
20832631
20862458
20880975
20909286

M

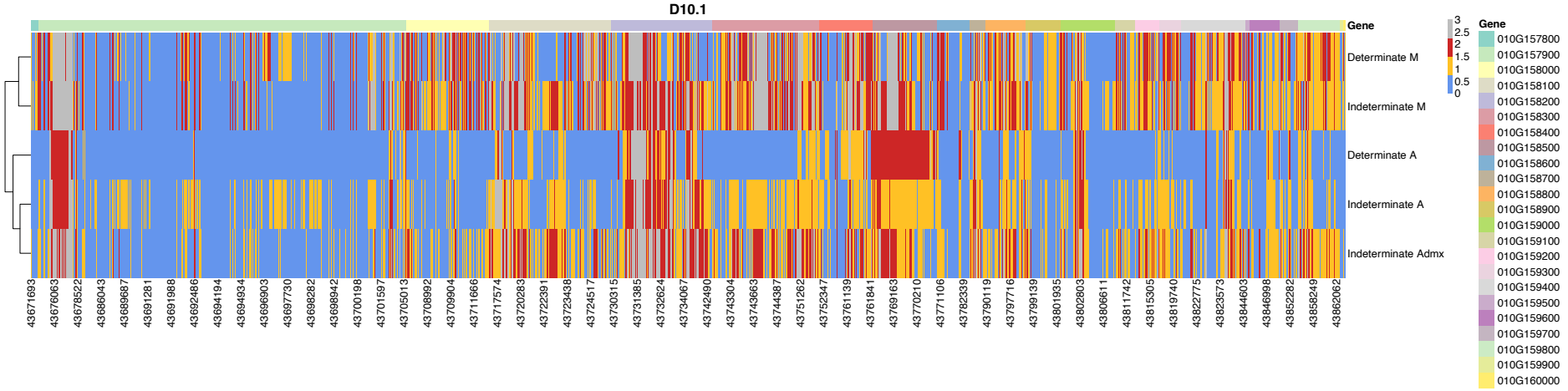


Figure 19; Heatmap of SNP mean dosage values for polymorphisms in the coding region of QTLs associated with determinacy from the GWAS of the whole (n=144) or Andean diversity panels. 0 are homozygous reference, 1 are heterozygous, 2 are homozygous alternative and 3 are uncalled genotypes (transformed NAs). Plotted with pheatmap v.1.0.13 with default Euclidean distance to cluster the rows and SNP location to order columns (Kolde 2025). The top row represents the gene which the SNPs fall within. SNPs have been averaged by their determinate trait and population structure at K2 (M; Mesoamerican, A; Andean, Admx; Admixture). SNPs with less than a 0.5 difference in variation between groups have been removed. The plots represent the whole diversity panel for QTLs (A) D1.1, (C) D1.3, (D) D1.4, (E) D1.5, (F) D1.6, (I) D7.1, (K) D8.2, (L) D9.1 and (M) D10.1. Also, the Andean diversity panel for QTLs (B) D1.2, (G) D2.1, (H) D3.1 and (J) D8.1.

2.4.9 Mean allele dosage from QTLs for photoperiod sensitivity

As in section 2.4.8 the mean SNP dosage for polymorphisms in CDS within QTLs (SNP mean) was calculated for groups of accessions. These were grouped based on photoperiod sensitivity (PS) and population structure at K=2, the Mesoamerican (M) and Andean (A) gene pools, and their admixed (Admx). These allowed the visual identification of differential allelic haplotypes among these groups (such as 'sensitive M' vs 'insensitive M').

Homozygous reference

The 'insensitive M' had a higher proportion of the HOMR than the 'sensitive M' in multiple genes, such as, Phvul.004G110200 (PS4.1, Figure 20A), Phvul.007G117500 (PS7.1, Figure 20D), Phvul.008G048300 and Phvul.008G047700 (PS8.1, Figure 20F). Also, in the Andean subpopulation the 'insensitive A' group had a higher proportion of HOMR than the 'sensitive A' in the genes Phvul.008G085000, Phvul.008G084900 (PS8.2, Figure 20G), Phvul.009G228900 (PS9.2, Figure 20I), Phvul.009G144600 and Phvul.009G144100 (PS9.1, Figure 20H).

All the insensitive groups (Admx, A and M) had a higher HOMR in multiple genes such as Phvul.004G110000 (PS4.1, Figure 20A), Phvul.005G076400 (PS5.1, Figure 20B) and Phvul.005G076800 (PS5.2, Figure 20C). However, the 'sensitive A' group had a higher proportion of HOMR in some genes compared to the 'insensitive A', such as Phvul.005G077000 (PS5.2, Figure 20C).

Homozygous alternative

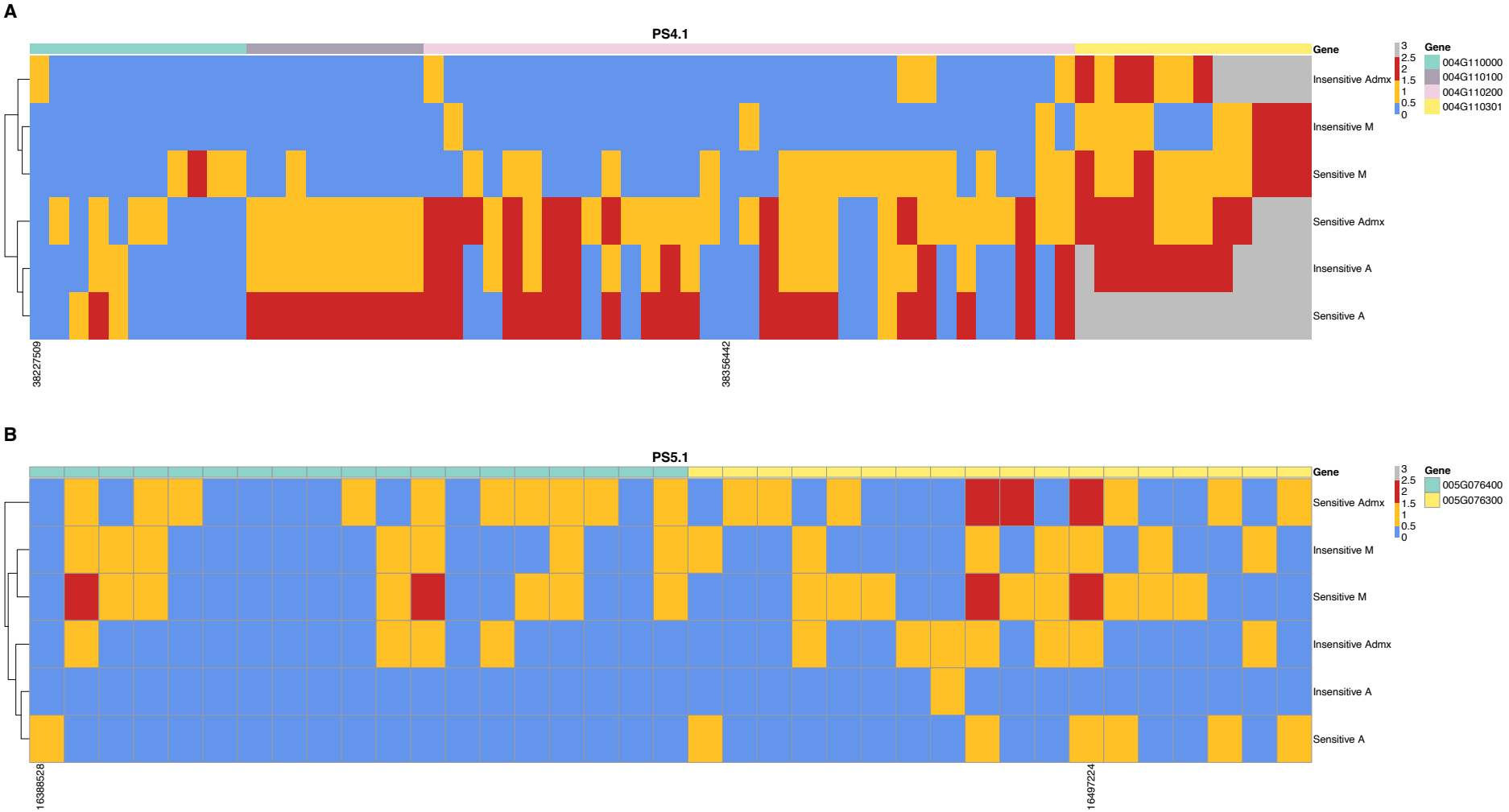
In the gene Phvul.005G076400 (PS5.2, Figure 20B), the HOMA alleles were only present in the 'sensitive M' group. On the other hand, the 'insensitive M' had a higher proportion of HOMA alleles compared to 'sensitive M' in the gene Phvul.007G157300 (PS7.2, Figure 20E). Other genes such as Phvul.011G002900 had a higher proportion of HOMA alleles in the 'insensitive A' and 'insensitive Admx' groups than the corresponding sensitive groups (PS11.1, Figure 20J). Finally, genes such as Phvul.005G077000 (PS5.2, Figure 20C) have no HOMA alleles in any groups.

Uncalled diversity

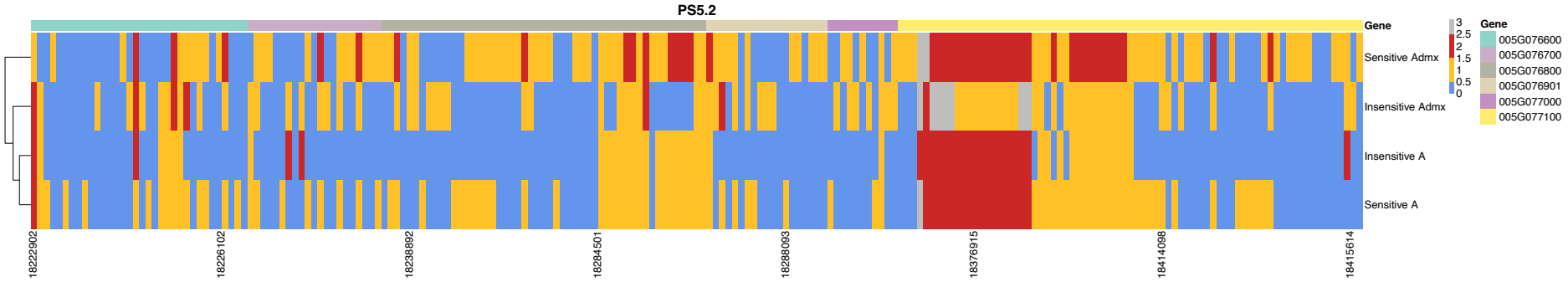
In the gene Phvul.005G077100 the 'insensitive A' is the only group to contain no uncalled loci, suggesting only this group contains the allele present in the reference genome (PS5.2, Figure 20). While in the gene Phvul.007G157300 neither Andean group had uncalled positions, suggesting it is a private Andean allele. As the Mesoamerican and Andean have different enough alleles to not align to each other (PS7.2, Figure 20E). The genes Phvul.009G144600, Phvul.009G144100 and Phvul.009G144700 only had uncalled regions in the 'sensitive M' group, which suggests this carries a significantly different allele for this gene, or this gene is absent (PS9.1, Figure 20H).

No clear differences

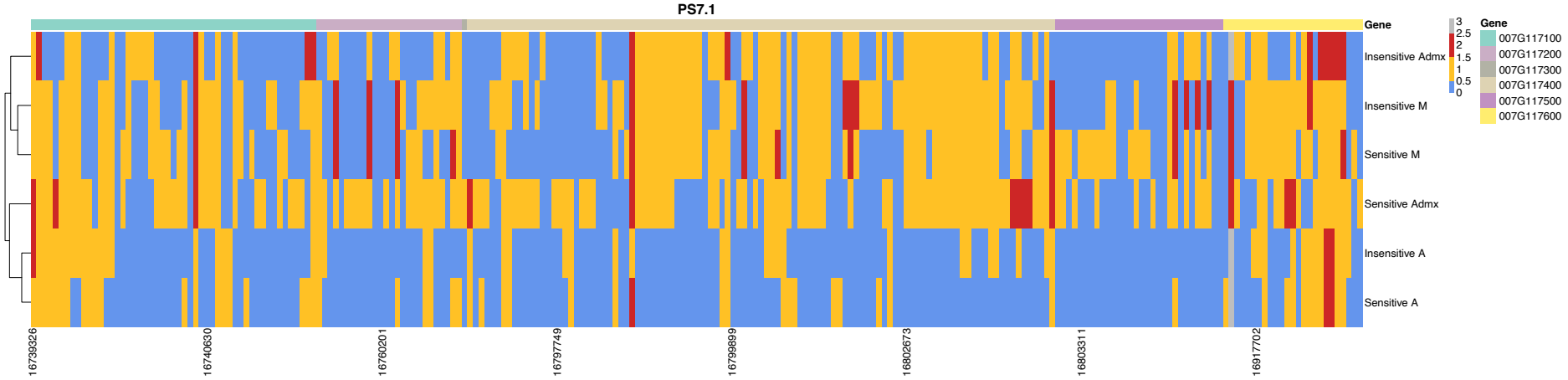
There are large proportions of uncalled loci in all the groups except the Mesoamerican gene pool in the gene Phvul.004G110301, suggesting it is private to this group of the diversity panel but not private to the Mesoamerican gene pool as the reference genome is Andean (PS4.1, Figure 20A). However, there are no differences between the sensitive or insensitive accessions in this gene. On the other hand, in the genes Phvul.007G117400, Phvul.007G117100 (PS7.1, Figure 20D), Phvul.007G156900 (PS7.2, Figure 20E), Phvul.008G084200 (PS8.2, Figure 20G), Phvul.009G144200 and Phvul.009G143700 (PS9.1, Figure 20H) there are no clear differences in proportions between the 'sensitive' and 'insensitive' types in their corresponding gene pools.


A
PS4.1
Gene
Insensitive Admx
Insensitive M
Sensitive M
Sensitive Admx
Insensitive A
Sensitive A
38227509
38356442
Gene
004G110000
004G110100
004G110200
004G110301
B
PS5.1
Gene
Sensitive Admx
Insensitive M
Sensitive M
Insensitive Admx
Insensitive A
Sensitive A
16388528
16497224
Gene
005G076400
005G076300

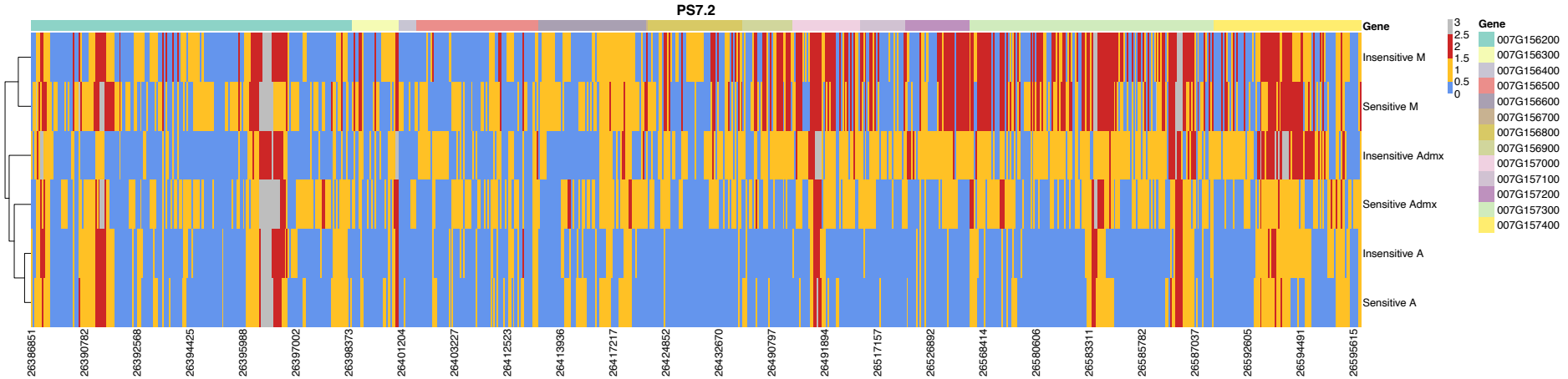
C


PS5.2
Gene
Sensitive Admx
Insensitive Admx
Insensitive A
Sensitive A
18222902
18226102
18238892
18284501
18288093
18376915
18414098
18415614
3
2.5
2
1.5
1
0.5
0
Gene
005G076600
005G076700
005G076800
005G076901
005G077000
005G077100

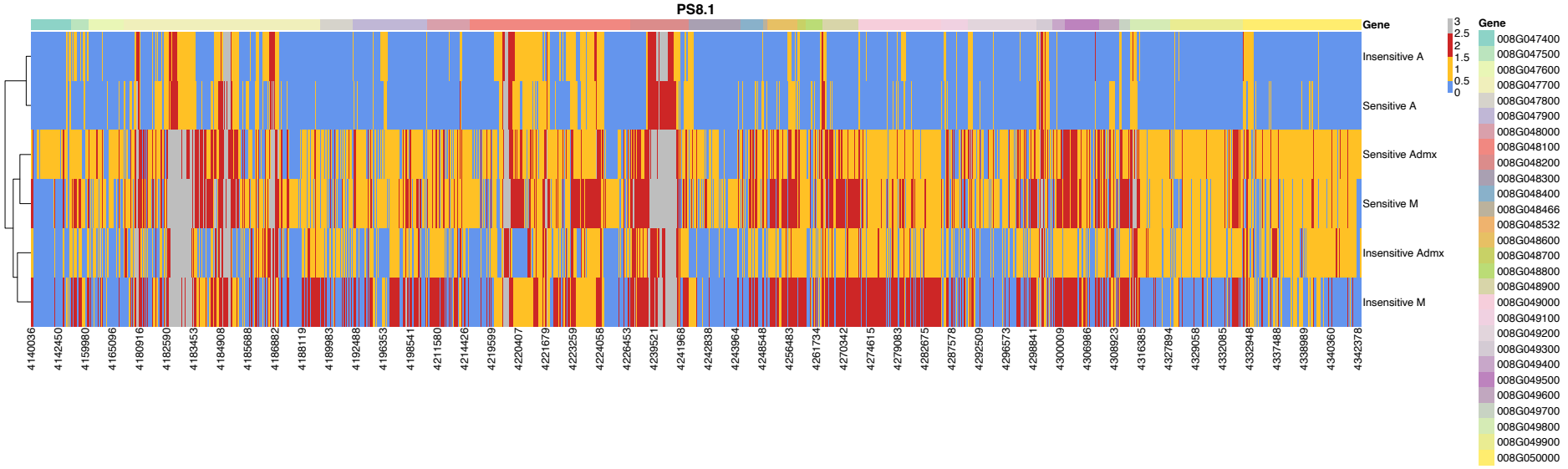
D


PS7.1
Gene
Insensitive Admx
Insensitive M
Sensitive M
Sensitive Admx
Insensitive A
Sensitive A
16739326
16740630
16760201
16797749
16799899
16802673
16803311
16917702
3
2.5
2
1.5
1
0.5
0
Gene
007G117100
007G117200
007G117300
007G117400
007G117500
007G117600

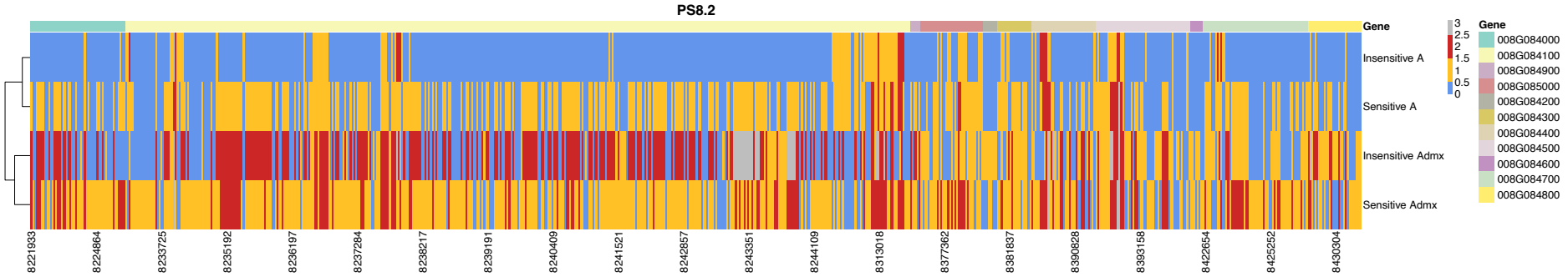
E


PS7.2
Gene
Insensitive M
Sensitive M
Insensitive Admx
Sensitive Admx
Insensitive A
Sensitive A
3
2.5
2
1.5
1
0.5
0
Gene
007G156200
007G156300
007G156400
007G156500
007G156600
007G156700
007G156800
007G156900
007G157000
007G157100
007G157200
007G157300
007G157400
26386851
26390782
26392568
26394425
26395988
26397002
26398373
26401204
26403227
26412523
26413936
26417217
26424852
26432670
26490797
26491894
26517157
26526892
26568414
26580606
26583311
26585782
26587037
26592605
26594491
26595615

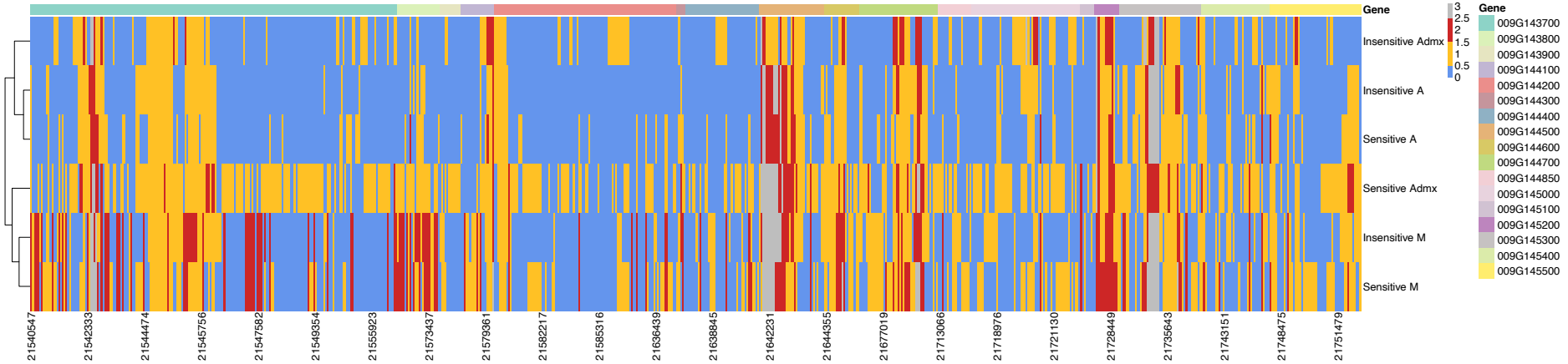
F


PS8.1
Gene
Insensitive A
Sensitive A
Sensitive Admx
Sensitive M
Insensitive Admx
Insensitive M
3
2.5
2
1.5
1
0.5
0
Gene
008G047400
008G047500
008G047600
008G047700
008G047800
008G047900
008G048000
008G048100
008G048200
008G048300
008G048400
008G048466
008G048532
008G048600
008G048700
008G048800
008G048900
008G049000
008G049100
008G049200
008G049300
008G049400
008G049500
008G049600
008G049700
008G049800
008G049900
008G050000
4140036
4142450
4159980
4165096
4180916
4182590
4183453
4184908
4185688
4186882
4188119
4189983
4192488
4196353
4198541
4211580
4214426
4219599
4220407
4221679
4223259
4224058
4226453
4239521
4241968
4242838
4243964
4248548
4256483
4261734
4270342
4274615
4279083
4282675
4287578
4292509
4296573
4298841
4300009
4306986
4308923
4316385
4327894
4329058
4332085
4332948
4337488
4338989
4340360
4342378

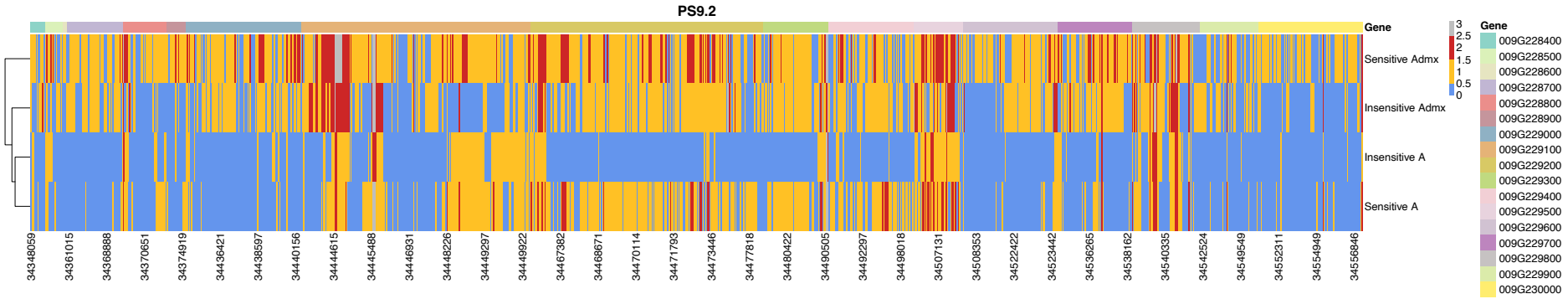
G


PS8.2
Gene
Insensitive A
Sensitive A
Insensitive Admx
Sensitive Admx
8221933
8224864
8233725
8235192
8236197
8237284
8238217
8239191
8240409
8241521
8242857
8243351
8244109
8313018
8377362
8381837
8390828
8393158
8422654
8425252
8430304
3
2.5
2
1.5
1
0.5
0
Gene
008G084000
008G084100
008G084900
008G085000
008G084200
008G084300
008G084400
008G084500
008G084600
008G084700
008G084800

H


PS9.1
Gene
Insensitive Admx
Insensitive A
Sensitive A
Sensitive Admx
Insensitive M
Sensitive M
21540547
21542333
21544474
21545756
21547582
21549354
21555923
21573437
21579361
21582217
21585316
21636439
21638845
21642231
21644355
21677019
21713066
21718976
21721130
21728449
21735643
21743151
21748475
21751479
3
2.5
2
1.5
1
0.5
0
Gene
009G143700
009G143800
009G143900
009G144100
009G144200
009G144300
009G144400
009G144500
009G144600
009G144700
009G144850
009G145000
009G145100
009G145200
009G145300
009G145400
009G145500

I


PS9.2
Gene
Sensitive Admx
Insensitive Admx
Insensitive A
Sensitive A
3
2.5
2
1.5
1
0.5
0
Gene
009G228400
009G228500
009G228600
009G228700
009G228800
009G228900
009G229000
009G229100
009G229200
009G229300
009G229400
009G229500
009G229600
009G229700
009G229800
009G229900
009G230000
34348059
34361015
34368888
34370651
34374919
34436421
34438597
34440156
34444615
34445488
34446931
34448226
34449297
34449922
34467382
34468671
34470114
34471793
34473446
34477818
34480422
34490505
34492297
34498018
34507131
34508353
34522422
34523442
34536265
34538162
34540335
34542524
34549549
34552311
34554949
34556846

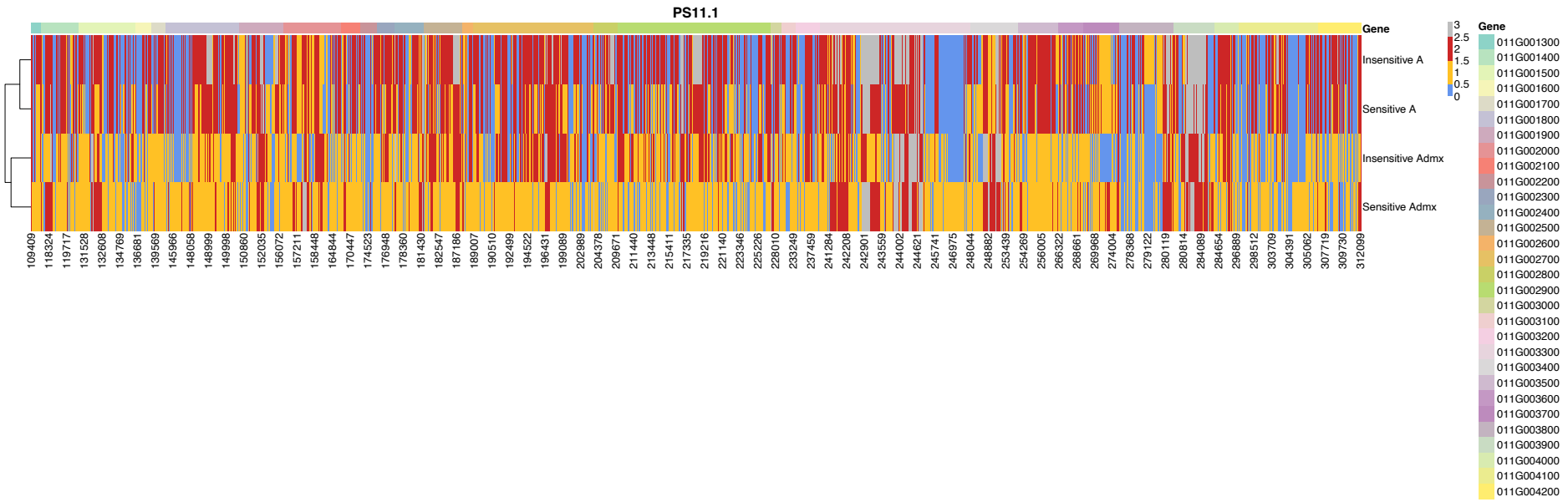


Figure 20; Heatmap of SNP mean dosage values for polymorphisms in the coding regions in QTLs associated with Photoperiod sensitivity from the GWAS of the whole (n=144) or Andean diversity panels. 0 are homozygous reference, 1 are heterozygous, 2 are homozygous alternative and 3 are uncalled genotypes (transformed NAs). Plotted with pheatmap v.1.0.13 with default Euclidean distance to cluster the rows and SNP location to order columns (Kolde 2025). The top row represents the gene which the SNPs fall within. SNPs have been averaged by their determinate trait and population structure at K2 (M; Mesoamerican, A; Andean, Admx; Admixture). SNPs with less than a 0.5 difference in variation between groups have been removed. The plots represent the whole diversity panel for QTLs (A) PS4.1, (B) PS5.1, (D) PS7.1, (E) PS7.2, (F) PS8.1 and (H) PS9.1. Also, the Andean diversity panel for (C) PS5.2, (G) PS8.2, (I) PS9.2 and (J) PS11.1.

2.5 Discussion

The common bean panel contains genetic diversity from the Andean (4 subgroups) and Mesoamerican (2 subgroups) gene pools. Including accessions from Colombia, that contain introgressive hybridisation and admixed diversity from the Andean and Mesoamerican gene pools. This work found an association between the population structure and agronomic traits, such as determinacy and photoperiod sensitivity.

In this study, genomic regions were identified that were connected to known and novel putative candidate genes, involved in developmental and reproductive pathways. Thirteen QTLs were associated with determinacy, and ten QTLs were associated with photoperiod sensitivity. One known QTL for determinacy was the *Fin* locus on chromosome 1, which is recognised for its pleiotropic effects in plant development. Other putative candidate genes were identified due to homology with *Glycine soja, Vigna species* and *Arabidopsis*. This includes *Phvul.008G170000* that encodes a putative FAF (fantastic four) domain-containing protein. The mean allele dosage could be used to infer allelic blocks within the population structure or phenotypes.

Consequently, GWAS are important in identifying MTAs and candidate genes, especially when accounting for population structure. By linking candidate genes to phenotypes, more targeted precision breeding approaches can be adopted to improve common bean traits under climate change. Nevertheless, this current study, and previous ones, highlight that for some genes and genomic regions, this will be difficult due to the high proportion of pleiotropic effects in common beans.

2.5.1 Environmental analysis

Plants have evolved to respond to light differences within the 24-hour day-night cycle. The length and timing of the light to dark cycle impacts plant development, such as flowering (Roeber et al. 2022). During the trials, flowering under long- and short-day length conditions was compared (Figure 16) to distinguish photoperiod sensitive and photoperiod insensitive phenotypes. Certain plants require long-day or short-day photoperiods to flower, or on the contrary are day-neutral (Figure 16A-C), these correspond to photoperiod sensitive (require a certain day length to flower) or photoperiod insensitive (evolved to flower independently from light conditions) (González-Delgado et al. 2025). Photoperiod insensitivity has helped in the dissemination of crops worldwide, as the plants can yield under various daylengths.

In Figure 16, the temperatures can be seen to go over 30 °C during certain periods in the long daylength trial. Temperatures greater than 30 °C during the day and 20 °C at night have been shown to impact the yields of common beans (Vargas et al. 2021). However, as seen in Figure 16 the temperatures did not exceed 30 °C during the day or 20 °C during the night, whilst the accessions had started flowering; consequently the temperature should not have impacted the timing of flowering. However, a limitation of this study is that humidity was not measured. Humidity can significantly influence transpiration and photosynthesis, therefore affecting plant growth and development. This should be considered in future studies (Lysenko et al. 2023).

2.5.2 Alignment and heterozygosity

During this study, analysis was completed with the Andean reference genome (Schmutz et al. 2014). This reference genome was selected for being the most complete at the time of analysis, and because the panel has a higher proportion of Andean accessions based on population structure analysis (Figure 8). The accessions also had higher alignments to the Andean reference genome (proportion paired > 87%, Figure 7) and no difference in metrics to the Mesoamerican reference genomes (DOE-JGI 2021) (proportion paired > 87%, Appendix 1). If there had been a large difference in the alignment metrics the analysis would have continued with both reference genomes.

The analysis of the heterozygosity (Figure 11A) revealed that the two weedy accessions in the diversity panel have higher levels of heterozygosity, suggesting outcrossing, as observed in other plant species (Hu et al. 2021b; Li et al. 2023b). The literature suggests that weedy accessions are hybridisations between crop wild relatives and landraces (Chacon-Sanchez et al. 2021). However, further research would be needed to draw conclusions on this, as the two weedy accessions in the diversity panel are not enough to understand diversity dynamics.

Also, the Mesoamerican accessions have a higher proportion of heterozygous sites than the Andean (<0.1) (Figure 11), this is supported by prior research (de Souza et al. 2023a). This suggests a founder effect during domestication or that Mesoamerican populations are undergoing more hybridisations (gene flow) with other gene pools. Finally, the six most heterozygous accessions (>0.3) are all from the region of Colombia called *Liborina*, suggesting a higher level of outcrossing in this region.

2.5.3 Population Structure and phenotyping

Subpopulations were delineated in the panel of 144 accessions, initially divided by domestication events, into the two Andean and the Mesoamerican gene pools (Figure 8, Figure 10) (Blair et al. 2013a; Kami et al. 1995). The Mesoamerican gene pool is generally more diverse (Mamidi et al. 2013; Schmutz et al. 2014) with less influence from domestication bottlenecks. Furthermore, the Mesoamerican gene pool, within the diversity panel, is more heterozygous, suggesting that the Andean gene pool has undergone fewer outcrossing events with other pools (gene flow) (Hoyos-Villegas et al. 2017). These crosses between gene pools occur during common bean dissemination, breeding programmes and selection based on market preferences by growers (Bellucci et al. 2023; Botero et al. 2021; de Almeida et al. 2020; Hoyos-Villegas et al. 2017). However, care needs to be taken when utilising market passport information. This is highlighted by the two 'Peruvian' accessions collected from markets that fall with the Mesoamerican subpopulation (Appendix 1).

Admixture was commonly observed in the panel, including 26 admixed Andean accessions, 5 admixed Mesoamerican accessions, and 11 Mesoamerican x Andean accessions. This supports the initial hypothesis that Colombia and neighbouring countries hold a large amount of common bean variation, including hybrids between both gene pools (Gori et al. 2022; Myers et al. 2000; Pironon et al. 2020). The wider crosses between gene pools compared to crosses within gene pools resulted in a larger observed heterozygosity in the hybrid accessions, supporting the outcrossing events and gene flow between gene pools. One implication of this study is that admixed Colombian hybrid landraces bridge Andean and Mesoamerican gene pools, and carry novel allelic and epistatic interactions that likely filtered out deleterious effects due to stronger purifying selection with increased recombination (Cichy et al. 2015). After all, recombination increases local effective population size (Ne) and limits Hill-Robertson interference (Hill and Robertson 2007). This suggests the Colombian hybrids have promising potential for breeding. However, the diversity panel may be biased and underestimating their prevalence in other regions due to the large number of Colombian accessions in the diversity panel.

Some traits were associated with demography, including determinacy and photoperiod sensitivity: C1 and C2 shared origin but could be separated by ancestry admixture analysis, and were characterised by different determinacy. C1 contained mainly determinate accessions, and C2 mainly indeterminate accessions. Furthermore, the population structure suggests that

Colombian farmers have not selected varieties based on the seed characteristics studied (e.g. seed size).

Indeterminate and photoperiod sensitive landraces were common, despite the combined selection for photoperiod insensitivity and determinacy resulting in common bean varieties with shorter flowering periods (DTF) and easier management. Prior research supports the correlation between DTF and phenotypes such as seed weight, determinacy and growth habit (Elias et al. 2021; Hoyos-Villegas et al. 2017; Moghaddam et al. 2016; Tar'an et al. 2002; Vargas et al. 2021). These phenotypes are related to apical meristems and floral development (Sablowski 2007).

The distribution of DTF values, in either summer or winter, were bimodal, i.e. had two peaks (Figure 15B and C). This likely occurred due to the determinate types flowering first, followed by the indeterminate beans (Coelho et al. 2023). The distribution also correlates to growth habits, as bush types typically flower earlier than climbing types (Ugwuanyi et al. 2022). Figure 10 supports that photoperiod sensitivity arose during domestication in both gene pools (Weller et al. 2019).

The Andean accessions within the diversity panel were large and medium seeded, while the Mesoamerican accessions were small and medium sized, agreeing with previous research (Blair et al. 2009). Among the Mesoamerican accessions, the Mesoamerican race is characterised by small-seeds, while the race Durango-Jalisco (DJ) is characterised by medium-seeds (Beebe et al. 2000; Blair et al. 2009; Giordani et al. 2022; Zhang et al. 2008). Separation of the diversity panel into subpopulation matching these races was not possible. This may be due to the limited number of Mesoamerican accessions in the panel, a limited genetic component underlying the seed size trait, or introgressions in the Mesoamerican Colombian accessions. Also, the major seed colour was correlated with other traits, such as the estimated seed weight. This suggests that seed colour may have been selected due to regional and consumer preferences alongside seed weight (Catarcione et al. 2023; Sadohara et al. 2025).

Interestingly, Ecuador accessions are often separated from Andean subgroups, suggesting that they are members of the PhI group or a possible sister species, Phaseolus debouckii (Chacon S et al. 2007; Rendon-Anaya et al. 2017a). Further to this, the wild Ecuador accession is separated from both gene pools (Figure 10 and Figure **11**), suggesting a separate ancestry originating from Ecuador or Peru (Bitocchi et al. 2012; Bitocchi et al. 2017). Finally, the C-EP group are mainly photoperiod sensitive (Figure 15F), possibly due to a different domestication

history or due to their Equatorial provenance, and not necessitating its evolution under fluctuating photoperiods.

2.5.4 Introgressions and population structure

The analyses shown in Figure 12 and Figure 13 provide evidence of introgressions between the Andean and Mesoamerican gene pools across the subpopulations. The introgression analysis identified structural differences between subpopulation genomes, helping to explain why the ADMIXTURE ancestry analysis separated them. The statistical tests support the visual patterns observed in the plots.

The adaptive introgressions provide a novel source of genetic variation, as the diversity from two backgrounds mixes and stabilises in a population, possibly with an impact on fitness (Suarez-Gonzalez et al. 2018). With the current level of analysis, introgressions could not be linked to the domestication phenotypes (photoperiod sensitivity or determinacy) but could be associated to the population structure. Further work could link the introgressions to other domestication phenotypes, such as seed size or responses to biotic stress (Bellucci et al. 2023; Wedger et al. 2024; Whitney et al. 2006).

In addition, VCFhunter identified regions of 'uncalled' (unknown) diversity within many accessions. These regions can arise for multiple reasons; therefore, caution should be taken when drawing conclusions without further validation. One explanation is that alleles may be private to a given accession or subgroup and are not represented in the 'ancestral' accessions selected from the Andean (A1) and Mesoamerican (M1 and M2) groups. For example, 'uncalled' alleles in the Colombian subgroups (C1, C2, and C-EP) may not be present in the 'ancestral' groups and could therefore represent private variation.

Regions may also be classified as 'uncalled' when the origin of the allelic block is ambiguous. The diversity panel was selected for its potential for admixed diversity between the Andean and Mesoamerican gene pools. Consequently, alleles may not fit the parameters of VCFhunter, as they did not clearly belong to one 'ancestral' origin, but instead a mixture of both Andean and Mesoamerican. As described in the Methods (section 2.3.5) a Colombian (C1 and C2) 'ancestral' group was tested alongside the Andean (A1) and Mesoamerican (M1 and M2). However, this resulted in large genomic regions being labelled 'uncalled', likely due to overlapping genetic diversity between the 'ancestral' groups. This meant few alleles were private to a group, so most alleles were labelled 'uncalled'.

Another reason is that the reference may be introducing bias. If the Andean reference genome does not contain sequences found in the accessions, or if certain regions mis-align or do not map correctly, then these regions will stay 'uncalled' throughout the VCFhunter pipeline. In such cases, the 'uncalled' diversity may not be private to the subpopulations but instead absent from the reference. These 'uncalled' regions should be explored further as they could contain novel structural variations which could impact phenotypes and plant fitness.

It may also be important to resolve the private diversity in the future to help prioritise genomic regions of interest in breeding programmes (Gasc et al. 2016; Meuwissen et al. 2020). However, there is also evidence that understanding the private diversity and precise impacts is unnecessary to leverage for genetic gain in breeding programmes therefore, selection should be phenotype based instead (Sanchez et al. 2023).

Higher levels of heterozygosity are observed in certain chromosomes in the subpopulations 'admixed AM', C2, M1 and 'admixed M', compared to other chromosomes in the corresponding subpopulation. This knowledge can help differentiate the subpopulations but also could reflect higher levels of cross over (hot spots), different selection pressures, or be due to the evolutionary history of the crop (Jaramillo-Correa et al. 2010; Li et al. 2023b). These regions may require further investigation to determine if they disproportionately impact diversity and adaptations.

Further to this, 'uncalled' regions may be due to genomic features including structural variations such as insertions, deletions, and copy number variation, as well as transposable elements. Transposable elements are frequently found in plant genomes from insertions and rearrangements, impacting alignment and gene function (Lozano-Arce et al. 2023). Divergence from the reference genome can happen due to admixture between the Andean and Mesoamerican gene pools or from diverged haplotypes, not well represented in the reference. Furthermore, there are limitations to the reference genomes currently available, including incomplete assemblies and missing sequences. Advances in long-read sequencing technologies and pangenome approaches are likely to improve genome resolution in future studies.

2.5.5 GWAS results

By leveraging this diversity panel and its trait segregation across the demographic stratification, thirteen QTLs for determinacy and ten QTLs for photoperiod sensitivity were prioritised (Figure 17 and Figure 18). Four of the QTLs for photoperiod sensitivity, and four for determinacy, were identified for the Andean subset, but not the whole panel. The Andean gene pool has adapted to lower latitudes than the Mesoamerican pool, resulting in differential selection for photoperiod sensitivity between the two gene pools. The linkage disequilibrium was estimated as 114Kb from an R^2 cut-off of 0.25, this value is consistent with WGS data of diversity panels rather than breeding populations (Ambachew et al. 2024; Campa et al. 2018; Diniz et al. 2019; Reinprecht et al. 2024). Linkage disequilibrium in common beans is impacted by the evolutionary and breeding history of the accessions in the diversity panel, therefore a 200kb region accounts for the higher resolution of WGS as well as allowing for LD (Moghaddam et al. 2016; Valdisser et al. 2017).

2.5.6 QTLS and Candidate Genes Associated with Determinacy

Three QTLs in Chromosome 1

A determinacy QTL was identified in chromosome 1 -Pv01- (D1.4-D1.6; Table 2), identified in other studies (da Silva et al. 2018; Kamfwa et al. 2019; Keller et al. 2022; Moghaddam et al. 2016; Sedlar et al. 2020; Vargas et al. 2021) as a hotspot of allelic variation, named the *Fin* locus. The *Fin* locus has been mapped to ~44.5Mb in chromosome 1 (Kamfwa et al. 2019; Pérez-Vega et al. 2010). This co-segregates with an upstream gene, *TFL1y* (*Phvul*.001G189200), a candidate gene for flowering, vegetative growth, rate of plant production and determinacy (Campa et al. 2018; Cichy et al. 2015; Delfini et al. 2021; González et al. 2016; Kwak et al. 2012; Kwak et al. 2008; Repinski et al. 2012). Consequently, the *Fin* locus has pleiotropic effects due to associations with many development traits such as determinacy, shoot biomass, days to flowering, days to maturity, plant architecture, embryo abortion, number of pods per plant, number of seeds per plant (seed yield and weight) (Delfini et al. 2021; González et al. 2016) and disease resistance (Delfini et al. 2021; González et al. 2016; Miklas et al. 2001; Soler-Garzón et al. 2024). Consequently, segregation for this QTL hotspot in Pv01 may prove difficult in breeding programmes due to these pleiotropic effects (Vargas et al. 2021).

Further candidate genes have been identified in this QTL, such as *Phvul.001G192200*. This gene is an ortholog of *LIGHT-REGULATED WD1* (*LWD1*), a gene involved in the circadian rhythm

pathway (Delfini et al. 2021; Moghaddam et al. 2016; Wu et al. 2008), or *Phvul.001G192300,* which is an ortholog of *SPINDLY* (*SPY*). *SPY* interacts with genes in the reproductive pathway (da Silva et al. 2018; Moghaddam et al. 2016; Tseng et al. 2004) and has been associated with days to maturity (Reinprecht et al. 2024).

Another QTL was identified on chromosome 1 (D1.3; Table 2) that contains the gene *Phvul.001G168700*. This gene is related to the PIF1 (Phytochrome interacting factor1) transcription factor isoform X1 in the legume *Vigna radiata* (Bateman et al. 2023). This bHLH (basic helix-loop-helix) transcription factor is involved in many light dependent pathways in plant development and interacts with circadian clock genes (Kim et al. 2016).

QTL D7.1 in Chromosome 7

The QTL at chromosome 7 (D7.1) was identified in the whole and Andean panel. The QTL contains the gene *Phvul.007G244700.* This is related to a transcriptional corepressor, Leunig-homolog (LUH) in *Vigna radiata* (Bateman et al. 2023). In *Arabidopsis*, Luenig-homologs have functional redundancy with Leunigs (LUGs), and are involved in embryo and floral development (Sitaraman et al. 2008). This QTL has been associated with seed size, seed weight and growth habit (da Silva et al. 2018; Elias et al. 2021; Keller et al. 2022; Kwak et al. 2008), suggesting it may have pleiotropic effects.

QTL D8.2 in Chromosome 8

The QTL identified on chromosome 8 (D8.2; Table 2) for determinacy overlaps with a region previously associated with plant architecture (da Silva et al. 2018). During this work no gene with a clear function was identified. However, in this study a possible candidate gene was identified for further investigation; *Phvul.008G170000.* This encodes a putative FAF (fantastic four) domain-containing protein. In *Arabidopsis,* FAF proteins regulate shoot meristem size and architecture (Wahl et al. 2010).

QTL D9.1 in Chromosome 9

The QTL D9.1 in chromosome 9 was identified in the whole and Andean panel. Nearby QTLs have been identified for yield and determinacy (Campa et al. 2018; Kamfwa et al. 2015). The gene *Phvul.009G138100* is found within this QTL and contains the significant MTA found by GAPIT (Wang and Zhang 2021). This gene has an insertion that possibly affects function (Cingolani et al. 2012). This gene is uncharacterised in common bean but has homology to the Root Meristem growth factor 9 from *Glycine soja* (Bateman et al. 2023; Goodstein et al. 2012). This growth factor is expressed in the roots and flowers, regulating and maintaining apical meristems, and therefore both root and floral development, seed size and leaf architecture (Chen et al. 2019; Shinohara 2021). Although this gene has previously identified as a candidate associated with Mesoamerican domestication (Schmutz et al. 2014), in this study the QTL was found in the Andean panel, suggesting that it has also played a role in the Andean domestication event.

QTL D10.1 in Chromosome 10

The QTL on chromosome 10 (D10.1) is located near to QTLs for plant height and number of nodules, and near genes associated with metabolic changes during domestication, once again suggesting pleiotropic effects (de Souza et al. 2023b; Delfini et al. 2021). Three of the genes within this region encode bHLHZip (Basic Helix-Loop-Helix Leucine Zipper) proteins: *Phvul.010G158500, Phvul.010G158300* and *Phvul.010G158200.* These bHLH transcription factors may be involved in the regulation of flowering genes (Zhou et al. 2019). The gene *Phvul.010G158500* displays non-synonymous modifications in the panel, including insertions, deletions and other variants linked to frameshift mutations and gained stop codons (Cingolani et al. 2012). Homology to *Vigna angularis* suggests this gene may be related to the transcription factor bHLH25, and possibly linked to a circadian rhythm-associated protein (Goodstein et al. 2012).

2.5.7 QTLS and Candidate Genes Associated with Photoperiod Sensitivity

QTL PS4.1 in Chromosome 4

One QTL for photoperiod sensitivity was found on chromosome 4 (PS4.1; Table 2) from the analysis on the whole panel. Within this QTL, four genes were identified, three of which (*Phvul.004G110200*, *Phvul.004G110301*, *Phvul.004G110000*) have non-synonymous mutations such as a stop lost, stop gained or a frameshift mutation in the panel (Cingolani et al. 2012). However, the genes are uncharacterised.

Two QTLs in Chromosome 5

Two QTLs were identified in chromosome 5: PS5.2 for the Andean panel and PS5.1 for the whole panel. PS5.2 overlaps with a previously identified QTL for seed weight, days to flowering and pod weight (Arriagada et al. 2023; Reinprecht et al. 2024). However, this previous analysis, that used a limited number of markers, did not identify a candidate gene. Based on sequence homology with *Vigna radiata,* the gene *Phvul.005G077000* was identified, which encodes a Proton gradient regulation 5 (PGR5) protein (Bateman et al. 2023). PGR5 is involved in plant growth under different light conditions due to interactions with Photosystem I, and consequently putatively associated with differentiating photoperiod sensitivity in the panel (Munekage et al. 2002). The QTL PS5.1 contained two genes, one of which, *Phvul.005G076300*, may encode a bidirectional sugar transporter, named SWEET protein. Evidence suggests SWEET proteins have essential roles in plant development, including in reproductive organs and bud growth (Gautam et al. 2022).

Two QTLs in Chromosome 7

Two QTLs were also identified on chromosome 7; PS7.1 and PS7.2, both in the Andean and the whole panel. The QTL PS7.2 contains the genes *Phvul.007G157400* and *Phvul.007G156200*. Homology with *Arabidopsis* suggests that *Phvul.007G157400* encodes a BANQUE3 BHLH161 protein. BANQUE3 is negatively regulated by *APETALA3* and *PISTILLATA* in petals and is involved in light-regulated responses and flowering time (Huala et al. 2001; Mara et al. 2010). *Phvul.007G156200* may encode the BHLH transcription factor PIF4 (Phytochrome Interacting Factor 4) based on homology with *Vigna radiata* and *Glycine soja* (Bateman et al. 2023;

Goodstein et al. 2012). PIF4 is a downstream signalling component integrating environmental cues such as light (Bateman et al. 2023).

The QTL PS7.1 overlaps with a previously identified QTL for plant production traits (González et al. 2016). The QTL includes the gene *Phvul.007G117400* which encodes a putative JUMONJI domain containing protein (Goodstein et al. 2012). JUMONJI proteins are involved in multiple plant developmental processes such as flowering and leaf senescence (Gan et al. 2014; Liu et al. 2019b). *Phvul.007G117400*'s homology with a JUMONJI16 orthologue in *Vigna radiata* also supports this role (Bateman et al. 2023).

Two QTLs in Chromosome 8

One of the QTLs found in chromosome 8 is PS8.1, from the whole panel. This QTL has been associated with determinacy (Campa et al. 2018), seed weight (Elias et al. 2021), days to flowering (Raggi et al. 2019) and pod number (Kamfwa et al. 2015). Due to the marker technology used, the QTL for seed weight was large, so had low resolution (Elias et al. 2021). The results (Figure 14) suggest a correlation between days to flowering, determinacy and photoperiod sensitivity under the same QTL. The significant MTA for this QTL was within the gene *Phvul.008G048300*. However, the function of this gene is currently unclear.

The other QTL found on chromosome 8 is PS8.2, which has previously been identified for seed weight (Blair et al. 2006). Genes within this QTL include *Phvul.008G085000, Phvul.008G084500, Phvul.008G084900* and *Phvul.008G084100*. *Phvul.008G085000* is homologous to *gibberellin 2-oxidase 8* in *Arabidopsis* (Huala et al. 2001). Gibberellin oxidases may respond to light intensity and can therefore be related to photoperiod sensitivity (Zhang et al. 2022). *Phvul.008G084100* is homologous to *CLAVATA3* in *Arabidopsis, a* gene that regulates shoot and floral meristem development (Clark et al. 1995; Hirakawa 2021). *Phvul.008G084900* is homologous to genes encoding ovate family proteins (OFPs). OFPs appear to be sensitive to light stimuli (Shahzaib et al. 2024). *Phvul.008G084500* has homology with *RAVEN/INDETERMINATE DOMAIN5* in *Arabidopsis,* which is linked to GA signalling pathways as well as other plant developmental pathways (Aoyanagi et al. 2020; Sanchez-Corrionero et al. 2019). *Phvul.008G085000* and *Phvul.008G084900* also both contain insertions or deletions with high impact non-synonymous mutations which, therefore, possibly affect function (Cingolani et al. 2012).

Two QTLs in Chromosome 9

A QTL was identified on chromosome 9 in the Andean panel (PS9.1). This was near a QTL associated with grain yield (Elias et al. 2021), post-harvest index (Sedlar et al. 2020), shoot biomass (Kamfwa et al. 2019), seed size (da Silva et al. 2018), days to flowering, and yield (Blair et al. 2006). Genes within the QTL included *Phvul.009G229100, Phvul.009G229200, Phvul.009G229700* and *Phvul.009G229900. Phvul.009G229100* is homologous to PIN3 transcription factor genes, involved in regulating root and shoot growth (Goodstein et al. 2012; Haga and Sakai 2012). Homology with *Arabidopsis* suggests *Phvul.009G229200* and *Phvul.009G229700* are involved in root growth (Huala et al. 2001), and that *Phvul.009G229900* encodes a *HAB1 (Hypersensitive To Aba1) homology to ABI (Abscisic Acid-Insensitive)1* gene involved in ABA signal transduction, which is regulated by circadian rhythm (Kamrani et al. 2022; Leitao et al. 2021b). The other QTL in chromosome 9 (PS9.2) was found in the whole panel and included the gene *Phvul.009G145100*, which was also related to an ABA response gene in *Arabidopsis*. A nearby QTL to PS9.2 was previously identified for days to flowering (Keller et al. 2022).

QTL PS11.1 in Chromosome 11

The QTL at chromosome 11 (PS11.1) was near a QTL for seed weight (da Silva et al. 2018) and a QTL for disease resistance (Banoo et al. 2020). This may be due to pleiotropic effects or low resolution of the previous analysis, using a limited number of markers. The gene *Phvul.011G004000* is in the QTL, which encodes a putative PHD (plant homeodomain) finger protein. PHDs have been found to be involved in the regulation of flowering time (Qian et al. 2021; Zhou et al. 2019). Other genes within the QTL are related to root or shoot growth. For example, homologous genes of *Phvul.011G003200* and *Phvul.011G003400* suggests that they are implicated in processes involved in root meristem development (Huala et al. 2001). *Phvul.011G003700* is an uncharacterised gene in common bean, but homology with *Arabidopsis* suggests it may be associated with phytochrome interacting factor 7 (PIF7) to regulate hypocotyl elongation (Huala et al. 2001; Leivar et al. 2008). However, there are many genes within this QTL and further research is needed to clearly distinguish a candidate gene.

2.5.8 Mean allele dosage

The allele dosage is calculated from the VCF file for each accession and SNP site, while the average is calculated by averaging these values for the accessions over a population or subpopulation. The mean allele dosage has been used in prior pipelines, (Wang and Fei 2025), often paired with the Fst score, per SNP site (Salojaervi et al. 2024) or the SNP density in windows of 50 SNPs (Lobaton et al. 2018). In this PhD, the mean allele dosage was plotted within the coding regions of genes in QTLs identified from a GWAS. This allows fixed genetic variations to be found and helps identify putative causative genes for phenotypes, alongside the usual analyses. Also, it helps to narrow down candidate genes for future molecular work such as transforming and mutating genes to experimentally assess their functional impact (Figure 19 and Figure 20).

In certain QTLs, such as D1.3, D1.4, D1.5, D7.1, D10.1, PS4.1, PS5.1 and PS5.2, large haplotype blocks in the indeterminate or determinate types were different based on mean allele dosage and, this can be observed in both gene pools. This suggests that certain genes within the QTLs are causative for the phenotypes. However, since determinacy and photoperiod insensitivity developed independently in the two gene pools, this is likely an example of convergent evolution (Weller et al. 2019). In most other QTLs the differences were only observed in particular subpopulations, and phenotypes (private alleles). This suggests that determinacy and photoperiod insensitivity evolved at a later stage after the subpopulation divergence (Kwak et al. 2012). Consequently, when identifying QTLs for phenotypes that are present in both domestication centres, it is important to have diversity covering both gene pools.

The dosage can also be used to infer introgressions (Vexler et al. 2024; Wang and Fei 2025). The reference genome is Andean (Schmutz et al. 2014), so a homozygous reference call (HOMR; dosage 0) is likely to suggest an Andean background. A homozygous alternative call (HOMA; dosage 2) could be hypothesised as from the Mesoamerican background (less frequently a private allele in the Andean background). For example, the ‘determinate M’ subgroup has a higher proportion of HOMR diversity in the QTL D1.5, suggesting an introgression from the Andean background that could be associated to the phenotypes. In the QTL PS8.1, the ‘insensitive M’ subgroup has a higher proportion of HOMR, not present in the other Mesoamerican group but shared with the Andean subgroups, the admixed are predominantly heterozygous. This spans multiple candidate genes, again suggesting an introgression from the Andean background into ‘insensitive M’.

Within the QTLs, there are genes that showed no clear differences in allelic mean dosage among the subpopulations. This does not necessarily discard them as candidate genes, as single non-synonymous substitutions can still have functional consequences at the protein level. Consequently, these analyses into the genes within QTLS should still be paired alongside homology and functional studies. Also, heterozygous sites were no discussed during this work, this is because the phenotype could be from either the reference or alternative allele. Without further work into the effects of dominance or additivity, phenotypic variation cannot be distinguished in these sites (Kalyta et al. 2023; Sanjak et al. 2017). Certain QTLs have large regions of uncalled variation. In some cases, such as Phvul.002G116500, the whole gene is likely absent (or a largely different allele) in the Andean gene pool.

In my PhD, the priority was supporting the identification of putative candidate genes within the QTLs. However, future work could include non-coding elements. There is evidence that non-coding regions such as promoters, silencers and enhancers can impact gene regulation and therefore phenotypes (Barrett et al. 2012; Chen et al. 2023). Further to this, other regulatory mechanisms have been found in prior research, such as epigenetic modifications (e.g. DNA methylation) (Zhang and Zhu 2025), post-transcriptional modifications (e.g. small RNAs)(Yu et al. 2026) and post-translational protein modifications (Han et al. 2022a).

Chapter 3

Natural variation in Colombian common beans reveals diverse water deficit response strategies and genetic loci through GWAS and phenotyping

Contributions: All work completed during the PhD by Kate E Denning-James and other authors supervised, reviewed and edited. Other PhD students and members of the lab group helped with sowing, harvesting and phenotyping, further detailed in acknowledgments.

3.1 Abstract

Water scarcity has been identified as one of the most significant global environmental risks for food security. In common beans (*Phaseolus vulgaris*), drought stress can cause yield losses of up to 50%, leading to significant economic losses and reduced food security. Since plants are unable to escape unfavourable conditions, breeding programs have focused on novel genetic diversity as a promising route to introduce greater drought resilience.

144 accessions of common beans were sequenced, from a variety of locations in South and Central America, concentrating on Colombia, where water deficit can be a major threat to productivity. The accessions were exposed to water deficit by stopping irrigation during vegetative growth. The effects of water deficit on phenology were assessed by measuring developmental changes using the BBCH scale, on leaf responses (transpiration and photosynthesis) using porometry and fluorometry, and on biomass and harvest parameters, including pod and foliar weight.

Within the panel multiple drought tolerance strategies were identified across genotypes, as well as drought sensitive accessions and putative drought tolerant accessions. For example, 'spenders' that accelerated development toward reproduction and 'savers' that delayed growth to conserve resources. Photoperiod sensitive accessions were also observed, for which drought stress overrides daylength signals, resulting in early pod production.

A GWAS on phenotypes collected from the water deficit-stressed accessions identified significant quantitative trait loci (QTLs) and putative candidate genes potentially regulating water deficit resilience mechanisms.

To combat the impacts of climate change on common bean yields, a multi-pronged approach will be required, including policy changes, improved infrastructure for water access and irrigation strategies and targeted breeding programmes. The results from this study provide genetic insights that could aid in developing drought tolerant common bean varieties which can contribute to future food security.

3.2 Introduction

3.2.1 Common Beans, Phaseolus vulgaris

Common beans, *Phaseolus vulgaris*, are one of the most important legumes for human consumption (Broughton et al. 2003). They have a high nutritional content of proteins and minerals, which is especially important in developing countries. Common beans have two centres of domestication, the Andean and Mesoamerican. The two domestication gene pools can cross creating admixed diversity and introgressions from one background to another. This admixed genetic diversity and the untapped diversity in landraces and crop wild relatives, provide novel sources of diversity to explore for resilience traits against water-deficit.

3.2.2 Drought stress in agricultural crops

Plants largely respond to drought in two ways, by avoidance or escape (Polania et al. 2022). Drought escaping plants rapidly complete their lifecycle before the onset of drought (Shavrukov et al. 2017), with progeny surviving as desiccated seeds. Drought avoidance strategies include ‘water spenders’ (anisohydric plants) that accelerate development and 'water savers' (isohydric plants) that delay growth to conserve resources.

Water spenders maintain photosynthesis during drought stress by keeping their stomata open at the expense of water loss. If this strategy enables reproduction and seed set (with parallels to drought avoiders), it can be a useful crop trait that prioritises yield under long-term droughts (Nesporová et al. 2024). Water savers are plants which reduce photosynthesis by closing stomata rapidly upon the detection of water deficit. Another strategy, the functional ‘stay-green’ (SG) response, delays leaf senescence, thereby maintaining chlorophyll levels, and prolonging photosynthetic assimilation.

The process of domestication, and modern breeding for high yielding varieties, have both limited the genetic diversity of our crops and inadvertently produced varieties with high stomatal conductance and low water use efficiency (Huang and Zeng 2024; Lei et al. 2023). Utilising and mobilising natural diversity during breeding programmes could address these imbalances to improve commercial cultivars’ drought stress resilience. Landraces and crop wild relatives are a rich and underexplored source of novel diversity, which could help to combat abiotic and biotic stresses (Renard et al. 2023).

3.2.3 Common beans and drought stress

Common beans are particularly susceptible to drought stress, with predictions suggesting ~60% of common bean production is impacted by it (Beebe et al. 2008; Villordo-Pineda et al. 2015). Drought stress is becoming more frequent in many regions of common bean production due to climate change and ENSO (El Niño Southern Oscillation) impacts. Prior research into drought tolerance in common beans have identified water deficit tolerant cultivars in the Mesoamerican gene pool (Polania et al. 2017a; Villordo-Pineda et al. 2015), the Andean gene pool (Dramadri et al. 2019) and in wild accessions (Cortés et al. 2013). However, drought tolerance strategies are not specific to the population structure. Exploring a diverse set of accessions from Colombia and neighbouring countries, utilising a GWAS, could reveal novel genetic diversity.

3.2.4 Research aims

This research aims to identify divergent drought tolerance strategies across a diverse panel of Colombian common beans. Within this diversity panel, different drought tolerance strategies are compared, identifying those genotypes worthy of further investigation. Further to this, GWAS is leveraged to identify genomic regions and putative candidate genes associated with resilient physiological and developmental traits in response to water deficit. Consequently, this study aims to elucidate how phenotypic traits, and developmental stages link different drought tolerance strategies, to inform future common bean breeding programmes.

3.3 Methods

3.3.1 Diversity panel and genotyping

A diversity panel of 144 common bean accessions, as described in section 2.2.3 (Denning-James et al. 2025), was used for a genome wide association study (GWAS). These accessions represent a large amount of the diversity from Colombia and neighbouring countries, as well as both Andean and Mesoamerican gene pools, and a variety of races.

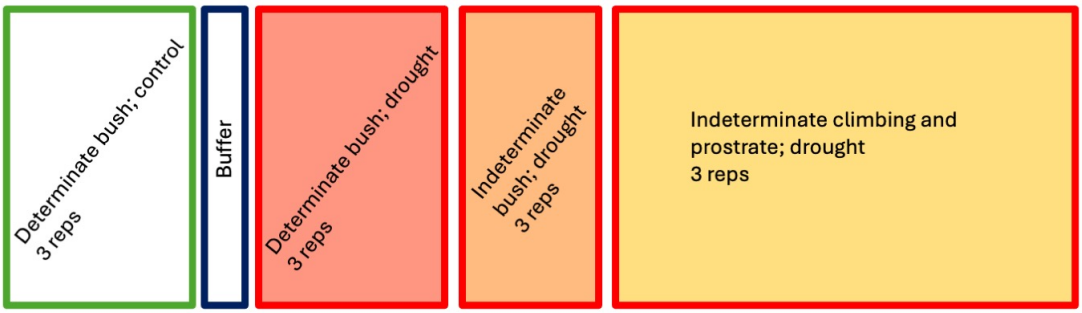
The genotypic data after whole genome re-sequencing contained 20.2 million variant loci (~17.1M SNPs and ~3.4M indels). This dataset was then filtered further for biallelic loci, a minor allele frequency of 1% and thinned to remove clustered SNVs within a window size of 5 bp, using BCFtools (v. 1.12) (Danecek et al. 2021). The resulting VCF for GWAS had 2,572,124 loci, as in section 2.3.3 and 2.3.6.

3.3.2 Experimental design

The experiment was executed outdoors at the NIAB research station in Histon, Cambridge (52.25, 0.10) from June to September 2023. An experimental unit was a single plant grown in a 5L pot. The accessions were organised by growth habit and arranged within each habit x treatment using a randomised block design (Chapter 2, (Denning-James et al. 2025)). For each treatment x growth habit, plants were distributed across three blocks, and pot positions were randomised within blocks to minimise environmental effects. 38 of the accessions were determinate bush, 9 were indeterminate bush, 86 were indeterminate climbing and 11 were prostrate.

There were two irrigation treatments, well-watered control and water deficit. Due to space limitations, only determinate bush accessions had both well-watered control treatments and water deficit treatments, with three biological replicates per accession per treatment. All other accessions (indeterminate bush, indeterminate climbing and prostrate) were subjected only to water deficit treatment, again with three biological replicates per accession. Consequently, the study is comparative focusing on relative performance between the accessions under water deficit. The controls were included to confirm that the cessation of irrigation imposed stress relative to irrigated plants and to provide a baseline. To prevent the movement of water between the well-watered and water deficit stressed accessions a 1 metre buffer zone was planted with a commercial variety of common bean (Figure 21).

A



B

Figure 21; Layout of the diversity panel (144 accessions with 3 water deficit replications), for the water deficit experiment at NIAB, Cambridge in 2023. 38 of the accessions were determinate bush, 9 were indeterminate bush, 86 were indeterminate climbing and 11 were prostrate.

All accessions were sown on 19/06/2023 (day 0) in 5 litre pots, containing Sylvamix compost (Nitrogen 250 mg/l, Phosphorus 80 mg/l; Potassium 300 mg/l, Melcourt Ltd, UK). This provided a basal nutrient supply, and no additional fertiliser or feed was applied during the experiment. Plants were thinned to 1 plant per pot on 06/07/2023 (17 DAS), after reaching BBCH stage 12, representing the growth of true leaves (Cavalcante et al. 2020). During early development, all the plants were irrigated daily, and pots were spaced at 15-20cm intervals.

Water deficit was initiated on 27/07/2023 (38 DAS), when irrigation was stopped. After this water deficit pots only received water via precipitation (Figure 21). The stress was imposed after the common beans had established and before the flowering developmental stage, to minimise the impacts from the transition to reproduction on results. On average the BBCH was below 55 on 28/07/2023 (39 DAS) (Wu et al. 2021). In contrast the control plants received irrigation throughout the trial and their development.

Table 3; Key dates within the water deficit experiment (June 2023 to September 2023), the stage of the trial and the number of days after sowing (DAS).

Date	Experimental stage	Days after sowing
19/06/2023	Sowing	0
06/07/2023	Thinning	17
21/07/2023	Normal irrigation	22 (Week 0)
27/07/2023	Irrigation removed	38
28/07/2023	Day 1 water deficit	39
01/08/2023	Day 4 water deficit First data recording	43 (Week 1)
04/08/2023	Day 7 water deficit	46
10/08/2023	Day 13 water deficit Second data recording	52 (Week 2)
18/08/2023	Day 21 water deficit Third data recording	60 (Week 3)
23/08/2023	Day 26 water deficit Fourth data recording	65 (Week 4)
25/08/2023	Day 1 Recovery	67
30/08/2023	Day 6 Recovery Fifth data recording	72 (Week 5)
06/09/2023	Harvesting	79
09/09/2023	Harvesting	82

3.3.3 Environmental conditions

During the experiment the mean precipitation was 1.73 mm a day, and this reduced to 1.39 mm when the irrigation was stopped (Visual Crossing Corporation 2024). This treatment represents a typical drought, whereby the water received did not meet the demands of the plants for optimal growth and development (Dramadri et al. 2019; Wortmann et al. 1998). On 25/08/2023 (day 67, week 5) the precipitation increased, thereby inducing a mock recovery period after the water deficit was relaxed. The mean temperature from sowing to harvesting was 17.6 °C and the average daylength during the experiment was 15.4 hours (Figure 22).

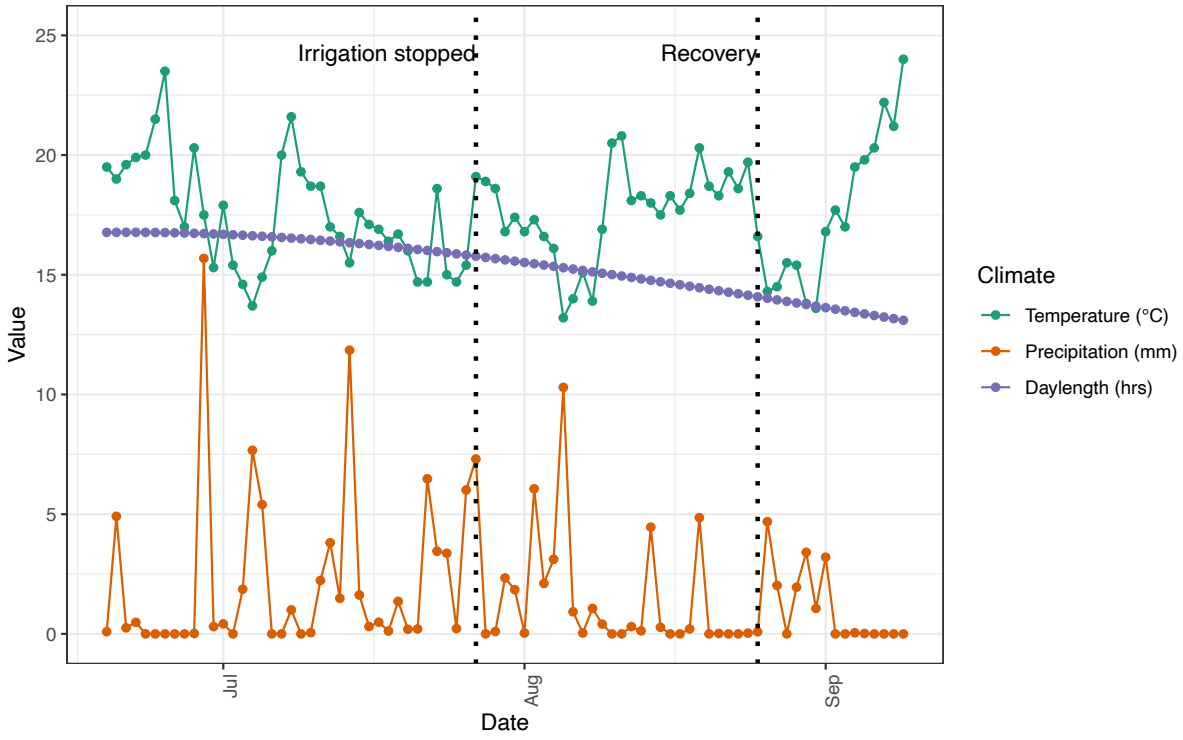


Figure 22; Temperature (°C), precipitation (mm) and daylength patterns (hrs) during the experiment in Cambridge, UK, from 19/06/2023 (sowing) to 09/09/2023 (harvesting) (Visual Crossing Corporation 2024). The dotted lines represent key events during the experiment. On 27/07/2023 (38 DAS, week 1) irrigation was stopped to induce the water deficit treatment and on 25/08/2023 (67 DAS, week 5) the plants experienced a mock recovery period due to precipitation.

After sowing, fifteen Teros-21 (Decagon Company) and twelve Teros-12 (Decagon Company) (Cominelli et al. 2024) soil sensors were randomly assigned to pots, distributed among the growth habits and replications (Table 4). Together, these collected hourly, supplementary information on soil water potential, soil moisture content, soil temperature and electrical conductivity. ZL6 data loggers were used to collect real time data recordings (Meter Group, Inc). Data from the sensors was collected and analysed in R studio, utilising the package ggplot (R Core Team 2024; Wickham 2016).

To test for differences in the soil metrics between growth habits and stress (control or water deficit) over time, a two-way ANOVA was completed with the stress x growth habit and week as fixed factors, including their interaction. Post-hoc pairwise comparisons of estimated marginal means were calculated for significant interactions with the emmeans package in R (Lenth 2025). The P-values were then adjusted with the Holm method. Group differences were visualised with letters generated with multcomp and numbers assigned for each week (Hothorn et al. 2008). Groups sharing the same letter x number were not significantly different at $\alpha = 0.05$, whereas groups with different letters differed significantly. Data was plotted in R with ggplot2, paletteer and ggthemes (Arnold 2024; Hvitfeldt 2021; Wickham 2016).

Table 4; Distribution of soil sensors among the growth habits and treatments during the trial.

Treatment	Growth habit	Sensor	Number of sensors per treatment
Control	Determinate bush	Teros-21	3
		Teros-12	2
Water deficit	Determinate bush	Teros-21	4
		Teros-12	3
	Indeterminate bush	Teros-21	2
		Teros-12	2
	Indeterminate climbing	Teros-21	6
		Teros-12	5

3.3.4 Phenotyping, data collection and data analysis

Leaf-level water loss and photosynthetic parameters were measured across the diversity panel during the experiment using a LI-COR LI-600P/F (porometer/fluorometer). However, during the study, 2 accessions (JSPinto and JSBlackwater), died during water deficit therefore, the trial continued with 142 accessions only. The porometer outputs included stomatal conductance to water vapour (g_{sw}), transpiration (*E*_apparent), leaf vapour pressure deficit (VPDleaf), relative humidity of the sample (rh_s), and leaf temperature (Tleaf), and the fluorometer outputs included electron transport rate (ETR), steady-state fluorescence (Fs), maximum fluorescence (Fm'), and quantum efficiency in light (PhiPS2).

LI-600 measurements were initiated when irrigation ceased (27/07/2023, 38 DAS) and were collected weekly between the hours 13:00 BST and 17:00 BST. For every plant in each treatment, spot measurements were taken on the central leaflet of fully expanded trifoliate leaves that showed no signs of yellowing or damage.

To compare how water deficit differentially affected common bean phenological development, BBCH was recorded weekly (Cavalcante et al. 2020). At the end of the growth experiment, plants were harvested at soil surface level (09/2023). Their pods were counted and the fresh weight for the pods and leaves were measured separately to evaluate the differentiation between reproductive and vegetative growth. Finally, seed type (wild, landrace, and heirloom), country of origin, population structure (admixture coefficients estimated by ADMIXTURE; K = 6), seed colour, photoperiod sensitivity and growth habit information were included from the previous analysis (Chapter 2, Denning-James (2025), and Appendices 1, 4 and 8).

The phenotypic data were tested for normal distribution using the Shapiro-Wilk test from the rstatix package (p-value > 0.05) (Kassambara 2023). The statistical analysis of variance of the phenotypic data between continuous datasets was completed in R with Spearman's rank correlation for non-normal distribution. Correlations between continuous and discrete data were calculated with Kruskal-Wallis for non-parametric data (Kassambara 2023; Kruskal and Wallis 1987; R Core Team 2024; Wei and Simko 2021). The correlation between two discrete datasets was calculated with Cramér's V and significance with chi-squared (Mangiafico. 2025). This was visualised using the R package 'corrplot' and 'ggplot2' (R Core Team 2024; Wei and Simko 2021; Wickham 2016). Analysis of the significant difference between growth habits across the porometer and fluorometer variables was completed with a two-way ANOVA as for the soil metrics (section 3.3.3).

3.3.5 Genome wide association studies

A genome-wide association study (GWAS) was completed using GAPIT v.3 (Wang and Zhang 2021) with three principal components for all phenotypes and data collections in Table 3 (Appendix 4). 'Trait' measurements derived from the LI-600 were labelled as 'initial response' to water deficit (Week 1 and Week 2) and 'recovery' from water deficit (Week 4 and Week 5). Other measurements included the differences between data collections (Week 1 to Week 4 and Week 1 to Week 5), averages across all the dates, sum of all the dates and the scaled sum of all the dates (to normalise the values using the built in R function). The data from harvesting such as pod weight, foliar weight and partitioning weight (the ratio between pod weight and foliar weight, to determine how plants allocate resources) were also included (Polania et al. 2016; Rehling et al. 2021). Phenotypic data was visualised using different methods to identify quantitative trait loci (QTLs) and to find QTLs that arose in more than one method.

The final two recordings (Weeks 4 and 5) were used to evaluate accession 'recovery' following water deficit. As is shown in Table 3, week 4 is the last recording of water deficit and week 5 is after the 'recovery' rainfall. Consequently, showing the difference in accession responses before and after the 'recovery', can help identify how or if the accession responded to the re-addition of water. Significance was calculated with the Wilcoxon signed-rank test for non-parametric paired data (Rosner et al. 2006). The first two recordings (week 1 and week 2) were used to analyse the 'response' to water deficit. The differences between recordings were used to measure the change, magnitude, speed, and intensity of accession 'responses' to water deficit and 'recovery' over time.

GAPIT was run on the whole diversity panel with the models Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) (Huang et al. 2019), Fixed and random model Circulating Probability Unification (FarmCPU) (Liu et al. 2016), Multiple Locus Mixed Linear Model (MLMM) (Segura et al. 2012) and Mixed Linear Model (MLM) (Zhang et al. 2010). BLINK, FarmPCU and MLMM were selected as multi-locus models for different heritability levels to improve the statistical power (Cebeci et al. 2023; Huang et al. 2019; Merrick et al. 2022). MLM was chosen as a single-locus analysis for comparison to the other models. To run GAPIT with the BLINK model, the analysis was completed with the parameter 'Random.model=TRUE' so that R^2 was not calculated for phenotypic variance explained values after GWAS. To determine the fit of the models to the phenotypic data quantile-quantile (QQ) plots were used. Plotting was completed in R using the packages 'ggplot2' (Ginestet 2011), 'qqman' (Turner 2018) and 'ggbreak' (Xu et al. 2021).

3.3.6 QTL identification and candidate gene prediction

Marker trait associations (MTAs) were investigated based on significance ($-\log_{10}$(p-value) > 7) and when the MTA was identified by two different phenotypic traits from any model. QTLs were defined as ±100 kbp from the MTA; this is based on an LD decay of 114 kb for the panel when R^2 = 0.25, as in Chapter 2 (Campa et al. 2018; Denning-James et al. 2025; Moghaddam et al. 2016; Reinprecht et al. 2024).

To prioritise causative genes within QTLs, the criteria previously tested in section 2.3.7 and Denning-James et al. (2025) was followed. The Andean reference genome, *Phaseolus vulgaris* G19833 v2.1, was visualised in JBrowse (Diesh et al. 2023; Schmutz et al. 2014) and used to identify the genes contained within the QTLs. SnpEff was used to identify non-synonymous 'high impact' mutations (Cingolani et al. 2012). eggNOG-mapper v2 provided functional annotation of the genes in the QTLs based on ortholog assignments (Cantalapiedra et al. 2021). GO terms GO:0009819, GO:2000070, GO:1902584, GO:0009414, GO:0009415, GO:0042630, GO:0042631, GO:0097207, GO:0006970, GO:0009651, GO:0009269, GO:0009992, and GO:0080148 were used to identify genes related to drought stress and abiotic stress responses.

The candidate genes were further explored using PhytoMine (Goodstein et al. 2012) (*Phaseolus vulgaris v.2*), BLAST (Camacho et al. 2009) against the non-redundant (nr) protein database at NCBI, and finally against the TAIR database, if no gene function could be identified in closer relatives (Huala et al. 2001). The loci were compared to previous studies and literature using PulseDB. QTLs and markers were mapped to the reference genome to estimate the conversion from cM to Mb, and visualised in JBrowse (Humann et al. 2019).

3.4 Results

3.4.1 Soil conductivity and moisture assessments

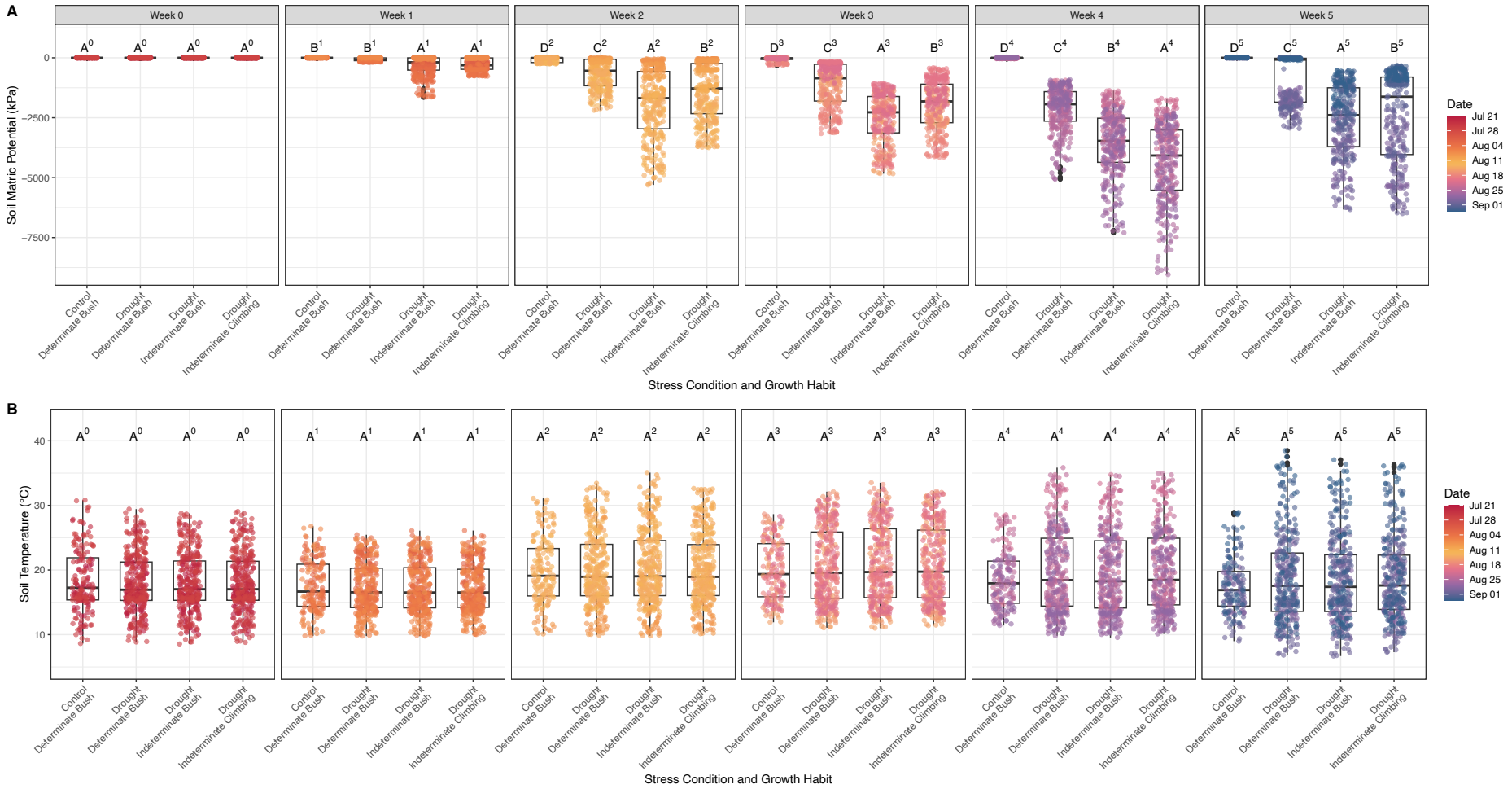
Figure 23 shows how the soil moisture metrics (soil matric potential, soil temperature, bulk electrical conductivity (EC) and water content) changed over time from thinning (17 DAS) to harvesting (79-82 DAS). Soil matric potential (kPa), bulk EC and water content clearly reflected when irrigation was stopped in the water-deficit stressed accessions.

During irrigation (week 0), all the growth habits started with a mean water content from 0.22 to 0.26 m^3/m^3, an average soil matric potential from -0.13 to -0.33 kPa and a Bulk EC from 0.17 to 0.08 mS/cm. The soil temperature ranged from 20.8°C in weeks 2 and 3 to 17.2°C in week 1 (Figure 23B). Soil temperature was never significantly different between the groups under control and water deficit treatments and stayed within 3°C from week 0 to week 5.

For the water-deficit treated plants, the Bulk EC decreased from an average of 0.08, 0.1 and 0.1 mS/cm in week 1 (removal of irrigation), to an average of 0.0067, 0.0075 and 0.0077 mS/cm in week 5, for the determinate bush, indeterminate bush and indeterminate climbing growth habits, respectively (Figure 23C). While the water content and soil matric potential also decreased from week 1 to week 4. In week 3, there was no significant difference in the water content and Bulk EC between the determinate bush and indeterminate climbing groups under water deficit. In week 4, there was no significant difference between the determinate bush, indeterminate bush or indeterminate climbing groups in water content or Bulk EC.

After the 'recovery' rainfall, there was a delayed response in the soil metrics, during which the soil matric potential, bulk EC and water content increased in all the groups between weeks 4 and 5 (Figure 23A, C and D). This delay suggests that the post water deficit precipitation was insufficient to balance the water lost. This was evidenced by the continued decline in soil matric potential, particularly for both groups of indeterminate accessions (Figure 23A).

The control determinate bush group experienced a water content from on average 0.22 to 0.14 m^3/m^3, a soil matric potential from -80.1 to -0.1kPa and a Bulk EC from an average of 0.03 to 0.17mS/cm. The soil matric potential and water content decreased during weeks 2 and 3, but the values were significantly different to those groups experiencing water deficit (Figure 23A and Figure **23**D). The bulk EC also decreased during weeks 2 and 3, but it was significantly different to the water deficit treatment in all weeks except week 3 (Figure 23C). This displays the significant differences in the soil metrics measured between the control and water deficit groups in the different growth habits (after irrigation was removed).


A
Week 0
Week 1
Week 2
Week 3
Week 4
Week 5
Soil Matric Potential (kPa)
0
−2500
−5000
−7500
Control Determinate Bush
Drought Determinate Bush
Drought Indeterminate Bush
Drought Indeterminate Climbing
Stress Condition and Growth Habit
Date
Jul 21
Jul 28
Aug 04
Aug 11
Aug 18
Aug 25
Sep 01
B
Soil Temperature (°C)
40
30
20
10

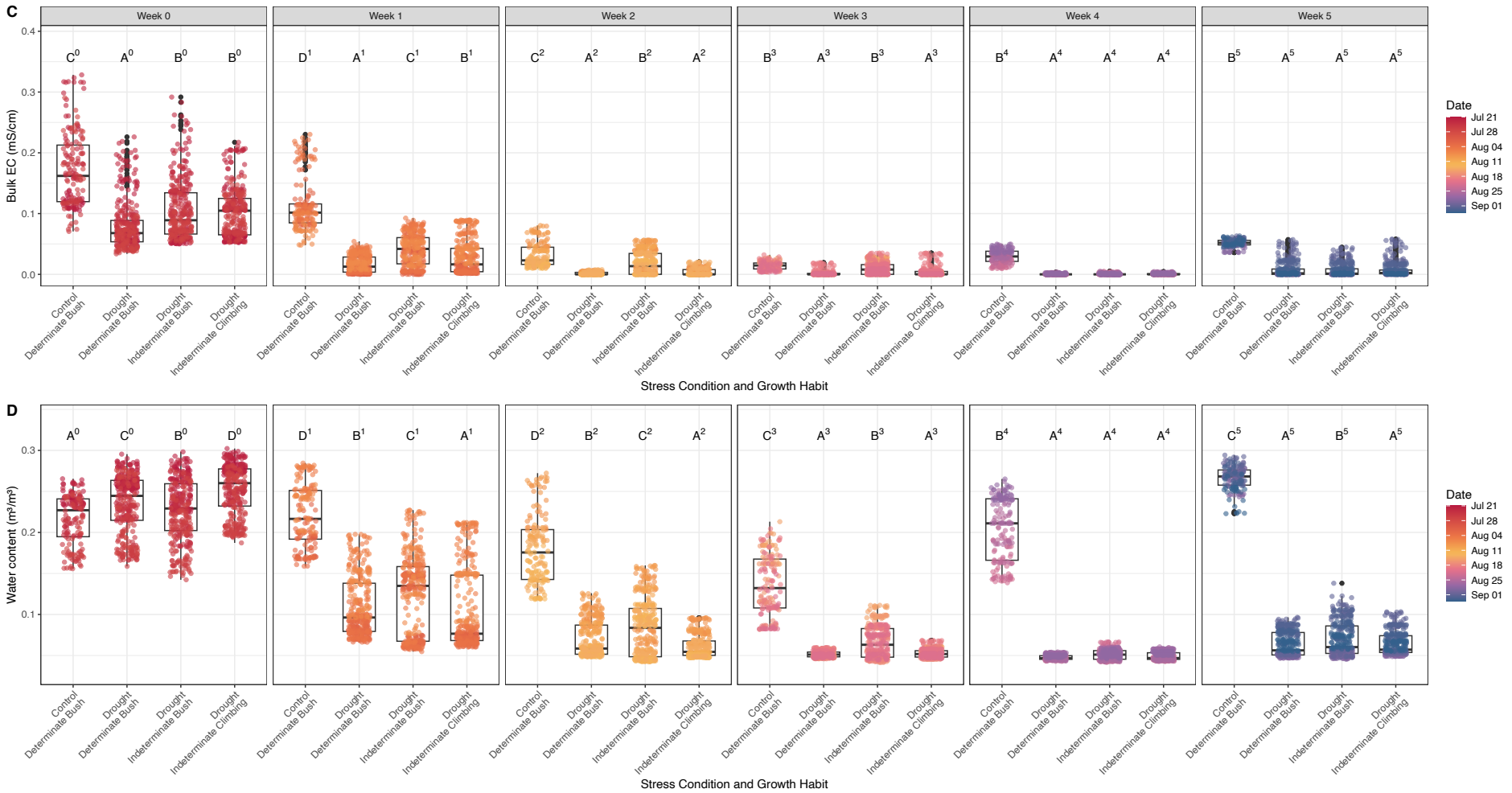


Figure 23; Bar plots of parameters collected from soil moisture sensors (Table 4). For (A) the Soil Matric Potential (kPa), (B) the soil temperature (°C), (C) the bulk electrical conductivity (EC, mS/cm) and (D) for the water content (m3/m3). The colours represent the dates from irrigation (week 0) to harvesting (week 5). Statistical analyses identified significant interactions between stress x growth habit over time. Groups were labelled by a letter and superscript number. Groups sharing the same letter were not significantly different at α = 0.05, whereas groups with different letters differed significantly. The superscript numbers in the labels corresponded to the week.

3.4.2 Phenotypic variation and correlations

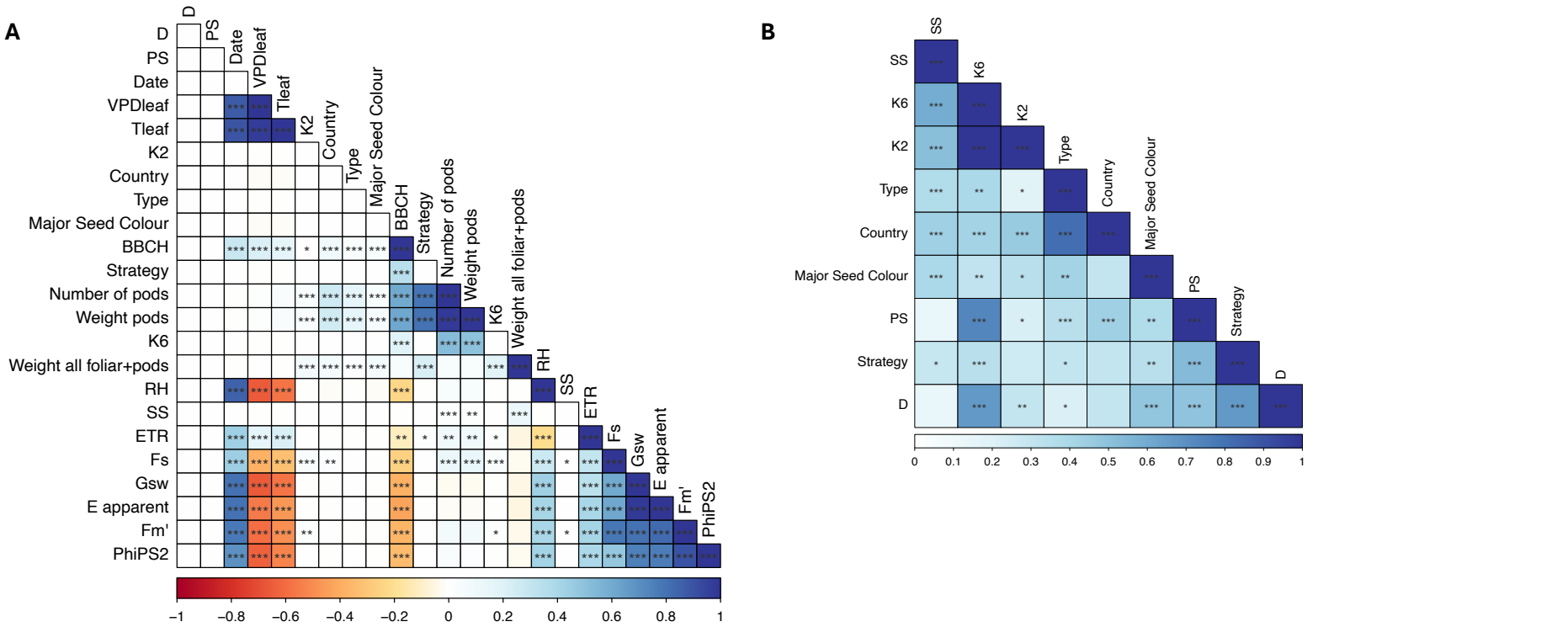


Figure 24; Correlation plot comparing phenotypes derived from the poromoter and fluorometer, BBCH (phenological developmental data), D (Determinacy), PS (Photoperiod Sensitivity), Date, VPDleaf (leaf vapour pressure deficit), Tleaf (leaf temperature), Country, Type (wild, heirloom, landrace), Strategy (drought tolerance strategy, Table 5), population structure at K6 and K2 (subgroups from ADMIXTURE), RH (relative humidity), SS (seed size), ETR (electron transport rate), Fs (steady-state chlorophyll fluorescence), GSW (g_{sw}, stomatal conductance), E apparent (E, transpiration), Fm' (maximum fluorescence yield), PhiPS2 (Φ_{PSII}, quantum yield of fluorescence). Data from water deficit treated plants (142 accessions) following the cessation of water application (five recordings between days 38 to 72, Appendix 4 and 5). (A) Correlation between continuous datasets, and correlation between continuous and discrete data, where discrete datasets were put to zero. (B) Correlation between discrete datasets.

Figure 24 shows the correlations between all the harvesting, porometer, fluorometer, population structure and growth habit phenotypes over multiple dates during, and after the cessation of water application on all water deficit treated plants (excluding controls), for comparative analysis. When investigating categorical values (Figure 24B), associations were determined between tolerance strategies and estimated seed size (ES = 0.3), Type (Wild, Landrace (1 commercial variety), or Heirloom; ES = 0.3), Country of origin (ES = 0.3), Major seed colour (ES = 0.34), K6 population structure ('Admixed K6', ES = 0.35) and growth habit (GH, ES= 0.5).

There are significant correlations ($p<0.01$) between many of the stomatal and photosynthetic parameters. Stomatal conductance (g_{sw}, r= -0.65, r=-0.57), transpiration (*E*_apparent, r=-0.56, r=-0.47), relative humidity (rh_s, r= -0.65, r=-0.58), steady-state chlorophyll fluorescence (Fs, r=-0.38, r=-0.33), maximum fluorescence yield (Fm', r=-0.6, r=-0.5), and quantum yield of fluorescence (Φ_{PSII}) (PhiPS2, r=-0.64, r=-0.53) all negatively correlate with leaf vapour pressure deficit (VPDleaf) and leaf surface temperature (Tleaf), respectively. VPDleaf and Tleaf positively correlate with each other (r=0.98) and VPDleaf positively correlates with BBCH (r=0.21), while g_{sw}, *E*_apparent, Fs, Fm, PhiPS2, and rh_s significantly negatively correlate. BBCH correlates positively with pod weight (r=0.62) and number of pods (r=0.61) and negatively correlates with g_{sw} (r=-0.39), *E*_apparent (r=-0.42), Fs (r=-0.26), Fm (r=-0.36), PhiPS2 (r=-0.34), growth habit (GH, r=-0.34). Pod number and pod weight positively correlate (r=0.99) and negatively correlate with GH (r=-0.48, r=-0.5) respectively (Figure 24A).

The average Fm' decreased for all growth habits from values above 300 in week 1 to average values below 150 in week 4 (Figure 25A). Fs has more subtle differences with the averages by growth habits ranging from 141-145 in week 1 to 100-124 in week 4 (Figure 25B). The leaf temperature increased from ~25°C in week 1 to <31°C in week 4 (Figure 25C). Stomatal conductance (g_{sw}) decreased throughout the water deficit stress from 0.2-0.3 in week 1 to ~0 in week 4 (Figure 25D). Transpiration (*E*) also decreased from 2.7-3.8 mmol $m^{-2}s^{-1}$ in week 1 to ~0 mmol $m^{-2}s^{-1}$ in week 5 (Figure 25E). Finally, PhiPS2 decreased from ~0.6 µmol $m^{-2}s^{-1}$ (week 1) to ~0.2 µmol $m^{-2}s^{-1}$ (week 5) (Figure 25F).

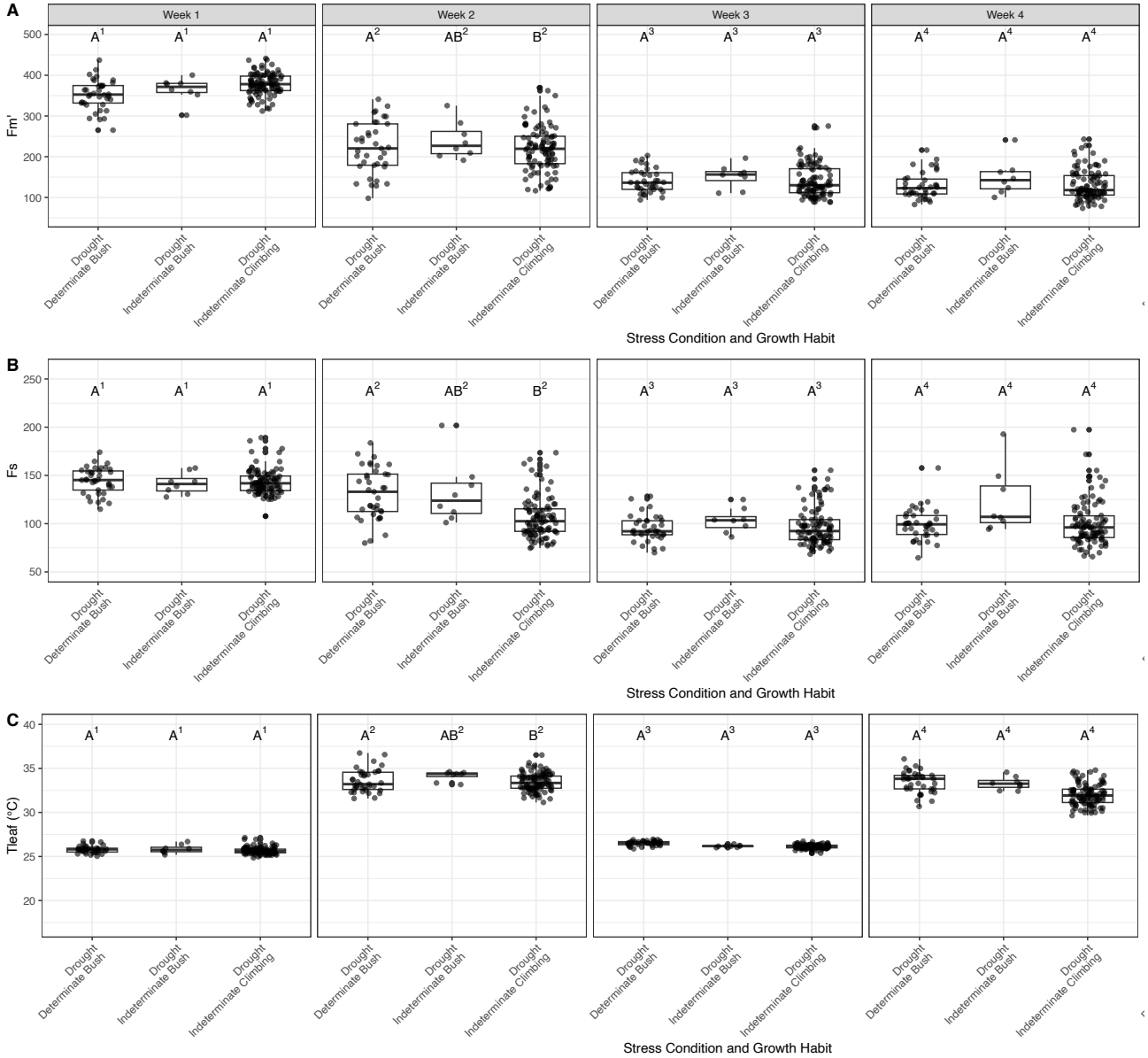
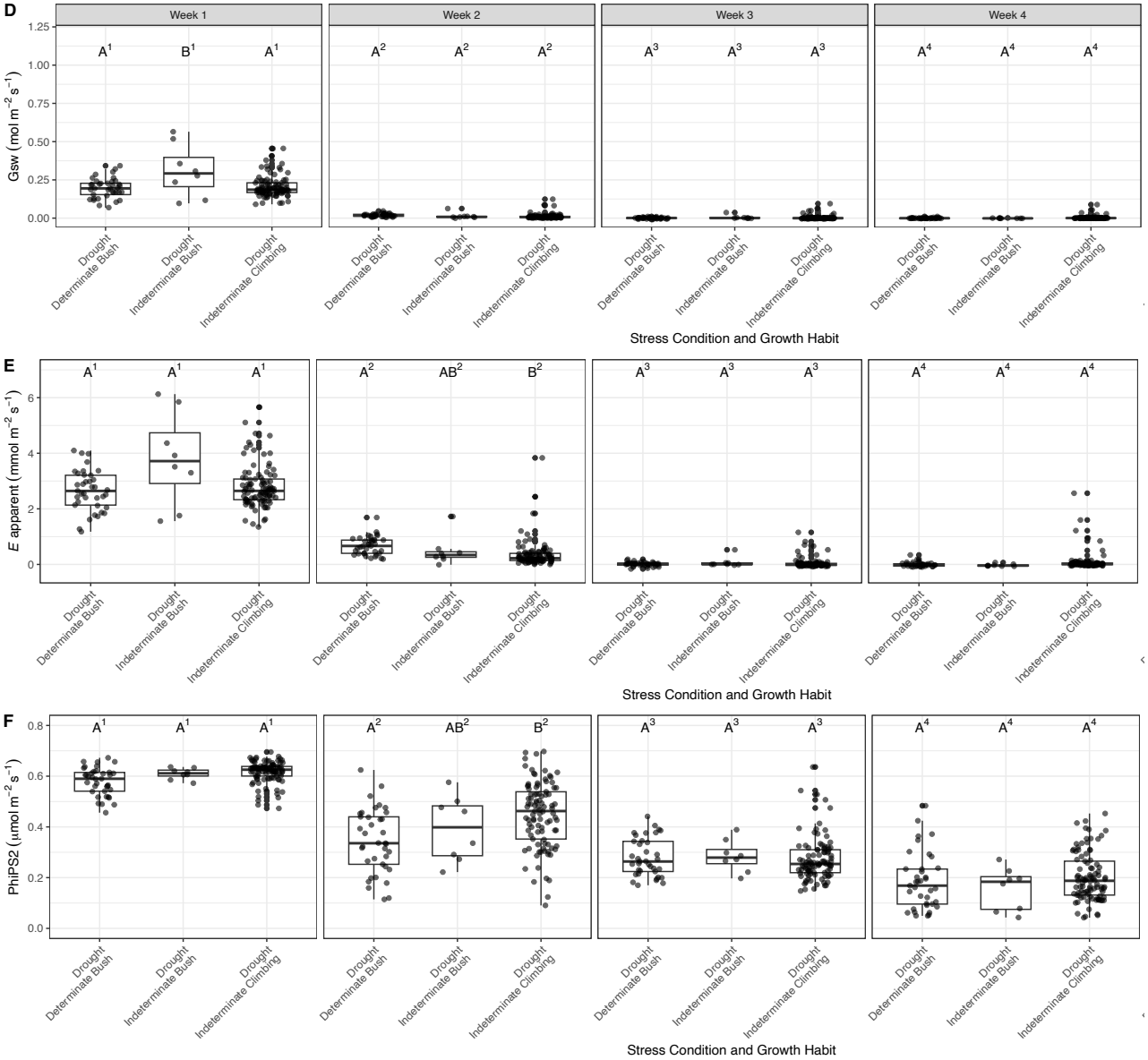


Figure 25; Bar plots of parameters collected from the porometer and fluorometer separated by growth habit from week 1 to week 4, for accessions experiencing water deficit (Table 3, 142 accessions; 37 determinate bush, 8 indeterminate bush, 97 indeterminate climbing and prostrate). (A) Fm' (maximum fluorescence yield), (B) Fs (steady-state chlorophyll fluorescence), (C) Tleaf (leaf temperature, °C), (D) GSW (g_{sw}, stomatal conductance, mol $m^{-2}s^{-1}$), (E) E apparent (E, transpiration, mmol $m^{-2}s^{-1}$) and (F) PhiPS2 (Φ_{PSII}, quantum yield of fluorescence, µmol m^{-2} s^{-1}). Groups sharing the same letter are not significantly different at α = 0.05, whereas groups with different letters differed significantly. The superscript numbers correspond to the week.


D
Week 1
Week 2
Week 3
Week 4
Gsw (mol m⁻² s⁻¹)
A¹ B¹ A¹
A² A² A²
A³ A³ A³
A⁴ A⁴ A⁴
Drought Determinate Bush
Drought Indeterminate Bush
Drought Indeterminate Climbing
Stress Condition and Growth Habit
E
E apparent (mmol m⁻² s⁻¹)
A¹ A¹ A¹
A² AB² B²
A³ A³ A³
A⁴ A⁴ A⁴
Stress Condition and Growth Habit
F
PhiPS2 (µmol m⁻² s⁻¹)
A¹ A¹ A¹
A² AB² B²
A³ A³ A³
A⁴ A⁴ A⁴
Stress Condition and Growth Habit

3.4.3 Drought stress strategies

Phenotypic data was collected for all accessions under water deficit and, where available, well-watered control conditions. Measurements included harvest information (fresh foliar weight and pod weight), development (BBCH and photoperiod sensitivity) and, porometer and fluorometer data (*E*, g_{sw}, rh_s, VPDleaf, Tleaf, PhiPS2, ETR, Fm, Fs). Collating this phenotypic data identified different drought tolerant strategies and responses, based on plant biomass, photosynthetic activity, reproductive output and developmental rate. The classification used measurements from week 4, as this was the longest period of water deficit before the recovery event. I have classified the responses as 'stay-green', 'saver', 'spender', 'prioritised yield', 'drought susceptible', 'determinate bush control growth' and 'no classification' (Table 5). Each accession was assigned to one category (Table 5).

'Stay-green' accessions had a pod weight below 13g (lowest 75%) and a high foliar weight (top 10%) (Figure 26A and B). This represents accession that maintain their canopy under water deficit. 'Savers' were photoperiod insensitive accessions that did not produce pods under water stress. This included those with a BBCH score below flowering (60) or those which displayed a difference in developmental stage (±10) during the recovery dates, consistent with stopping development under stress.

'Spenders' were accessions that continued development to reproduction and maintained photosynthetic activity under water deficit. The accessions included those which were photoperiod sensitive with pods, accessions with a BBCH greater than 60 (flowering) and those with an ETR (electron transport rate measurement and indicator of photosynthetic efficiency) greater than 90 on week 4, the longest period of water deficit (Figure 26C and D). Accessions classified as 'prioritised yield' had a high reproductive output under stress or a high allocation of resources to the pods. They had a pod weight in the top 10%, a BBCH greater than 60 (flowering) or the proportion of pod weight to foliar weight in the top 5%.

'Drought susceptible' stopped growth under water deficit. The accessions had no pods, a low foliar weight in the bottom 5% and a BBCH below 50 (budding). This group don't maintain their biomass or switch to reproductive group. 'Determinate bush control growth' included accessions that had similar growth and development (BBCH) to the controls under water deficit. They were considered different if there was a BBCH difference greater than ±10, therefore covering a whole developmental stage. Finally, any accessions which did not fit into these categories had 'no classification' (Table 5, Figure 26A-D).

The population structure within the different strategies was also investigated. All categories contained members from both the Andean and Mesoamerican gene pool, except 'stay-green' which only included Andean accessions (Appendix 4).

Table 5; The drought stress strategies within the diversity panel were identified based on phenotypic characteristics from comparative analysis on accessions under water-deficit conditions (142 accessions). Including photoperiod sensitivity and growth habit data (photoperiod insensitive and photoperiod sensitive). Data from week 4 (four weeks since irrigation was removed) was selected as it reflected the maximum length of water deficit before the 'recovery' event.

Drought stress strategy	Phenotypes	Number of accessions	Population Structure
Stay-green	Foliar weight > 120 g (top 10%) & Pod weight < 13 g (bottom 75%)	7	7 Andean
Saver	Photoperiod Insensitive without pods / BBCH < 60 (week 5) (flowering) / Difference between the BBCH on week 4 and week 5 greater than 10 (1 developmental stage)	84	9 Admixed 18 Mesoamerican 57 Andean
Spender	Photoperiod Sensitive with pods / ETR > 90 µmol m^{-2} s^{-1} (week 4) (top 50%) & BBCH > 60 (flowering)	6	3 Mesoamerican 3 Andean
Prioritised yield	Pod weight > 31 g (top 10%) & BBCH > 60 (Flowering) / Proportion pod weight to foliar > 0.195 (top 5%)	19	7 Mesoamerican 12 Andean
Drought-susceptible (growth stopped)	Pod weight = 0 g & Foliar weight < 67 g (bottom 5%) & BBCH < 50	5	1 Admixed 3 Mesoamerican 1 Andean
Determinate bush control growth	Determinate bush accessions which have a difference between control and drought on week 4 greater than 10 (a developmental stage)	15	1 Mesoamerican 14 Andean
No classification	Without further work and controls they cannot be placed in a different category	6	1 Admixed 2 Mesoamerican 3 Andean

To evaluate the accessions responses to the 'recovery' rainfall, the differences between the water-deficit in week 4 (before 'recovery) and during the 'recovery' in week 5 were analysed across multiple traits (Figure 27). Significant differences ($p<0.001$) between week 4 and week 5 were identified for the electron transport rate (ETR), leaf temperature (Tleaf), transpiration (*E*_apparent), quantum yield of Photosystem II (PhiPS2), Fs (steady-state chlorophyll fluorescence), relative humidity (rh_s), g_{sw} (stomatal conductance) and Fm' (maximum fluorescence yield).

The median decreases between week 4 and week 5 for ETR (from 91.9 to 73.7 µ mol m^{-2} s^{-1}) and Tleaf (from 32.4°C to 20.5°C). The median increases between week 4 and week 5 for transpiration (*E*) (from 0.008 to 0.8 mmol m^{-2} s^{-1}), PhiPS2 (0.2 to 0.5 µmol m^{-2} s^{-1}), Fs (from 97.4 to 121.7), humidity (RH) (from 30% to 46.3%), Gsw (g_{sw}) (from 0.0002 to 0.07 mol m^{-2} s^{-1}) and Fm' (119.5 to 259.2). While BBCH has no significant difference from week 4 (39) to week 5 (36) (Figure 27E).

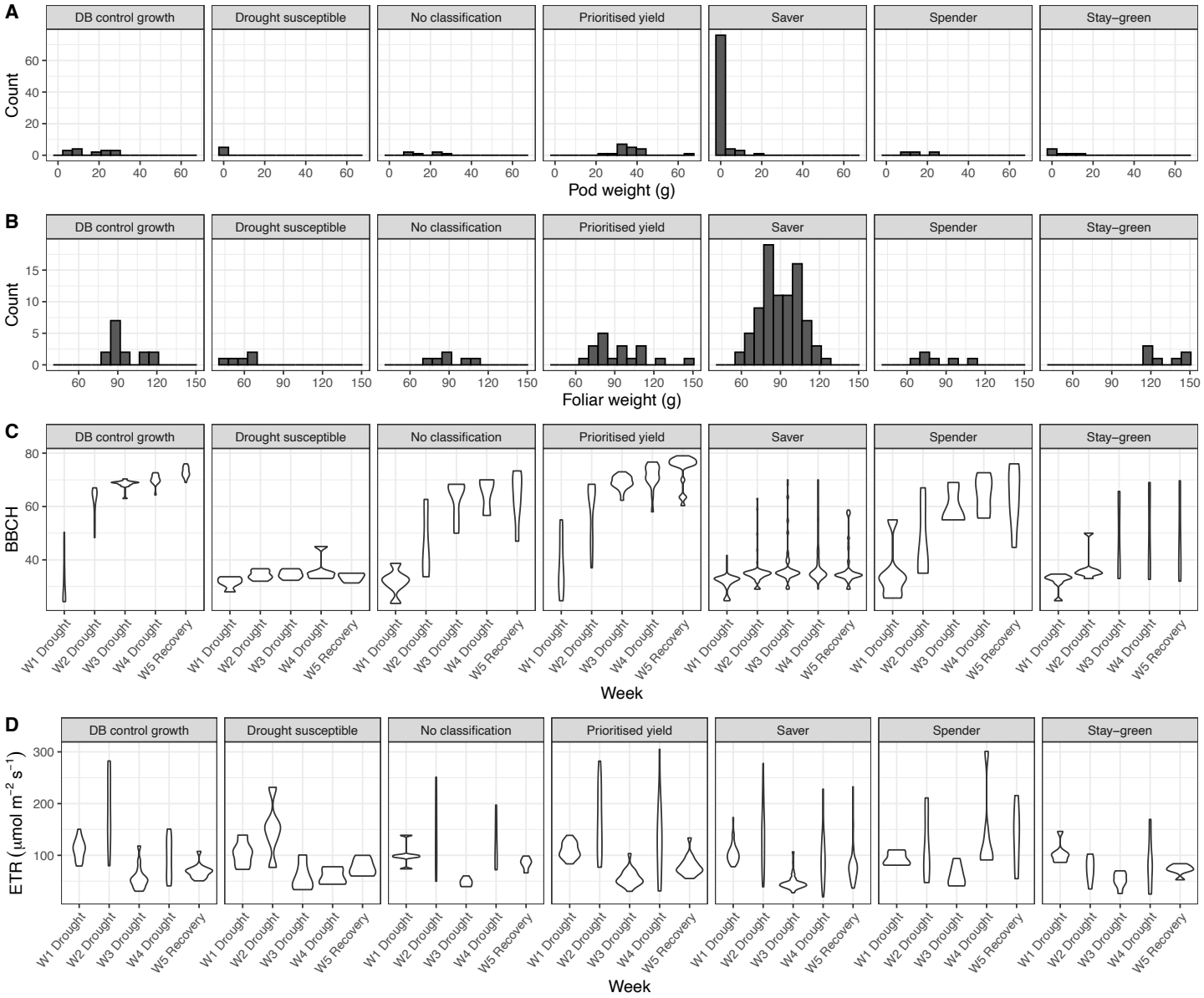


Figure 26; (A, B) Histograms for comparative analysis of the diversity panel undergoing water deficit stress treatment (n=142). Histograms for (A) pod weight (g) and (B) foliar weight (g). (C, D) Violin plots for five weeks of recording water-deficit and recovery (Table 3) for (C) BBCH (development) and (D) ETR (electron transport rate, µmol m^{-2} s^{-1}). Grouped by the drought tolerance strategies identified in Table 5.

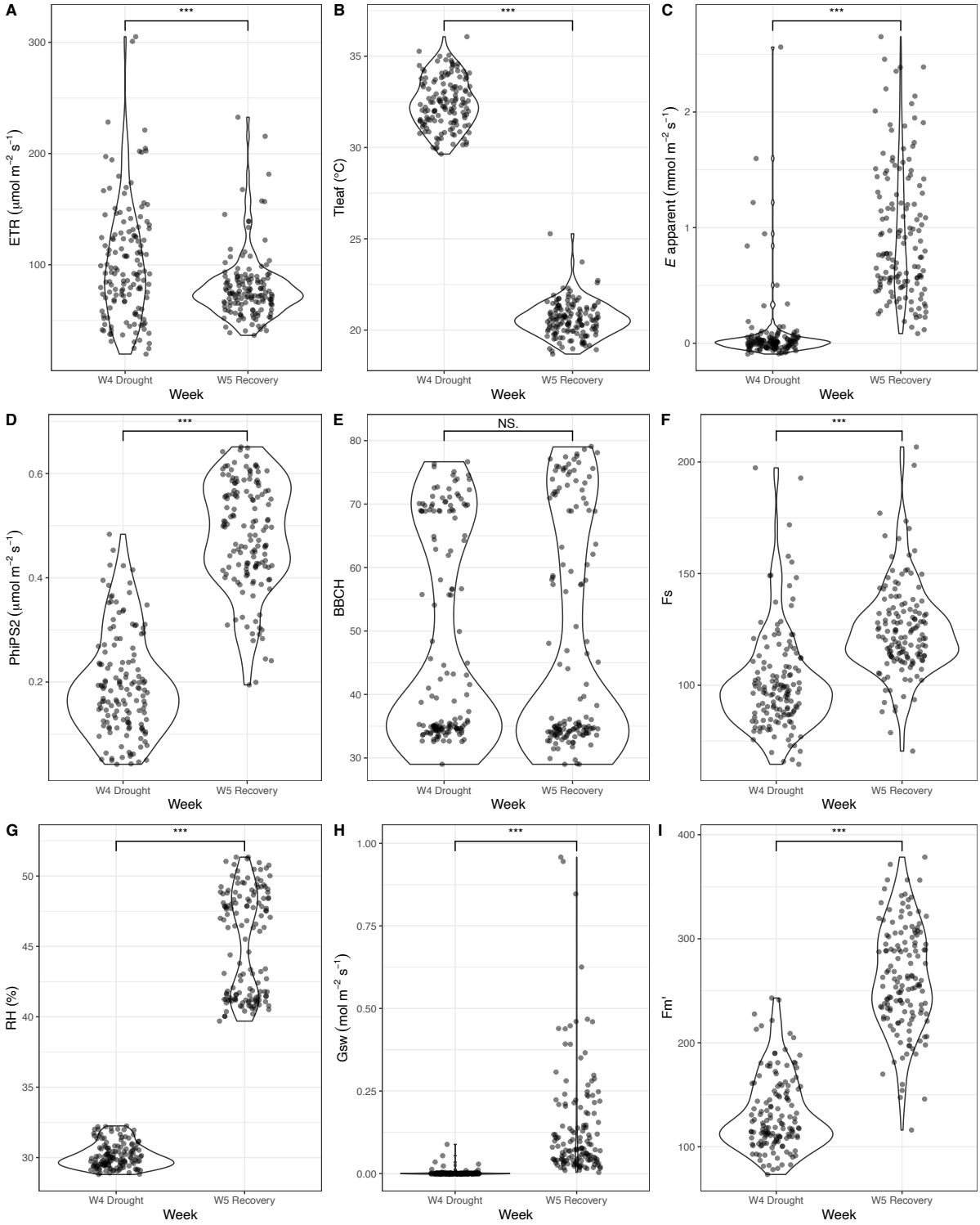


Figure 27; Violin plots comparing leaf porometer and fluorometer responses between water deficit (Week 4) and recovery (Week 5) for all accessions (n=142). (A) ETR (electron transport rate, µmol m^{-2} s^{-1}), (B) Tleaf (leaf temperature, °C), (C) *E* apparent (*E*, transpiration, mmol $m^{-2}s^{-1}$), (D) PhiPS2 (Φ_{PSII}, quantum yield of fluorescence, µmol m^{-2} s^{-1}), (E) BBCH (development), (F) Fs (steady-state chlorophyll fluorescence), (G) RH (relative humidity, %), (H) Gsw (g_{sw}, stomatal conductance, mmol $m^{-2}s^{-1}$) and (I) Fm' (maximum fluorescence yield). Significance calculated with Wilcoxen signed-rank test for non-parametric paired data (Rosner et al. 2006) (NS; not significant, * p <0.05, ** p<0.01, *** p<0.001).

3.4.4 GWAS based on drought responses

The GWAS was performed using the models BLINK, FarmCPU, MLM and MLMM (Figure 28, Table 6) (Huang et al. 2019; Liu et al. 2016; Segura et al. 2012; Wang and Zhang 2021; Zhang et al. 2010).

The QQ plots were analysed to determine the suitability of the model fit (Figure 29), and MTAs were considered significant when $-\log_{10}$(p-value) > 7. A total of 36 QTLs were then selected for being present in more than one phenotype (Table 6). Within these QTLs there were 465 genes (listed in Appendix 6 and 7). For convenience, the QTLs were grouped in categories based on the collection method or complex trait associated; porometer (Figure 28A (g_{SW}) and Figure 28B (E)), development (Figure 28C, harvesting and BBCH), fluorometer (ETR), VPDleaf and Tleaf (Figure 28D), response (Figure 28E (weeks 1 and 2)) and recovery (Figure 28F (weeks 4 and 5)).

Twelve QTLs were found for g_{SW}, on chromosomes Pv01, Pv02, Pv04, Pv05, Pv06, Pv08 and Pv11 (Figure 28A). ETR, VPDleaf and Tleaf were associated with 3 QTLs on chromosomes Pv01 and Pv11. VPDleaf and Tleaf have the same 2 QTLs and are significantly correlated (Figure 24, Figure 28D). *E*_apparent was associated with 8 QTLs on chromosomes Pv02, Pv05, Pv06, Pv08 and Pv09 (Figure 28B). 8 QTLs were found for development (BBCH and harvesting) on chromosomes Pv02, Pv04, Pv07, Pv08, Pv09 and Pv11 (Figure 28C). The 8 QTLs related to drought response (first 2 dates after water application stopped, 01/08/2023 and 10/08/2023) were found on chromosomes Pv02, Pv04, Pv05, Pv06, Pv08 and Pv11 (Figure 28E). Finally, 17 QTLS were identified related to responses within the mock recovery period; these were identified on chromosomes Pv01, Pv02, Pv03, Pv04, Pv05, Pv07, Pv08, Pv09, Pv10 and Pv11 (Figure 28F).

After identifying genes within the QTLs, GO terms related to drought or water stress were searched for. This identified thirteen genes, (Phvul.001G058600, Phvul.002G288700, Phvul.004G007100, Phvul.004G122000, Phvul.005G138000, Phvul.006G005100, Phvul.006G170600, Phvul.006G171100, Phvul.007G167200, Phvul.008G092800, Phvul.008G275700, Phvul.009G192900 and Phvul.011G013900). Of the thirteen genes identified in the GO analysis, four of these genes (Phvul.001G058600, Phvul.004G122000, Phvul.005G138000, Phvul.011G013900) have SNPeff ‘high impact’ mutations corresponding to frameshift mutations (Cingolani et al. 2012).

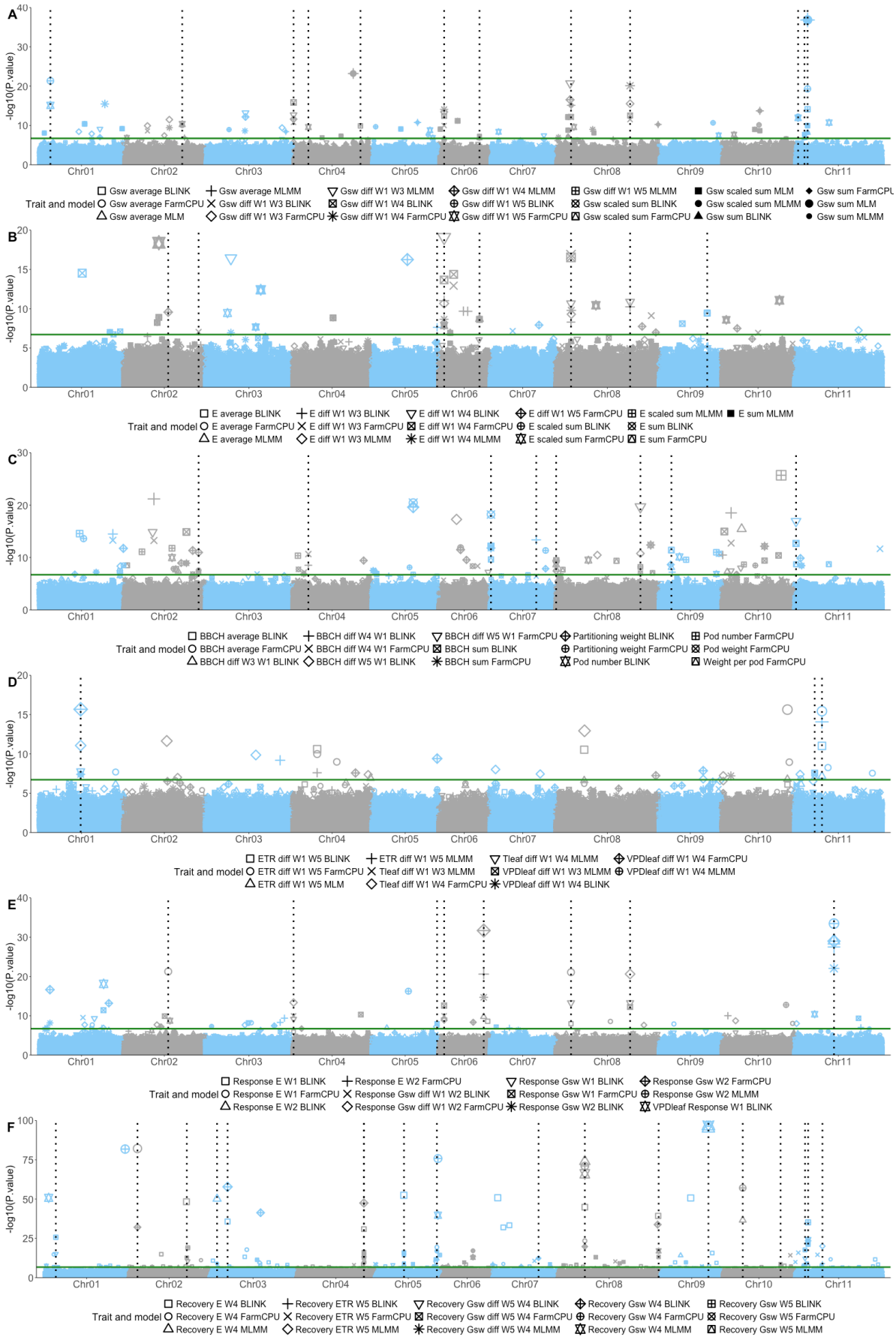
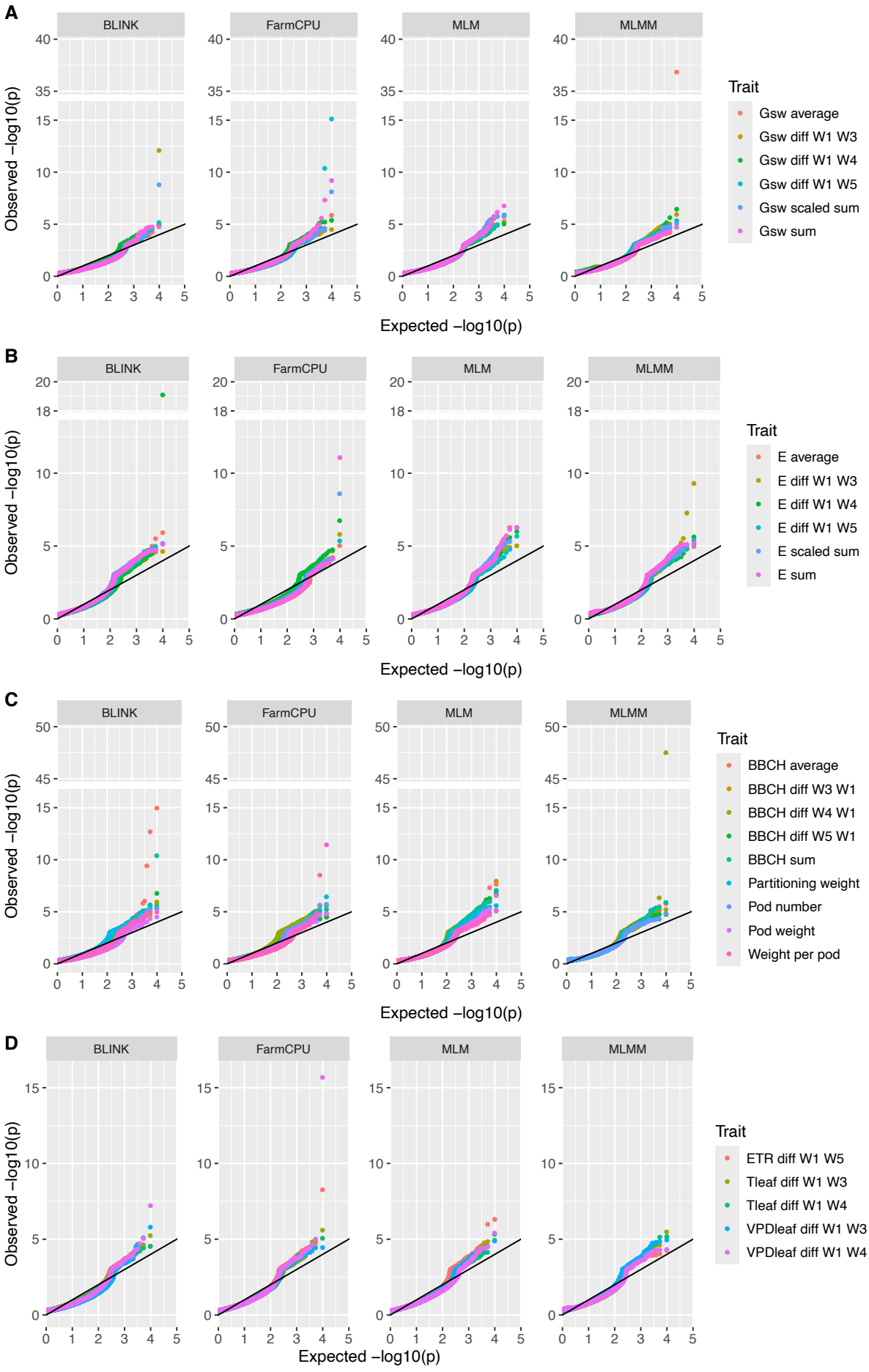


Figure 28; Manhattan plots highlighting markers significantly associated with the traits; (A) Gsw (g_{sw}, stomatal conductance, porometer), (B) *E* (transpiration, porometer), (C) development

(BBCH and harvesting), (D) ETR (electron transport rate, fluorometer), VPDleaf (leaf vapour pressure deficit) and Tleaf (leaf temperature), (E) 'responses' to drought (weeks 1 and 2) and (F) 'recovery' (weeks 4 and 5). The difference (diff) was calculated to see how the traits changed over time and to see differences in the magnitude of change between weeks (Table 3). Week 1 (W1), week 2 (W2), week 3 (W3), week 4 (W4) and week 5 (W5). The analyses were completed with GAPIT and the models FarmCPU, BLINK, MLMM or MLM on 142 accessions (Huang et al. 2019; Liu et al. 2016; Segura et al. 2012; Wang and Zhang 2021; Zhang et al. 2010). The X-axis represents the genomic position of markers, and the Y-axis is the -log 10 of the P-values for association with the phenotype. The vertical lines correspond to QTLs associated with at least two phenotypes. Point size correlates to -log10(P-value). Quantile-quantile (QQ) plots are provided in Figure 29.


A
BLINK
FarmCPU
MLM
MLMM
Observed −log10(p)
Expected −log10(p)
Trait
Gsw average
Gsw diff W1 W3
Gsw diff W1 W4
Gsw diff W1 W5
Gsw scaled sum
Gsw sum
B
BLINK
FarmCPU
MLM
MLMM
Observed −log10(p)
Expected −log10(p)
Trait
E average
E diff W1 W3
E diff W1 W4
E diff W1 W5
E scaled sum
E sum
C
BLINK
FarmCPU
MLM
MLMM
Observed −log10(p)
Expected −log10(p)
Trait
BBCH average
BBCH diff W3 W1
BBCH diff W4 W1
BBCH diff W5 W1
BBCH sum
Partitioning weight
Pod number
Pod weight
Weight per pod
D
BLINK
FarmCPU
MLM
MLMM
Observed −log10(p)
Expected −log10(p)
Trait
ETR diff W1 W5
Tleaf diff W1 W3
Tleaf diff W1 W4
VPDleaf diff W1 W3
VPDleaf diff W1 W4

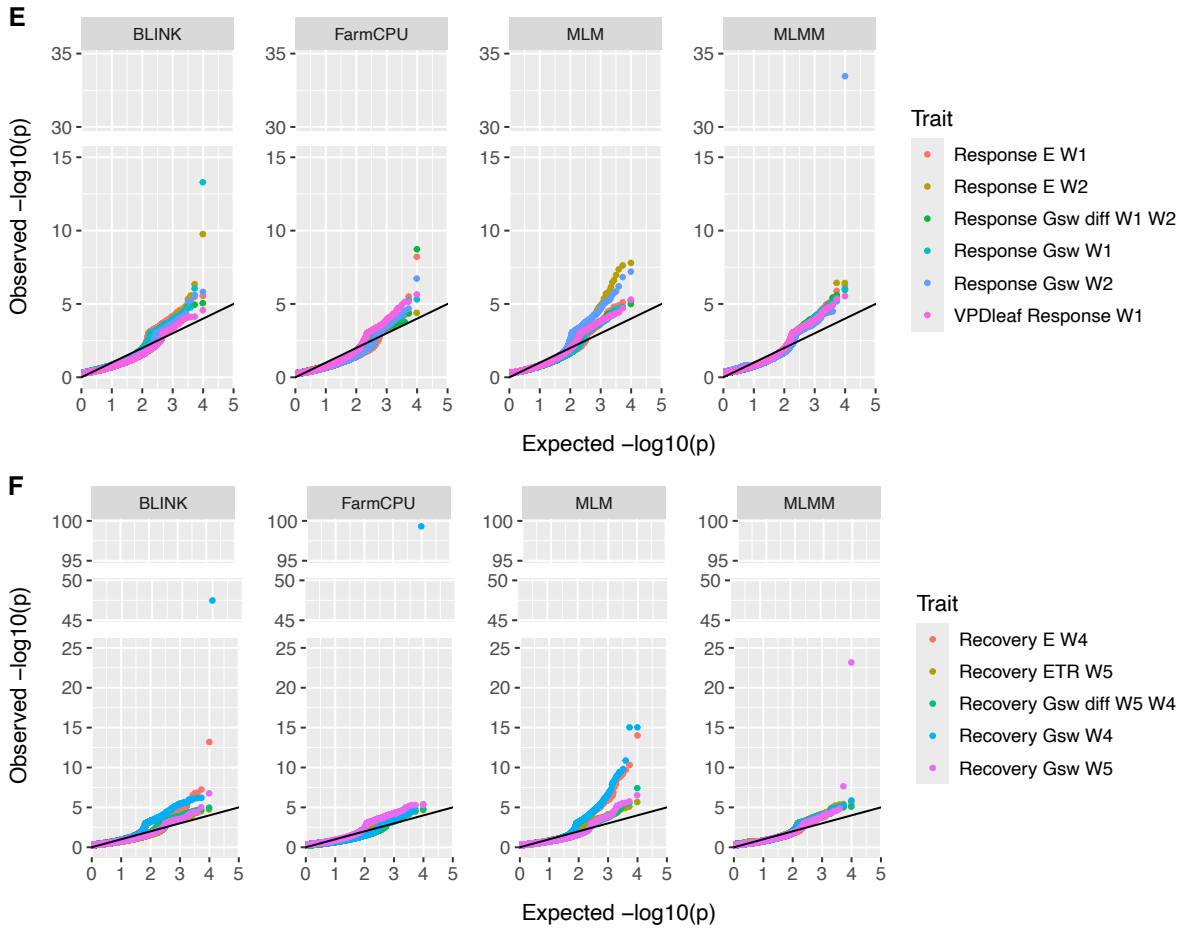


Figure 29; Quantile-quantile (QQ) plots are provided for Manhattan plots in Figure 28. (A) Gsw (g_{sw}, stomatal conductance, porometer), (B) *E* (transpiration, porometer), (C) development (BBCH and harvesting), (D) ETR (electron transport rate, fluorometer), VPDleaf (leaf vapour pressure deficit) and Tleaf (leaf temperature), (E) 'responses' to drought (weeks 1 and 2) and (F) 'recovery' (weeks 4 and 5). The difference (diff) was calculated to see how the traits changed over time and to see differences in the magnitude of change between weeks (Table 3). Week 1 (W1), week 2 (W2), week 3 (W3), week 4 (W4) and week 5 (W5). The analyses were completed with GAPIT and the models FarmCPU, BLINK, MLMM or MLM on 142 accessions (Huang et al. 2019; Liu et al. 2016; Segura et al. 2012; Wang and Zhang 2021; Zhang et al. 2010). The X-axis represents the expected -log 10 of the P-values and the Y-axis is the observed -log 10 of the P-values for association with the phenotype. The Y-axis in plots A-C and E-F have breaks to improve visualisation of the distribution and extreme values.

Table 6; The 36 QTLs for g_{sw} (porometer), *E* apparent (porometer), development (BBCH and yield), ETR (electron transport rate, fluorometer), VPDLeaf (leaf vapour pressure deficit), Tleaf (leaf temperature), water deficit 'response' (weeks 1 and 2) and 'recovery' from water deficit (weeks 4 and 5).

Name	Chromosome	Start (bp)	End (bp)	Traits
Dr1.1	Chr01	6,671,000	6,880,000	g_{sw} Recovery g_{sw}
Dr1.2	Chr01	25,129,000	25,338,000	Tleaf VPDleaf
Dr2.1	Chr02	5,343,000	5,543,000	Recovery g_{sw} Recovery Transpiration €
Dr2.2	Chr02	27,067,000	27,276,000	Transpiration Response Transpiration (*E*)
Dr2.3	Chr02	35,612,000	35,821,000	g_{sw} Recovery g_{sw}
Dr2.4	Chr02	45,604,000	45,813,000	Partitioning Weight Transpiration (*E*)
Dr3.1	Chr03	4,510,000	4,719,000	Recovery g_{sw} Recovery Transpiration (*E*)
Dr3.2	Chr03	10,951,000	11,160,000	Recovery g_{sw} Recovery Transpiration (*E*)
Dr4.1	Chr04	401,000	610,000	g_{sw} Response g_{sw}
Dr4.2	Chr04	9,414,000	9,623,000	BBCH g_{sw}
Dr4.3	Chr04	41,187,000	41,396,000	g_{sw} Recovery g_{sw}
Dr5.1	Chr05	17,733,000	17,941,000	g_{sw} Recovery g_{sw} Recovery Transpiration (*E*)
Dr5.2	Chr05	37,918,000	38,127,000	Recovery g_{sw} Recovery Transpiration (*E*)
Dr5.3	Chr05	39,838,000	40,047,000	Transpiration (*E*) Response g_{sw} VPDleaf

Dr6.1	Chr06	3,218,000	3,427,000	g_{sw} Response g_{sw} Transpiration (*E*)
Dr6.2	Chr06	24,785,000	24,994,000	g_{sw} Transpiration (*E*)
Dr6.3	Chr06	27,359,000	27,568,000	Response g_{sw} Response Transpiration (*E*)
Dr7.1	Chr07	541,000	750,000	BBCH Pod Weight
Dr7.2	Chr07	28,150,000	28,359,000	BBCH Recovery g_{sw}
Dr8.1	Chr08	234,000	443,000	BBCH Pod Number
Dr8.2	Chr08	9,293,000	9,502,000	Transpiration (E*)* g_{sw} Response g_{sw} Response Transpiration (*E*)
Dr8.3	Chr08	16,512,000	16,721,000	Recovery g_{sw} Recovery Transpiration (*E*)
Dr8.4	Chr08	45,321,000	45,530,000	Transpiration (E) g_{sw} Response g_{sw}
Dr8.5	Chr08	51,625,000	51,834,000	BBCH Pod Number Pod Weight Partitioning Weight
Dr8.6	Chr08	61,805,000	62,014,000	Recovery g_{sw} Recovery Transpiration (*E*)
Dr9.1	Chr09	7,448,000	7,657,000	BBCH Pod Weight
Dr9.2	Chr09	29,293,000	29,502,000	Transpiration (*E*) Recovery g_{sw} Recovery Transpiration (*E*)
Dr10.1	Chr10	12,119,000	12,328,000	Recovery g_{sw} Recovery Transpiration (*E*)

Dr10.2	Chr10	35,366,000	35,557,000	Recovery g_{sw} Recovery Transpiration (*E*)
Dr11.1	Chr11	898,000	1,107,000	BBCH Pod Weight
Dr11.2	Chr11	6,038,000	6,247,000	g_{sw} Recovery g_{sw}
Dr11.3	Chr11	7,934,000	8,143,000	g_{sw} Recovery g_{sw} Recovery Transpiration (*E*)
Dr11.4	Chr11	12,256,000	12,465,000	VPDleaf Tleaf Response VPDleaf
Dr11.5	Chr11	16,712,000	16,921,000	ETR Recovery ETR
Dr11.6	Chr11	24,038,000	24,247,000	Response g_{sw} Response Transpiration (*E*)

3.5 Discussion

3.5.1 Indicators of drought stress

This study is comparative between accessions experiencing water deficit. The soil moisture sensors provided a quantitative measure of the level of water stress the plants experienced. In line with previously published water deficit experiments, a soil matric potential of -800 kPa or less, and a water content of less than 0.1 m^3/m^3 are sufficient for drought stress (Figure 23) (Ganesan et al. 2024; Gebregiorgis and Savage 2006; Lui and Mihara 2024; Medynska-Juraszek et al. 2021; Yang et al. 2024). Consequently, the limited water received by the accessions in the water deficit experiment failed to meet the demands of the plants for normal growth and development (Dramadri et al. 2019; Wortmann et al. 1998).

In the later weeks of the trial, there was no statistical difference in the water content between the stressed determinate bush, indeterminate bush or indeterminate climbing groups (Figure 23D). This supports that there was not a significant difference in the water content experienced by the different growth habits under water deficit. However, the water content and Bulk EC at

week 2 were significantly different between the growth habits experiencing water deficit (Figure 23C and D). Indeterminate and determinate accessions have divergent behaviours such as in vegetative growth, development and water conductance, suggesting that the observed differences between the growth habits were due to the indeterminate cultivars accessing more water from the soil (Campos et al. 2021).

Leaf porometry and fluorometry also provided quantitative measures of the levels of water deficit that the plants experienced (Figure 25). Overtime, the Fs (steady state fluorescence) and Fm' (maximum fluorescence yield) decreased from week 1 to week 4, indicative of photoinhibition in response to water deficit (Guidi et al. 2019; Jat et al. 2024; Ramírez-Estrada et al. 2023). Also, higher leaf temperatures (Tleaf) under drought negatively correlate with other leaf traits, reducing stomatal conductance (g_{sw}) and photosynthetic efficiency (e.g. lower ETR, and Fm', Figure 24 and Figure 25). These responses are indicative of the onset of water deficit stress during this trial. Further to this, the work supports the use of leaf temperature as a proxy for stomatal conductance under drought stress, a phenomenon exploited in canopy-level infra-red thermal imaging (Driever et al. 2023; Yu et al. 2015).

Conclusions can also be drawn from the correlation plot between phenotypes (Figure 24). BBCH, a score of the phenological development of the plant, has a strong positive correlation with the harvest information and negatively correlates with stomatal and photosynthetic traits (Cavalcante et al. 2020). This is likely because the BBCH scale correlates with pod production, while maturation under drought negatively impacts stomatal conductance and transpiration (*gsw* and *E*_apparent, respectively) (Jahan et al. 2023; Müllers et al. 2022; Urban and Urban 2024).

During this work, water-deficit was experienced in the absence of other stresses such as heat stress. The soil temperature stayed constant throughout the experiment and was never significantly different between control or the water-deficit treatments, and the air temperature stayed in the range of 12 to 25°C (Figure 22 and Figure 23). However, combinatory stress is common in nature, and the stresses are more frequently overlapping under climate change (Sato et al. 2024).

Another consideration is that this trial took place in pots rather than in the field, to manage water leakage between accessions. This limitation was controlled across the growth habits and treatments but is likely to have impacted the root structure of the plants. Certain cultivars have shown phenotypic plasticity in root growth under water deficit conditions, sometimes

developing more extensive root systems. Indeterminate types, which generally produce larger root systems than determinate types, may therefore have been more strongly affected inside the pots (Cerutti et al. 2023; Velho et al. 2018).

3.5.2 Drought stress strategies

The panel contains both the Andean and Mesoamerican gene pools. Differences in susceptibility to drought among the two domestication pools had already been observed (Vicente-Serrano et al. 2020). Within the diversity panel, multiple different drought tolerance strategies were found. Classified as 'stay-green', 'saver', 'spender', 'prioritised yield', 'drought susceptible', 'Determinate bush control growth' and 'no classification' (Table 5, Figure 26) in this study. However, many of the strategies have developed in both gene pools, likely through local adaptation to the ecogeographic environment.

Interestingly, there is a correlation between the major seed colour and water deficit strategies in this panel, which has been identified previously in literature with common beans (Hussaini et al. 2021). This is possibly indirectly selected for by the market preferences of a country, so the colour is associated with the agro-ecological conditions in which the varieties are selected for and maintained. Similar to associations observed between seed colour, seed weight and seed size within gene pools (Giordani et al. 2022). Further to this the seed colour can be a proxy for seed coat traits, such as the concentration of flavonoids and phenolics, but also seedling growth and germination (Kavas et al. 2022; Vidak et al. 2022). These compounds have in some cases been linked to stress resistance, including drought tolerance (Rao and Zheng 2025; Yaneva et al. 2024). Further to this, the country of origin showed an association with the drought tolerance strategy (Figure 24) however, due to the movement of accessions and the population structure this would need further investigation to see if this is a true correlation with drought tolerance strategies.

The panel contains accessions that are 'drought susceptible', as they had low pod and foliar weights following four weeks of water deficit. Common beans are renowned for experiencing severe yield losses under drought stress, hence the need to identify those genotypes with different drought tolerance strategies (Labastida et al. 2023; Rosales et al. 2012). 'Savers' (isohydric) are plants which reduce gas exchange and photosynthesis under drought stress, due to closing their stomata earlier (Polania et al. 2016). These accessions pause development under drought tolerance or abort pod development to wait for better conditions which are more

appropriate for reproductive development. In the panel, this was evidenced by the BBCH pausing during water-deficit, therefore not reaching pod development and, in some cases, development resuming during the 'recovery'. Without the re-addition of water, these accessions tend to have lower seed yields (Hamabwe et al. 2024). This strategy is best for longer-term mild droughts, when water will return to allow the plant to produce yields rather than focusing on vegetative biomass (Bandurska 2022; Nesporová et al. 2024). A large proportion of the panel are 'savers' suggesting, this strategy is better adapted for the ecogeographic environment they were collected from (Arregocés et al. 2025).

'Spenders' (anisohydric) are plants that are fast growing under drought stress conditions. They maintain gas exchange and photosynthesis, allowing growth under stress conditions in order to avoid the onset of worsening conditions. In the phenotyping, this is evidenced by their continuing development (increasing BBCH), combined with high transpiration and photosynthesis rates. The spending strategy is best for producing yields under long-term drought stress, capitalising on the prevailing conditions before more unfavourable conditions set in, and prioritising biomass (Delfin et al. 2021; Nesporová et al. 2024). If the accessions are able to complete their life cycle before the conditions get worse, then this strategy is best for longer-term drought stress when precipitation may not return, similar to drought escaping (Shavrukov et al. 2017). However, if the lifecycle is not completed, then this strategy may result in total crop failure, and is best for milder, shorter droughts, when water will return to continue development. A subset of these accessions were those which 'prioritised yield', which under drought stress, maintained yield and grain development (Polania et al. 2016). These accessions have seed yields in the top 10% under the water deficit conditions, so could be determined 'drought tolerant' (Hamabwe et al. 2024). Those with high reproductive output have reduced levels of abortion under drought stress and maintain fertilisation.

Other strategies included 'stay-green'. In this panel, the 'stay-green' accessions were Andean, however they have been previously identified in other common bean populations (Labastida et al. 2023; Schmit et al. 2019). Our panel is skewed towards the Andean gene pool, consequently it is possible that our genetic diversity didn't cover the Mesoamerican 'stay-green' genetic diversity. 'Stay-green' accessions maintain photosynthesis under drought stress, and under certain droughts, may alleviate yield loss. In this trial, this strategy was identified at harvest time in accessions with very high foliar weight but very low or absent pod weight. 'Stay-green' plants can also be categorised as delaying senescence and chlorophyll degradation under drought stress, so that when water returns, they can quickly continue development (Thomas and Ougham 2014). This strategy is best for shorter droughts when water will return quickly. In other

crops the ‘stay-green’ strategy has been linked to higher yields under drought stress, and therefore further investigation is needed in common beans to understand the yield of the stay-green accessions after drought stress and during a recovery, comparative to controls (Kamal et al. 2019; Padilla-Chacón et al. 2019).

These different strategies have developed for, and are beneficial under, different agro-ecogeographic conditions and farming methods (Blum 2015; Polania et al. 2016). All strategies are beneficial and of interest for breeding programmes depending on the type of drought prevalent in a region.

During this study, I found that using developmental information was useful to distinguish strategies, and this could be complimented with the porometer and fluorometer data to determine levels of stress. Porometer and fluorometer associated traits (Tleaf, g_{sw}, E_apparent) indicate the level of water stress the plants were experiencing. Accessions with the highest scores for photosynthesis (top 10%) included those with a variety of strategies, as well as those which were ‘drought susceptible'. These trait evaluations alone were unable to differentiate the different drought tolerance strategies without the addition of developmental information.

This panel includes examples of various strategies for drought tolerance. However, the assignment of these strategies remains speculative, and would require further experimental work and comparisons under contrasting drought and environmental scenarios. Additionally, control groups for the indeterminate accessions would help to improve certainty and progress the investigation into drought tolerance beyond comparative analysis.

Finally, phenotypic analyses could be expanded to include root traits, as research links root phenotypes to strategy (Polania et al. 2017a). Roots are often the first organ to respond to drought stress in plants, however due to difficulties with sampling and observation, foliar organs are often investigated instead (El Bey et al. 2024; Irshad et al. 2024). Root architecture traits (root branching and depth) are often correlated with drought tolerance strategies in common beans therefore, exploring this could uncover further pathways (Polania et al. 2017b). Also, drought stress can impact the symbiotic relationship with rhizobia, affecting nodulation, nitrogen fixation, carbon allocation and gene expression (da Silva et al. 2024; Fenta et al. 2020). Direct analysis of epidermal traits, including stomatal density (Egesa et al. 2024) and size, may also be beneficial (Polania et al. 2022). These could further complement the other above-ground traits, which were proven for estimating drought responses, including the stay green trait, rates of photosynthesis, transpiration and reproductive output (Sofi et al. 2021).

A multi-faceted approach will be best to combat the effects of global drought stress on common bean yields. Incorporating government policies that discuss how to utilise water resources more efficiently (Li et al. 2020), new irrigation methods (such as exploring drip irrigation (Barezzi et al. 2024)), soil supplements (such as coconut charcoal (Lui and Mihara 2024)) and breeding programmes for beneficial drought tolerant traits.

3.5.3 Recovery from water deficit

A separate trait to investigate is recovery from drought stress. How crop plants recover photosynthesis and growth after water deficit affects their ability to continue normal development and, consequently, yields (Wang et al. 2019). Recovery is an important trait to investigate, as drought is often intermittent or can be alleviated through irrigation, and other water management practices. Where drought is short-lived, drought recovery, plays a significant role in yields (Delfin et al. 2021).

Within this study, the differences in accessions under drought stress, and then following a timely precipitation event were investigated. This event allowed some of the plants to start recovering from the drought stress. The large differences between the week 4 and week 5 suggest that the re-addition of water instigated a quick recovery response, photosynthetically (Figure 27).

One of the traits that did not significantly change between these dates (week 4 and week 5) was plant development, as defined by BBCH (Figure 27E). This is likely because the accessions that had not flowered did not have enough time with the re-addition of water to significantly accelerate their development, such as to flowering. Those plants that had flowered during the water deficit had already committed to producing pods or maturing.

However, this recovery period was not sufficiently long or extensive enough to see an impact on yield; only the shorter-term physiological changes triggered by the recovery, were investigated. During future work, a comparison of the effects of intermittent drought, recovery, and terminal drought stress would further clarify their relationships. In such a scenario, the role of plant stress ‘memory’ (priming) and how it affects responses to future stresses will be similarly important (Jacques et al. 2021).

3.5.4 Candidate genes

By leveraging this diversity panel, 36 QTLs were found associated with responses and recovery to water deficit (Figure 28, Table 6). Within these QTLs there were 465 genes (listed in Appendix 6 and 7). To prioritise causative genes within QTLs, I followed the criteria that I previously tested in Chapter 2, and published in Denning-James et al. (2025). This consisted of focusing on genes with 'high impact' mutations in the coding region (mostly frame shifting, where an insertion or deletion changes the translation) or associated to 'drought related' or 'abiotic stress related' GO terms.

Phvul.001G058600 (QTL Dr1.1) is related to *chloroplastic chaperone activity of the BC1 complex* (*CABC1)* (Goodstein et al. 2012), and may modulate chlorophyll degradation in response to oxidative stress, including salt stress (Borkiewicz et al. 2020; Huala et al. 2001; Qin et al. 2020). Phvul.004G122000 (Dr4.3) may encode a dehydration response element binding protein (*DREB1*) (Goodstein et al. 2012). DREBs are found across the plant kingdom (Han et al. 2022b) (rice (Wang et al. 2022), wheat (Mei et al. 2022), tomato (Tao et al. 2022)). Another DREB gene, Phvul.008G092800 (orthologous to *DREB2*) was also found within the QTLs (Dr8.2). Both *DREB1* and *DREB2* are involved in responses to abiotic stresses including heat, drought and salt (Akbudak et al. 2018; Akhtar et al. 2012; Guttikonda et al. 2014).

Other identified genes are associated with drought stress. Phvul.005G138000 (Dr5.2) has homology to *Arabidopsis* AT3G45660, a member of the NAXT NPF subfamily which is upregulated under drought and involved in xylem transport in response to osmotic stress (Huala et al. 2001; Li et al. 2016; Li et al. 2010). Phvul.011G013900 (Dr11.1) is likely to encode a calcium transporting ATPase (Goodstein et al. 2012; Huala et al. 2001). Homology with *Arabidopsis* gene AT3G57330.1 suggests associations with drought and salt tolerance possibly via the control of guard cell physiology (He et al. 2024; Su et al. 2024).

Transcription factor-encoding genes in the QTLs include *SHINE* orthologue Phvul.010G092300, an APETALA2/ Ethylene Responsive Factor (AP2/ERF) controlling waxy cuticle and stomatal development (Girón-Ramírez et al. 2021; Khoudi 2023), MYBs (Phvul.004G121500, Phvul.009G192600, Phvul.011G084500), and Phvul.008G275300, an orthologue of *WRKY20* known to modulate guard cell ABA and ROS signalling as well as cuticular wax biosynthesis in response to drought (Li et al. 2023a; Luo et al. 2013; Rushton et al. 2012). ERFs, MYBs, WRKY and protein kinases are genes which are involved in responses to drought stress and drought tolerance by interacting with downstream genes in signalling pathways (Umezawa et al. 2006; Wu et al. 2024).

Genes related with the phytohormones auxin (Phvul.005G172900, Phvul.005G173000, Phvul.006G142200, Phvul.006G142300) (Liu et al. 2019a; Wang et al. 2021; Yin et al. 2023) and gibberellic acid (Phvul.002G193300, Phvul.005G173500, Phvul.007G167600) were also identified (Chu et al. 2022). Phvul.007G167600 is an ortholog of *GIBBERELLIN-INSENSITIVE DWARF 1* (*GID1*), a partially-ABA-dependent GA receptor that regulates stomatal development and ABA biosynthesis under drought stress in rice (Du et al. 2015). Other ABA-associated genes identified in these pathways include PPRs (Phvul.001G059500, Phvul.003G071600, Phvul.008G277600, Phvul.009G034500, Phvul.011G068200, Phvul.011G069000) including a DYW subgroup PPR gene, (Phvul.011G107600) linked to soybean drought responses (Su et al. 2019). Similarly, homologues to protein kinase-encoding genes involved in drought and ROS signalling were identified (Phvul.002G289000, Phvul.005G173100, Phvul.008G093200, and Phvul.008G124300) (Cai et al. 2025; Liu et al. 2024; Liu et al. 2020; Wu et al. 2024; Xu et al. 2018). Enhanced photosynthetic efficiency under drought may arise from superior responses to oxidative stress (ROS). Together with the transcription factors, crosstalk between auxin, gibberellic acid and ABA regulate an array of signalling components in downstream pathways critical to water deficit responses (Aslam et al. 2022; da Silva et al. 2024; Labastida et al. 2023; Salehin et al. 2019; Su et al. 2019; Verma et al. 2022).

The candidate genes identified in this diversity panel show promising relevance to drought stress resilience. Nevertheless, further functional experimental analyses are necessary to confirm whether these putative genes play similar roles in common beans as their orthologues in other crops such as soybean and Arabidopsis. Also, the QTLs are still relatively large, covering many genes therefore, future work could include fine-mapping and higher-resolution mapping populations to aid in distinguishing causal genes from the candidates. Functional validation approaches, such as gene silencing or gene expression analyses, could then be used on these genes (Azizi et al. 2020; Gutierrez and Torres 2025). As the porometry and fluorometry analyses selected the largest proportion of QTLs, combining these approaches in future studies will be important to identify differences in stomatal and photosynthetic drought responses alongside developmental datasets.

3.6 Conclusions

Altogether, this study highlights the power of combining both phenological and physiological analyses when investigating the diversity of plant drought responses. By measuring developmental (long term responses) and short-term instantaneous responses to water deficit, both divergent strategies were identified in response to drought as well as a broad range of candidate genes underlying those responses. Leveraging large-scale trait data with large scale populations not only provides novel insights into the mechanisms of plant drought response, but I believe will provide new breeding targets for future sustainable agriculture under climate change.

Chapter 4

Future work and Further Discussion

4.1 Key findings chapter 2

During my PhD, I showed that the diversity panel contains genetic diversity from both the Andean (four subgroups) and Mesoamerican (two subgroups) gene pools, including Colombian accessions that display introgressive hybridisation and admixture. This indicates a large amount of genetic diversity with fewer geographic effects.

Variation in domestication traits, such as determinacy and photoperiod insensitivity were found to be prevalent in both gene pools. Genomic regions associated with these traits included both known and novel putative candidate genes. Many of which were identified due to homology with other species such as *Glycine soja, Vigna species* and *Arabidopsis*. These findings contribute to current literature and have practical relevance to breeding programmes. This is because determinacy and photoperiod insensitivity are pleiotropic traits, under linkage drag, they are important for yield stability and stress resistance under different ecogeographic conditions, and they have facilitated the dissemination of common beans.

Introgressions between the gene pools, alongside the mean allele dosage in regions of the genomes, support frequent gene flow (low isolation). They can help to infer private allelic diversity and mutations within the population structure or link to phenotypes, including functional variation at a SNP level (non-synonymous substitutions). This work also highlights the importance of GWAS for identifying MTAs, especially when properly accounting for population structure. By linking candidate genes to phenotypes, hopefully more targeted precision breeding approaches can be adopted to improve common bean traits under climate change. Future work should include functional validation in multiple *Phaseolus* genomes (pangenomes) and exploring protein modelling to determine how the SNPs or indels impact protein function or folding. All the data generated during my PhD is available and can be utilised in future work to investigate gene function and then be utilised in breeding programmes. This understanding is important for the conservation, management and utilisation of the genetic resources through breeding programs.

The diversity panel also included two weedy accessions; these may contain interesting novel diversity and population structure. Providing a path for trait introgressions between crop wild relatives and landraces as evidenced by the improvement of drought tolerance in a CIAT breeding programme crossing tepary bean with common bean. Possibly making genetic diversity in wild relatives more accessible in future breeding programmes. However, further research is needed into weedy accessions to determine their genetic diversity and phenotypes.

4.2 Key findings chapter 3

The diversity panel was also investigated under water deficit, using porometry, fluorometry and developmental datasets to capture complex genotype responses. Across the different subpopulations, multiple drought tolerance strategies were found. These strategies did not correspond directly to population structure, reflecting the local adaptation to a wide range of ecogeographic conditions in Central and South America.

Different strategies are better suited to different environments and production regimes (small holders vs industrial), which is an important consideration for bean breeding programmes. Drought is a complex trait and strongly impacts common bean yields. The PhD expands current knowledge on drought tolerance strategies in common beans and putative candidate genes. Accessing the genetic diversity in landraces with the 'correct' drought strategy, could improve yields or yield stability under drought stress in commercial varieties through breeding programmes.

QTLs were identified for the different evaluated responses, as well as traits associated with recovery from water deficit (despite not being a trait initially planned for), in common beans to understand how the common beans respond to drought stress. This work could be taken further as in chapter 2, linking the genes to introgressions, SNPs or indels; and also, by functionally validating the candidate genes to understand the genetic basis for the strategies.

The work is mainly comparative, with only controls for determinate bush accessions, due to space limitations. To build on this, another study could also include controls for the indeterminate growth habits to directly resolve differences between control growth and water deficit development. Future, field trials could add to the data from this study by not restricting root development, to better understand whole-plant responses. In addition, future work could

explore combinatory stress, such as drought and heat, which commonly co-occur in real production systems, particularly under climate change.

Altogether, this study highlights the power of combining both phenological and physiological analyses when investigating the diversity of plant drought responses. By measuring developmental (long-term responses) and instantaneous responses to water deficit, both divergent strategies were identified in response to drought and a broad range of candidate genes underlying those responses. Leveraging large-scale trait data with large scale populations not only provides novel insights into the mechanisms of plant drought response, but it is hoped to provide new breeding targets for future sustainable agriculture under climate change.

4.3 Future Implications

This work highlighted the genetic diversity within common bean accessions, their domestication traits, and differing drought tolerance strategies in response to water deficit. This work can be integrated, using data on determinacy and photoperiod sensitivity with drought associated loci. Allowing for breeding programmes to breed common beans that are climate-resilient and adaptable across different ecogeographic environments. This genetic diversity would safeguard yields under different climate change scenarios.

Although this work provides important insights, there are several limitations. First, the reference genome may introduce bias, potentially leading to missing or misrepresented variation, particularly in 'uncalled' regions. Second, GWAS analyses are constrained by population structure and sample size, which can reduce statistical power and increase the likelihood of false positives or undetected loci. Additionally, candidate gene identification relied heavily on homology with model species such as Arabidopsis thaliana, which may not fully capture gene function in common bean.

In Chapter 3, phenotyping was conducted in pots, restricting root development and limiting the ability to capture natural plastic responses to water deficit. The absence of control groups across all growth habits further constrained comparisons. Also, drought stress was applied in isolation, whereas in field conditions multiple stresses often co-occur. Finally, while computational approaches enabled the identification of numerous candidate genes, functional

validation was beyond the scope of this study, limiting the ability to confirm causal relationships between genotype and phenotype.

In the future, short term aims should include prioritising candidate genes for future work using integrative approaches, including transcriptomics, protein modelling, and machine learning-based prediction of gene function. While this study identified a large number of candidate genes, experimental validation remains a bottleneck. The application of machine learning approaches to predict gene function and phenotype could help reduce candidate gene lists prior to validation. Targeted validation of key loci using gene expression analyses or genome editing approaches such as CRISPR/Cas9 will also help to confirm their roles in domestication and drought response.

Medium term aims could include expanding genomic resources through the development of Phaseolus vulgaris pangenomes, alongside utilising long-read sequencing technologies, to improve the resolution of structural variation and 'uncalled' regions. This sequence diversity panel can be utilised alongside other whole-genome sequenced common bean accessions to improve prediction capabilities and better understand genomic diversity.

My PhD has also highlighted the importance of genebanks in conserving genetic diversity ex situ. I investigated the genomes of 144 common bean accessions, yet the CIAT gene bank alone contains ~30,000 accessions, most of which remain unexplored and therefore, is a reservoir of untapped diversity. Greater utilisation and characterisation of these resources, combined with multi-environment field trials, will allow validation of drought tolerance strategies, including combined stress scenarios such as heat and drought.

In the longer term, integrating genomic, phenotypic, and environmental data into predictive breeding programmes will enable the development of climate-resilient common bean varieties. My PhD has highlighted the power of computational approaches in identifying hundreds of potential candidate genes for traits of interest. However, taking this work forward experimentally is difficult, as it is not possible to validate all the genes, creating a large bottleneck. New machine learning approaches for predicting phenotypes, gene function and gene structure, offers a way to reduce the lists of genes before experimental validation, improving the efficiency of downstream experimental work.

Alongside this, where possible, genetic diversity should be conserved both *ex situ* and *in situ* within genetic hotspots to allow the species to evolve to ever-changing conditions. Crop wild relatives, which have not undergone domestication bottlenecks, represent an important

reservoir of diversity. As these are often found at farm boundaries, local education and management practices could support their conservation and future utilisation in breeding programmes.

Altogether, integrating genomic diversity, functional validation and new computational technologies will help in developing common beans which can produce yields in a large variety of ecogeographic conditions and are climate resilient through breeding programmes.

Appendices

All appendices and scripts are available at https://github.com/kdennj/PhD_2025

Appendix 1; *Subpopulation and phenotypic annotation for the 144 common bean accessions used in studying determinacy and photoperiod sensitivity.*

Appendix 2; *Genes and non-synonymous genetic variants within each of the identified QTLs associated with determinacy.*

Appendix 3; *Genes and non-synonymous genetic variants within each of the identified QTLs associated with photoperiod sensitivity.*

Appendix 4; *All phenotypes and transformations input into GAPIT to identify significant QTLs for tolerance to water-deficit.*

Appendix 5; *The phenotypes input into the correlation plot, after water application was stopped.*

Appendix 6; *All genes found within the 36 QTLs, their locations, if SNPeff high mutations are present and their annotations.*

Appendix 7; *TAIR and Phytomine annotations for all genes within the QTLs for water deficit response and recovery.*

G3, 2025, jkaf090
https://doi.org/10.1093/g3journal/jkaf090
Advance Access Publication Date: 19 April 2025
Plant Genetics and Genomics

Genome-wide association mapping dissects the selective breeding of determinacy and photoperiod sensitivity in common bean (*Phaseolus vulgaris* L.)

Kate E. Denning-James [1,2], Caspar Chater,[2,3] Andrés J. Cortés [4,5], Matthew W. Blair,[6] Diana Peláez,[4] Anthony Hall [1], José J. De Vega [1,*]

[1]Earlham Institute, Norwich, NR4 7UZ, UK
[2]Royal Botanic Gardens, Kew, Richmond TW9 3AE, UK
[3]Department of Plants, Photosynthesis and Soil, School of Biosciences, The University of Sheffield, Sheffield S10 2TN, UK
[4]Corporación Colombiana de Investigación Agropecuaria (AGROSAVIA)—C.I. La Selva, Rionegro 054048, Colombia
[5]Facultad de Ciencias Agrarias—Departamento de Ciencias Forestales, Universidad Nacional de Colombia—Sede Medellín, Medellín, Colombia
[6]Department of Agricultural and Environmental Sciences, College of Agriculture, Tennessee State University, 3500 John A Merritt Blvd, Nashville, TN 37209, USA

*Corresponding author: Email: jose.devega@earlham.ac.uk

The common bean (*Phaseolus vulgaris* L.) is a legume pulse crop that provides significant dietary and ecosystem benefits globally. We investigated 2 key traits, determinacy and photoperiod sensitivity, that are integral to its management and crop production, and that were early selected during the domestication of both Mesoamerican and Andean gene pools. Still, significant variation exists among common bean landraces for these traits. Since landraces form the basis for trait introgression in prebreeding, understanding these traits' genetic underpinnings and relation with population structure is vital for guiding breeding and genetic studies. We explored genetic admixture, principal component, and phylogenetic analyses using whole-genome sequencing to define subpopulations and gene pools. We used genome-wide association mapping (GWAS) to identify marker-trait associations in a diversity panel of common bean landraces. We observed a clear correlation between these traits, gene pool, and subpopulation structure. We found extensive admixture between the Andean and Mesoamerican gene pools in some regions. We identified 13 QTLs for determinacy and 10 QTLs for photoperiod sensitivity and underlying causative genes. Our study identified known and novel causative genes and a high proportion of pleiotropic effects for these traits in common bean, and likely translatable to other legume species.

Keywords: common bean; legume; determinacy; photoperiod; GWAS; domestication; Plant genetics and genomics

Introduction

The common bean is a global staple that provides significant dietary and economic services by improving health and nutrition while helping to reduce poverty, specifically in developing countries. Common beans have also been labeled as one of the essential crops to mediate climate change due to their lower environmental impact and protection of food and nutritional security (Foyer *et al.* 2016). Common beans are cultivated mainly as grain legumes, but the immature seeds, pods, and leaves are also eaten (Blair *et al.* 2010; Ganesan and Xu 2017). There are hundreds of varieties, and the prevailing type grown in a country depends on market preferences (Rawal and Navarro 2019). Common beans are rich in essential dietary components, such as protein, minerals, fiber, and micronutrients (Patto *et al.* 2015; Blair, Izquierdo, *et al.* 2013; Castro-Guerrero *et al.* 2016; Ganesan and Xu 2017), and protect against some forms of malnutrition, including stunting in children and micronutrient deficiencies (Jha *et al.* 2015; Suarez-Martinez *et al.* 2016; Ganesan and Xu 2017; Bernardi *et al.* 2023). As legumes, common beans have a symbiotic relationship with nitrogen-fixing bacteria, allowing them to fix atmospheric nitrogen and enhance nitrogen levels in the soil, thereby reducing the need for expensive chemical fertilizers while improving yields (Mylona *et al.* 1995; Cusworth *et al.* 2021; Mupangwa *et al.* 2021; Phiri and Njira 2023). Despite its widespread usability, trait segregation within and among bean landraces is still widespread, especially for critical agronomic traits such as growth habit and photoperiod.

The common bean underwent 2 separate domestications resulting in 2 gene pools: Andean and Mesoamerican. In addition, there are different races, intermediate species, and admixed accessions due to genetic isolation, fragmentation, and artificial selection for different morphological traits. The gene pools of common beans grow in a large variety of environments in the neotropics. These ecogeographic conditions, together with isolation by distance, have disrupted the gene flow between wild and domesticated common beans, and between the different gene pools (Santalla *et al.* 2004; Beebe *et al.* 2012). Consequently, there are large differences in their life history traits, morphology, and genetics (Gepts and Debouck 1991; Broughton *et al.* 2003; Beebe *et al.* 2012; Bitocchi *et al.* 2017). Another difference is cultivars are commonly autogamous and annual, while wild common beans and

Received on 20 January 2025; accepted on 03 April 2025

related species can be perennial and allogamous (Debouck et al. 1993; Schier et al. 2019; Chacon-Sanchez et al. 2021).

Photoperiod insensitivity and determinacy arose separately in both gene pools during the domestication of common beans, likely co-selected by growers (Weller et al. 2019; Repinski et al. 2012). Wild common beans tend to be indeterminate and photoperiod sensitive, requiring a particular day length to flower. Indeterminate growth is advantageous in the wild due to competition with surrounding vegetation, while photoperiod sensitivity (PS) was likely reinforced by divergent natural selection and local adaptation. On the other hand, photoperiod insensitivity was selected (likely unconsciously) as cultivated common beans were spread along a greater range of latitudes and environments. Determinacy, a developmental feature that causes common beans to have a terminal inflorescence when switching to a reproductive state (Cavalcante et al. 2020), optimized agricultural management and harvesting efficiency. Determinate common beans tend to have a bush growth habit with reduced branching and vining abilities compared with the indeterminate varieties (Kwak et al. 2012), therefore translocating biomass resources into an increased fitness output. While indeterminate and photoperiod sensitive landraces are common, the combined selection for photoperiod insensitivity and determinacy resulted in common bean varieties with shorter flowering periods, earlier maturation, and easier management during harvesting (Daba et al. 2016; González et al. 2016). Photoperiod insensitivity and determinacy are advantageous traits from an agronomical point of view due to earlier harvesting and shorter exposure to unfavorable weather patterns under climate change, consequently providing better food security for communities (Perez et al. 2020; Botero and Barnes 2022).

Modern breeding programs are moving beyond a yield-centered paradigm to target resistance to biotic and abiotic stress, and also nutritional quality (Singh and Schwartz 2010; Assefa et al. 2019; Caproni et al. 2020; Kachinski et al. 2022). Landraces and crop wild relatives offer a promising reservoir of genetic diversity for these traits by introgression from the landraces into the elite genetic background (Tai et al. 2014; Hu et al. 2021; Suarez, Polania et al. 2021; Suarez, Urban, et al. 2021). However, understanding the genetic diversity, population structure, patterns of adaptations, and how these correlate with determinacy and photoperiod insensitivity is required to guarantee the retention of these key domesticated traits within future breeding cycles, given their association with crop management and production (Beebe et al. 2012).

Common beans in Colombia are diverse regarding growth habits and PS. Colombia is the northernmost part of the Andean gene pool and south of the Mesoamerican and may act as a region of confluence between them. Consequently, it has been proposed that the region has a large amount of admixture and introgressive hybridization (Tohme et al. 1996; Blair et al. 2007; Blair, Cortes, et al. 2013; Leitao, Bicho, et al. 2021). Admixture and hybridization lead to introgressions from differential parental origins, introducing new alleles and novel epistatic interaction into a population, allowing for new trait combinations that could merge exotic variation from diverse germplasm with more agronomically desirable traits such as determinacy and photoperiod insensitivity.

Considering the above hypothesis, we characterized 144 representative landraces from Colombia and neighboring countries, together with controls from other regions, using whole-genome re-sequencing. We utilized genome-wide association mapping (GWAS) to identify significant SNPs for photoperiod insensitivity and determinacy in this diversity panel. The novelty of this work lies in that prior research commonly focused on the Mesoamerican diversity rather than the Andean, due to the greater genetic diversity in the former, and had ignored admixed materials as an essential source of variation. Furthermore, research has rarely utilized whole-genome sequencing of common bean accessions to undertake a GWAS on determinacy and photoperiod insensitivity phenotypes. Instead, previous work has mostly used QTL mapping and low-density marker panels, resulting in poor resolution (Kwak et al. 2008; González et al. 2016; García-Fernández et al. 2021).

Materials and methods

Diversity panel

The diversity panel was comprised of 144 genotypes mainly from Colombia and surrounding countries in Central and South America (Fig. 1). The panel contained accessions from elite backgrounds, landraces, heirlooms, weedy, and wild materials. The material was sourced from the International Centre for Tropical Agriculture (CIAT)'s genebank, the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)'s genebank, and heirlooms bought from the catalogs from "Jungle Seeds" (JungleSeeds 2020) and (Beans and Herbs 2020) in 2020. The panel was chosen to include control accessions from the Andean and Mesoamerican gene pools and races, while representing diverse seed coat colors and varying genetic backgrounds from Colombia and neighboring countries to focus on putatively admixed varieties.

Genotyping

The genotypes were whole genome re-sequenced using Illumina short reads. The accessions were grown at the Norwich Research Park (Norwich, UK) in 2021 until the expansion of the first true leaf, after which they were snap-frozen (~50–100 mg). The genomic DNA extraction for short-read sequencing from each accession was completed using a Qiagen DNAeasy kit (Qiagen, Germany). The DNA concentration of the samples was quantified for quality control using the Tecan Plate Read Infinite F200 Pro for a fluorometry-based assay. The sequencing of the samples was completed by Genomic services at Earlham Institute (Norwich, UK). LITE libraries, a cost-effective low-volume variant of the standard Illumina TruSeq DNA protocol, were constructed for the 144 accessions and were sequenced using 2 NovaSeq 6000 S4 v 1.5 flow cells with 150 bp paired-end reads, following the protocol in (Kirkwood et al. 2021).

Phenotyping

All 144 common bean accessions were evaluated at the Norwich Research Park (Norwich, UK) in temperature-controlled glasshouses. The experiments were conducted in 2 seasons; summer 2022 with long daylength (16:8) and winter 2023 with short daylength (12:12). The accessions were organized in a randomized block design with 3 or 2 replications, respectively. Management was conducted according to recommendations for common bean cultivation.

The diversity panel was characterized for the days to flowering (DTF), seed size (SS), weight of 100 seeds (E100_SW; estimated based on the weights of seeds harvested and projected to 100 seeds), determinacy (D; terminal flower bud presence) (Cavalcante et al. 2020), and PS (flowering in none, 1 or both seasons). DTF was split into the 2 seasons due to PS in certain accessions and PS was characterized in 3 ways for the GWAS.

The statistical analysis of variance (1-way ANOVA) of the phenotypic data was done in R, then the Pearsons's correlation coefficient was calculated and visualized using the R package "corrplot" (Wei and Simko 2021).

Fig. 1. Distribution of the 127 common beans with location data that were used in this study. The coordinates of the capital city were used for those without coordinate data. Produced with QGIS.

Preprocessing genotype data

The raw sequence reads were processed with TrimGalore (v. 0.5.0) (Krueger *et al.* 2023) to remove adapters and poor-quality reads, and then quality checked using FastQC (Wingett and Andrews 2018) and MultiQC (Ewels *et al.* 2016). The trimmed reads were aligned to the Andean reference genome, *Phaseolus vulgaris* G19833, v2.1 (Schmutz *et al.* 2014) downloaded from Phytozome (Goodstein *et al.* 2012) with BWA-MEM (v 0.7.13) (Li and Durbin 2009) and "-M -R" to add read group information and allow compatibility with GATK. SAMtools (v 1.7) combined, compressed, and sorted the aligned files (Danecek *et al.* 2021). Picardtools (https://broadinstitute.github.io/picard/) (v 2.1.1) marked duplicates and BamTools indexed the alignments (Barnett *et al.* 2011). The percentage of alignments were calculated at this stage. The genotype data were divided into 10 Mbp regions (Garrison and Marth 2012) (v 1.0.2) to run the Genome Analysis ToolKit (GATK v 4.2) haplotype caller with default parameters (Van der Auwera and O'Connor 2020). This identified 20.2 million variant loci (~17.1 M SNPs and ~3.4 M indels).

Population structure analysis

The resulting VCF file from GATK using the Andean reference ("Andean VCF") was filtered further with BCFtools to retain calls with a minimum depth of 5 reads per variant call (FMT/DP ≥ 5), a locus call quality over 30, maximum missing calls per locus of 5%, to keep only biallelic SNP locus, and for a minor allele frequency over 2%. The resulting VCF had ~9 million SNP loci. Then, the VCF was filtered for a maximum heterozygosity of 20% per locus using TASSEL 5 (v. 20230314) (Bradbury *et al.* 2007). This was then filtered for linkage disequilibrium (LD) (based on LD decay) and thinned with a window size of 10 bps using BCFtools prune.

The population structure of the panel was analyzed using ADMIXTURE (v 1.3.0) (Alexander and Lange 2011) on a subset of 88,786 SNP loci. ADMIXTURE was run for K = 2 to K = 10 and the ideal number of K was determined using the cross-validation error. Accessions were allocated a group when their membership coefficient (q) was greater than 0.7. Plotting was completed in R using the packages "ggplot2' (Ginestet 2011).

Genome-wide association study

The "Andean VCF" from GATK was filtered with BCFtools (v 1.12) (Danecek *et al.* 2021) for biallelic loci, a minor allele frequency of 1% and thinned with a window size of 5 bp. To understand the genetic relationship between accessions, we used a principal component analysis (PCA) generated with GAPIT v.3 (Wang and Zhang 2021) on a subset of 2,572,124 loci.

A genome-wide association study investigated marker-trait association for determinacy and photoperiod insensitivity phenotypes

using GAPIT v.3 (Wang and Zhang 2021) with 3 principal components. We ran with the models Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) (Huang et al. 2019), Fixed and random model Circulating Probability Unification (FarmCPU) (Liu et al. 2016), and Mixed Linear Model (MLM) (Zhang et al. 2010). BLINK and FarmCPU were identified as the best multi-locus models for different heritability levels, improving statistical power (Huang et al. 2019; Merrick et al. 2022; Cebeci et al. 2023). While MLM was chosen for single-locus analysis as a baseline for comparison to BLINK and FarmCPU.

GAPIT was run on the whole panel (144 accessions) and on the Andean subpanel (as defined at K2 ADMIXTURE; 108 accessions). To run BLINK, GAPIT completed the analysis with the option "Random.model = TRUE" as not to calculate R^2 for phenotypic variance explained values after GWAS. The quantile-quantile (QQ) plots were used to understand the suitability of the models to the data. Plotting was completed in R using the package 'ggplot2' (Ginestet 2011).

Selecting significant loci, candidate gene mining, and functional annotation

Significant marker-trait associations (MTAs) were investigated further when they had a $-\log_{10}$(P-value) over 7 and were confirmed by 2 models from GAPIT. QTLs were defined as ±100 kbp from the MTA based on the estimated LD decay distances in common bean diversity panels and by using a $r^2 = 0.25$ cutoff (estimated decay as 114 kb) (Moghaddam et al. 2016; Valdisser et al. 2017; Campa et al. 2018; Raggi et al. 2019; Wu et al. 2020, 2024; Ugwuanyi et al. 2022; Reinprecht et al. 2023). This is shorter than the calculated recombination rate in common bean of 3.72 cM/Mb (Bhakta et al. 2015). LD decay was estimated for the diversity panel (mean $R^2 = 0.27$) and subpopulation at K = 2 (Andean mean $R^2 = 0.21$, Mesoamerican mean $R^2 = 0.2$) using PopLDdecay software following Wu et al. (2020) (Zhang et al. 2019).

Identified loci were compared with the Andean reference genome, *Phaseolus vulgaris* G19833 v2.1 in JBrowse (Schmutz et al. 2014; Diesh et al. 2023) while considering "highimpact" mutations identified by SnpEff (Cingolani et al. 2012). Once genes were identified, their putative function was explored using PhytoMine (Goodstein et al. 2012) (*Phaseolus vulgaris* v.2), BLAST (Camacho et al. 2009) against the nonredundant protein database at NCBI, and finally against the TAIR database if no gene function could be identified in close relatives (Huala et al. 2001). The loci were compared with previous studies and literature. PulseDB was used for comparison, particularly for QTLs and markers related to developmental and flowering phenotypes (Humann et al. 2019). QTLs and markers were mapped to the reference genome to estimate the conversion from cM to Mb in JBrowse.

Results

Population structure

The diversity panel split into the 2 gene pools, the Andean and Mesoamerican (Figs. 2a and 3a). At K6 (Fig. 2b), the Mesoamerican group split into 2 subpopulations (M1 and M2), while the Andean subgroup split into 4 subpopulations. Two of these subpopulations included only accessions from Colombia and were named C1 and C2. A subpopulation containing accessions from Colombia and Ecuador/Peru was named C-EP. The remaining subpopulation was named A1. In the PCA (Fig. 3a), PC 1 explained 38.8% of the variation in our diversity splitting the 2 gene pools, while PC2 accounted for 5.06% of the variation, splitting the Mesoamerican subgroups (M1 and M2) and separating C-EP from the other Andean subgroups. A total of 11 accessions were classified as admixed between the Andean and Mesoamerican gene pools (Admx_AM), as they had an ancestry composition lower than 70% from either of the origins ($q < 0.7$). The Admx_AM accessions were all indeterminate and produced a variety of seed sizes. Seven were landraces and 2 were wild. There was also a mix of photoperiod sensitive and insensitive accessions.

The Colombian subgroups (C1 and C2; Fig. 2b) contained medium and large seeded landraces. However, the subpopulations distinguished by determinacy; C1 contained mainly insensitive determinate accessions while C2 contained sensitive indeterminate accessions. The A1 group contained large and medium seeded landraces that were mainly photoperiod insensitive. The C-EP population contained accessions from Ecuador, Peru, and Colombia. This group contained large-seeded indeterminate landraces and also included accessions from races previously identified to be from the Andean gene pool. The Mesoamerican subgroups (M1 and M2; Fig. 2b) were also distinguished by phenotypic data. They both contained indeterminate and determinate accessions; however, M1 was mainly medium seeded while M2 was mainly small seeded. This is summarized in Table 1 and Supplementary Table 1.

Colombian accessions can be found within all the subgroups and admixed groups at K = 6 (Fig. 2b). While the admixture accessions are mainly from Colombia, while 1 sample is a wild "Ecuador" accession.

The Andean accessions had a lower proportion of heterozygous sites (<0.1) than the Mesoamerican accessions, which were more heterozygous (Fig. 3b). The 6 highly heterozygous accessions (>25% of the loci) were found within the Andean X Mesoamerican hybrid (Admixed-AM) subpopulation (Fig. 3b) and were from Colombia. Finally, the outlier accession with the lowest alignment to the Andean reference genome and low proportion of heterozygous sites was a wild accession from Ecuador.

Phenotypic variation and correlations

The correlation coefficient was estimated for each pair of traits (Fig. 4), averaged over 2 seasons or studied in both years. There was a positive correlation between DTF from winter and summer ($r = 0.57$). Both DTF were negatively correlated with PS [$r = -0.72$ (DTF_S22), $r = -0.77$ (DTF_W23)] and D [$r = -0.35$ (DTF_S22), $r = -0.43$ (DTF_W23)]. Population structure at either 2 or 6 ancestries (K2, K6) was positively correlated with D [$r = 0.32$ (K6), $r = 0.37$ (K2)] but negatively correlated with SS [$r = -0.44$ (K6), $r = -0.4$ (K2)] and E100_SW [$r = -0.37$ (K6), $r = -0.47$ (K2)]. SS was not correlated with DTF_S22, DTF_W23, D, or PS ($r = -0.13$, $r = -0.07$, $r = -0.12$, $r = 0.09$). However, E100_SW was positively correlated with PS ($r = 0.18$) and SS ($r = 0.87$) but negatively correlated with DTF_S22 ($r = -0.22$). Then D and PS were positively correlated ($r = 0.45$).

Figure 5, a–c showed the distributions of the phenotyping for traits E100_SW, S22_DTF, and W23_DTF, respectively. The seed weights (Fig. 5a) were normally distributed, while the DTF in summer and winter (Fig. 5, b and c) were binomial distributions; the peaks were around 42- and 54-days postsowing in summer, and around 70- and 90 days in winter. When analyzing the phenotypes by subpopulation, we can see that C-EP (Fig. 2b) did not flower during winter in the UK, W23_DTF, as was mainly photoperiod sensitive. This is further supported by the correlation plot (Fig. 4). Furthermore, determinacy, photoperiod insensitivity, and DTF are correlated. The determinate accessions flower earlier than the indeterminate, supporting the binomial distribution.

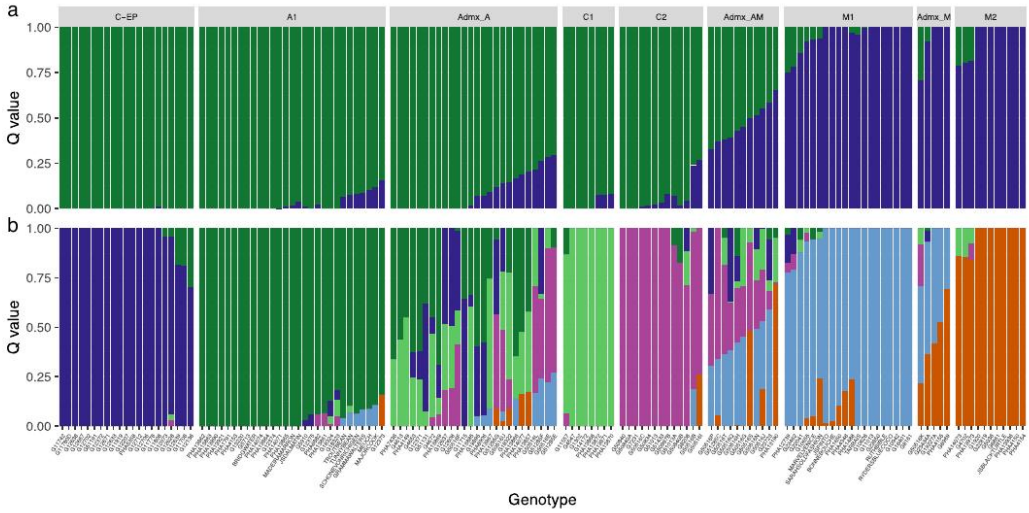


Fig. 2. Analysis of the population structure of 144 accessions belonging to our diversity panel focusing on Colombia at K = 2, Andean or Mesoamerican groups a) and K = 6 b). (C-EP) accessions mainly from Peru, then Ecuador and Colombia; (A1) Andean accessions from a variety of South American countries; (C1) mostly determinate Colombian landraces; (C2) indeterminate Colombian landraces; (M1) mainly medium seeded** from Central America and Colombia; (M2) mainly small seeded** from Central America and Colombia. (Admx_AM) Andean X Mesoamerican hybrids; (Admx_A) and (Admx_M) admixed accessions between subpopulations (ancestry composition q < 0.7 at K = 6). **P < 0.01 using a 2-tailed student t-test with unequal variance.

GWAS for determinacy

The GWAS was performed using the models BLINK, FarmCPU, and MLM with GAPIT (Fig. 6, a and b). The QQ plots (Fig. 6, c and d) provided evidence that the selected models were well fitted to identify significant MTAs for the dataset. We identified 13 MTAs with a significant *P*-value ($-\log_{10}$(*P*-value) > 7), corresponding to 13 QTLs. We focused on 7 significant MTAs that were identified for the whole panel based on the criteria laid out in the methods (vertical lines in Fig. 6). The 7 QTLs were found on chromosomes Pv01, Pv07, Pv08, Pv09, and Pv10 (Table 2). Five of the 7 QTLS were also identified for the Andean subset.

Putative candidate genes were identified for determinacy based on the significant MTAs and corresponding QTL windows. The identified genes and QTLs are listed in Supplementary Tables 2 and 3.

GWAS for PS

The GWAS was performed using the BLINK and FarmCPU models with GAPIT (Fig. 7, a and b). The QQ plots (Fig. 7, c and d) provide evidence that the selected models are fitted to identify significant MTAs for the dataset. We identified 10 QTLs ($-\log_{10}$(P-value) > 7). We focused on 6 QTLs for the whole panel based on criteria laid out in the methods. The MTAs were found on chromosomes Pv04, Pv05, Pv07, Pv08, and Pv09 (vertical lines in Fig. 7). Six QTLs were identified for the Andean subset panel in Chromosomes Pv05, Pv07, Pv08, Pv09, and Pv11. The QTL in Pv04 and Pv09 were found in the full dataset only. The QTL in Pv9 and Pv11 were found in the Andean subset only. Candidate genes were identified for the significant MTAs and their corresponding QTLs. The identified genes and QTLs are listed in Supplementary Tables 2 and 3.

Discussion

We delimited subpopulations in a panel of 144 accessions, initially divided by domestication event into the 2 Andean and the Mesoamerican gene pools (Figs. 2 and 3) (Blair, Cortes, *et al.* 2013; Kami *et al.* 1995). The Mesoamerican gene pool is generally more diverse (Mamidi *et al.* 2013; Schmutz *et al.* 2014) with less influence from domestication bottlenecks. Furthermore, the Mesoamerican gene pool within our diversity panel is also more heterozygous, suggesting that the Andean gene pool has undergone fewer outcrossing events. These crosses between gene pools occur during common bean dissemination, breeding programs and selection based on market preferences (Hoyos-Villegas *et al.* 2017; de Almeida *et al.* 2020; Botero *et al.* 2021; Bellucci *et al.* 2023). However, care needs to be taken when utilizing market sampling information. This is highlighted by the 2 "Peruvian" accessions collected from markets that fall with the Mesoamerican subpopulation (Supplementary Table 1).

Admixture was commonly observed in the panel, including 26 admixed Andean accessions, 5 admixed Mesoamerican accessions, and 11 Mesoamerican × Andean accessions. This supports our initial hypothesis that Colombia and neighbouring countries hold large common bean variation, including hybrids between both gene pools (Gori *et al.* 2022; Myers *et al.* 2000; Pironon *et al.* 2020). The wider crosses between gene pools compared with within gene pools resulted in a larger observed heterozygosity in the hybrid accessions, supporting the outcrossing events and movement between gene pools. One implication of this study is that admixed Colombian hybrid landraces bridge Andean and Mesoamerican gene pools, and novel allelic and epistatic interactions likely filtered out deleterious effects (Cichy *et al.* 2015) due to stronger purifying selection with increased recombination. After all, recombination increases local effective

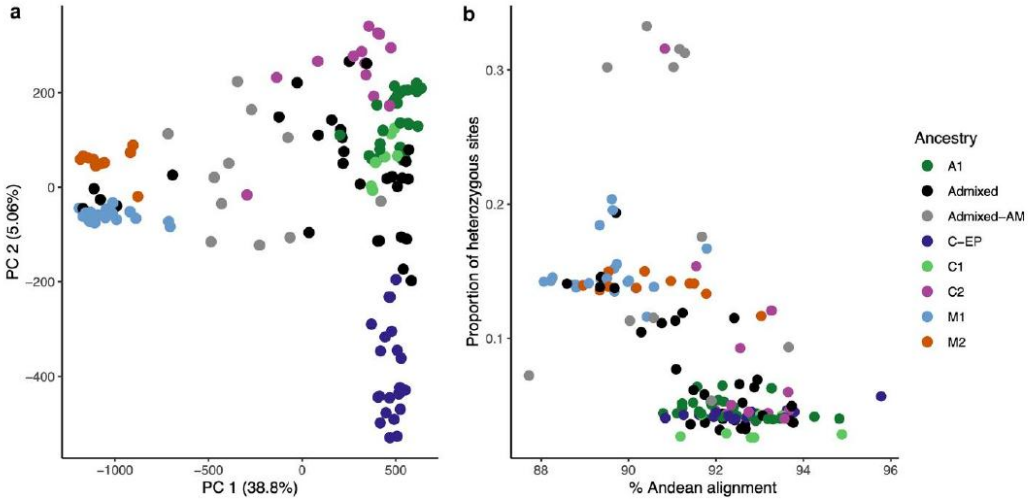


Fig. 3. a) Principle component analysis (PCA) plot of PC1 against PC2. b) Proportion of heterozygous sites against the percentage of read pair alignment to the Andean reference genome G19833 (Schmutz *et al.* 2014). The colors illustrate the population structure of our diversity panel.

Table 1. Phenotypic characteristics associated with each subpopulation.

Subpopulation	Gene pool	Determinancy	Photo. sen.	Seed size	Origin
C1	Andean	Mainly determinate	Insensitive	Mainly large	Colombia and Ecuador
C2	Andean	Indeterminate	Mainly sensitive	Mainly large	Colombia
A1	Andean	Both	Mainly insensitive	Mainly large	South America, Heirlooms, Colombia
C-EP	Andean	Indeterminate	Sensitive	Large	Colombia, Ecuador, Peru
Admix_A	Andean	Mainly indeterminate	Both	Mainly large	Colombia and South America
M1	Mesoamerican	Mainly indeterminate	Both	Mainly medium**	Central America, Colombia, Heirlooms, Peru
M2	Mesoamerican	Mainly indeterminate	Mainly insensitive	Mainly small**	Central America, Colombia
Admix_M	Mesoamerican	Mainly indeterminate	Insensitive	Small and medium	Colombia, Brazil, Heirlooms, Central America
Admix_AM	AxM hybrids	Indeterminate	Mainly sensitive	Mainly medium	Colombia and Ecuador

population size (Ne) and limits Hill–Robertson interference (Hill and Robertson 2007). This suggests the Colombian hybrids have promising potential for breeding. However, the diversity panel may also be biased and underestimating their prevalence in other regions due to the large number of Colombian accessions in our diversity panel.

We observed some traits associated with demography, including determinacy and PS: C1 and C2 shared origin but could be separated by ancestry admixture analysis, and were characterized by different determinacy, as C1 contained mainly determinate accessions, and C2 mainly indeterminate accessions. Furthermore, the population structure suggests that Colombian farmers have not selected varieties based on the seed characteristics studied (e.g. SS) (Botero *et al.* 2021).

Indeterminate and photoperiod sensitive landraces were common, despite the combined selection for photoperiod insensitivity and determinacy resulting in common bean varieties with shorter flowering periods (DTF) and easier management. Prior research supports the correlation between DTF and phenotypes such as seed weight, determinacy and growth habit (Tar'an *et al.* 2002; Moghaddam *et al.* 2016; Hoyos-Villegas *et al.* 2017; Elias *et al.* 2021; Vargas *et al.* 2021). These phenotypes are related to apical meristems and floral development (Sablowski 2007).

We observed the distribution of DTF values, in either summer or winter, were bimodal, i.e. had 2 peaks (Fig. 5, b and c). This likely occurred due to the determinate types flowering first and then followed by the indeterminate beans (Coelho *et al.* 2023). The distribution also correlates to growth habits as bush types typically flower earlier than climbing types (Ugwuanyi *et al.* 2022). Figure 2a supports that PS arose during domestication in both gene pools (Weller *et al.* 2019).

The Andean accessions within our diversity panel were large and medium seeded while the Mesoamerican accessions were small and medium sized, which supports previous research (Blair *et al.* 2009). Among the Mesoamerican accessions, the Mesoamerican race is characterized by small-seeds, while the Durango–Jalisco race is characterized by medium seeds (Beebe *et al.* 2000; Zhang *et al.* 2008; Blair *et al.* 2009; Giordani *et al.* 2022). We could not separate our diversity panel into subpopulations matching these races due to a lack of Mesoamerican diversity in the panel, a limited genetic component for the SS trait, or introgressions occurring in the Mesoamerican Colombian accessions.

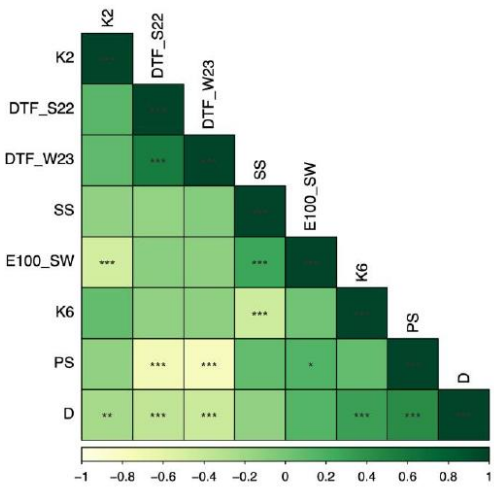


Fig. 4. Pearson correlation coefficients among five agronomic traits and population structure measured in 144 common bean genotypes grown at the Norwich Research Park, Norwich, UK in 2022 and 2023. K6, K6 subgroups from ADMIXTURE; K2, K2 subgroups from ADMIXTURE; D, determinacy; PS, photoperiod sensitivity; SS, seed size; E100_SW, estimated weight of 100 seeds; DTF_W23, DTF from winter 2023; DTF_S22, DTF from summer 2022. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Interestingly, Ecuador accessions are often separated from Andean subgroups, suggesting that they are members of the PhI group or a possible sister species *Phaseolus debouckii* (Chacon-Sanchez *et al.* 2007; Rendon-Anaya *et al.* 2017). Further to this, the wild Ecuador accession is separated from both gene pools (Figs. 2 and 3), suggesting a separate ancestry originating from Ecuador or Peru (Bitocchi *et al.* 2012; Bitocchi *et al.* 2017). Finally, the C-EP group (Fig. 2b) are mainly photoperiod sensitive (Fig. 5f), possibly due to a different domestication history or due to their quatorial provenance not necessitating evolution under fluctuating photoperiods.

By leveraging this diversity panel and its trait segregation across the demographic stratification, we prioritized 13 QTLs for determinacy and 10 QTLs for PS. Four of the QTLs for PS, and 4 for determinacy, were also identified only for the Andean subset, but not the whole panel. The Andean gene pool has adapted to lower latitudes than the Mesoamerican pool, resulting in differential selection for PS between the 2 gene pools. The LD was estimated as 114 kb from an R^2 cutoff of 0.25, this value is consistent with WGS data of diversity panels rather than breeding populations (Campa *et al.* 2018; Diniz *et al.* 2018; Reinprecht *et al.* 2023; Ambachew *et al.* 2024). LD in common beans is impacted by the evolutionary and breeding history of the accessions in the diversity panel; therefore, a 200 kb region accounts for the higher resolution of WGS as well as allowing for LD (Moghaddam *et al.* 2016; Valdisser *et al.* 2017).

During this study we completed analysis with the Andean reference genome (Schmutz *et al.* 2014). This reference genome was selected for being the most complete at the time of analysis and because our panel has a higher proportion of Andean accessions based on population structure analysis (Fig. 2). The accessions also had higher alignments to the Andean reference genome (92.5% ± 1 and 89.9% ± 1.1% for the Andean and Mesoamerican subpopulations, respectively) and no difference in metrics to the Mesoamerican reference genomes (Supplementary Table 1).

QTLs and candidate genes associated with determinacy

Three QTLs in chromosome 1

We identified a determinacy QTL in chr 1 -Pv01- (D1.4-D1.6; Table 2), identified in other studies (Moghaddam *et al.* 2016; da Silva *et al.* 2018; Kamfwa *et al.* 2019; Sedlar *et al.* 2020; Vargas *et al.* 2021; Keller *et al.* 2022) as a hotspot of allelic variation, named the *Fin* locus. The *Fin* locus has been mapped to ~44.5 Mb (Pérez-Vega *et al.* 2010; Kamfwa *et al.* 2019). This co-segregates with an upstream gene, *TFL1y* (*Phvul*.001G189200), a candidate gene for flowering, vegetative growth, rate of plant production, and determinacy (Kwak *et al.* 2008, 2012; Repinski *et al.* 2012; Cichy *et al.* 2015; González *et al.* 2016; Campa *et al.* 2018; Delfini *et al.* 2021). Consequently, the *Fin* locus has pleiotropic effects due to associations with many development traits such as determinacy, shoot biomass, DTF, days to maturity, plant architecture, embryo abortion, number of pods per plant, number of seeds per plant (seed yield and weight), and disease resistance (Miklas *et al.* 2001; González *et al.* 2016; Delfini *et al.* 2021; Soler-Garzón *et al.* 2024). However, segregation for this QTL hotspot in Pv01 may prove difficult in breeding programs due to these pleiotropic effects (Vargas *et al.* 2021).

Further candidate genes have been identified in this QTL, such as *Phvul.001G192200*. This gene is an ortholog of *LIGHT-REGULATED WD1* (*LWD1*), a gene involved in the circadian rhythm pathway (Wu *et al.* 2008; Moghaddam *et al.* 2016; Delfini *et al.* 2021), or *Phvul.001G192300*, which is an ortholog of *SPINDLY* (*SPY*). SPY interacts with genes in the reproductive pathway (Tseng *et al.* 2004; Moghaddam *et al.* 2016; da Silva *et al.* 2018) and has been associated with days to maturity (Reinprecht *et al.* 2023).

Another QTL we identified on Pv01 (D1.3; Table 2) contains the gene *Phvul.001G168700*. This gene is related to the phytochrome interacting factor 1 (PIF1) transcription factor isoform X1 in the legume *Vigna radiata* (Bateman *et al.* 2023). This bHLH transcription factor is involved in many light-dependent pathways in plant development and interacts with circadian clock genes (Kim *et al.* 2016).

QTL D7.1 in chromosome 7

The QTL at Pv07 (D7.1) was identified in the whole and Andean panel. The QTL contains the gene *Phvul.007G244700*. This is related to a transcriptional corepressor, Leunig-homolog in *Vigna radiata* (Bateman *et al.* 2023). In *Arabidopsis*, Luenig-homologs have functional redundancy with Leunigs (LUGs), and are involved in embryo and floral development (Sitaraman *et al.* 2008). This QTL has been associated with SS, seed weight, and growth habit (Kwak *et al.* 2008; da Silva *et al.* 2018; Elias *et al.* 2021; Keller *et al.* 2022), suggesting it may have pleiotropic effects.

QTL D8.2 in chromosome 8

The QTL identified on Pv08 (D8.2; Table 2) for determinacy has previously been identified for plant architecture (da Silva *et al.* 2018). However, no gene with a clear function was identified. We have, however, identified a possible candidate gene for further investigation; *Phvul.008G170000*. This encodes a putative fantastic 4 (FAF) domain-containing protein. In *Arabidopsis*, FAF proteins regulate shoot meristem size and architecture (Wahl *et al.* 2010).

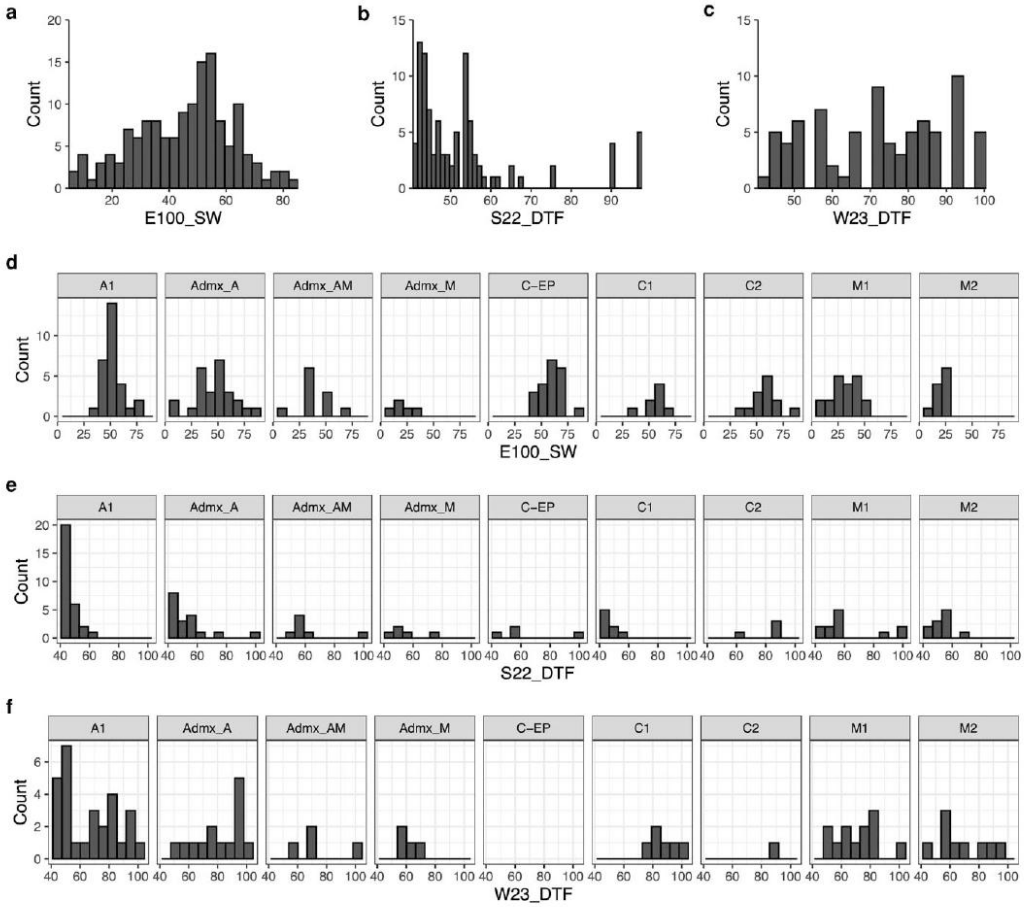


Fig. 5. Frequency distribution of seed weight and days to flower traits evaluated in 2 seasons in a common bean diversity panel. a) E100_SW, estimated weight of 100 seeds; b) phenological DTF in the summer 2022 (S22_DTF) and c) in the winter 2023 (W23_DTF) at the Norwich Research Park, excluding those which did not flower. The distributions were split into the subpopulations from K6 ADMIXTURE. d) E100_SW***; e) S22_DTF***; f) W23_DTF*. Completed a 1-way ANOVA for E100_SW, S22_DTF, and W23_DTF. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

QTL D9.1 in chromosome 9

The QTL D9.1 in chr 9 was identified in the whole and Andean panel. Nearby QTLs have been identified for yield and determinacy (Kamfwa *et al.* 2015; Campa *et al.* 2018). The gene *Phvul.009G138100* is found within this QTL and contains the significant MTA found by GAPIT (Wang and Zhang 2021). This gene has an insertion that possibly affects function (Cingolani *et al.* 2012). This gene is uncharacterized in common bean but has homology to the root meristem growth factor 9 from *Glycine soja* (Goodstein *et al.* 2012; Bateman *et al.* 2023). This growth factor is expressed in the roots and flowers, regulating and maintaining apical meristems, and therefore both root and floral development, SS, and leaf architecture (Chen *et al.* 2019; Shinohara 2021). Although it has previously been identified as a candidate gene associated with Mesoamerican domestication (Schmutz *et al.* 2014), we found the QTL in the Andean panel, suggesting that it has also played a role in the Andean domestication event.

QTL D10.1 in chromosome 10

The QTL on Pv10 (D10.1) is located near QTLs for plant height and number of nodules and near genes associated with metabolic changes during domestication, once again suggesting pleiotropic effects (Delfini *et al.* 2021; de Souza *et al.* 2023). Three of the genes within this region encode bHLHLZip proteins: *Phvul.010G158500*, *Phvul.010G158300*, and *Phvul.010G158200*. These bHLH transcription factors may be involved in the regulation of flowering genes (Zhou *et al.* 2019). The gene *Phvul.010G158500* displays nonsynonymous modifications in our panel, including insertions, deletions, and other variants linked to frameshift mutations and gained stop codons (Cingolani *et al.* 2012). Homology to *Vigna angularis* suggests this gene may be related to the transcription factor bHLH25, and possibly linked to a circadian rhythm-associated protein (Goodstein *et al.* 2012).

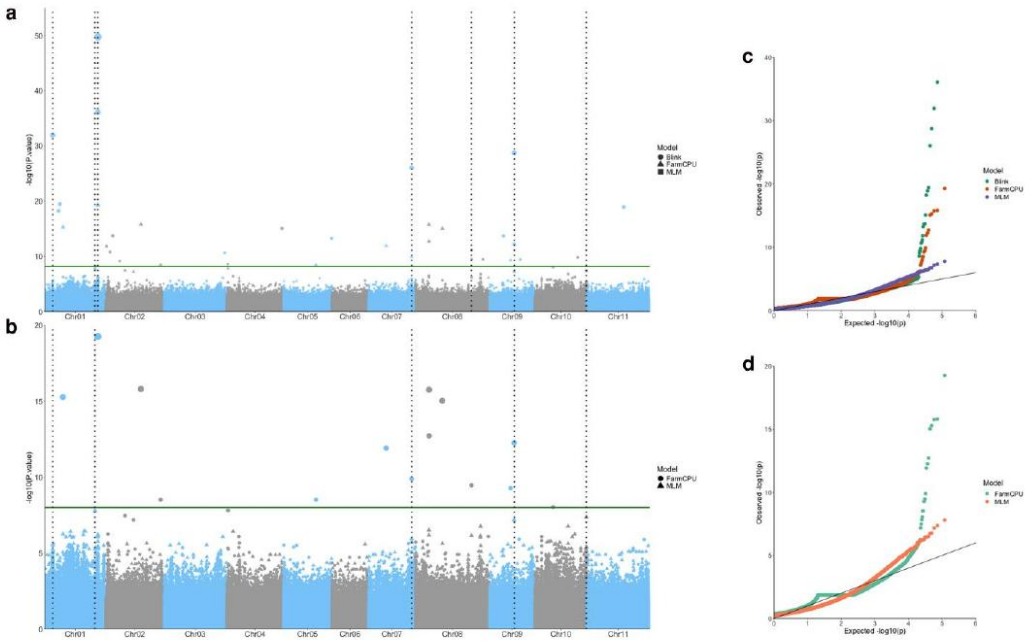


Fig. 6. Manhattan plots highlighting markers significantly associated with determinacy on (a) the whole panel and b) the Andean subpanel. The analyses were completed with GAPIT and the models are FarmCPU, BLINK, or MLM (Huang *et al.* 2019; Liu *et al.* 2016; Wang and Zhang 2021; Zhang *et al.* 2010). The X-axis represents the genomic position of markers and the Y-axis is the −log 10 of the *P*-values for association with the phenotype. The vertical lines correspond to QTLs found by at least 2 models. Point size correlates to −log10(*P*-value). Quantile-quantile (QQ) plots are provided for c) the whole panel and d) the Andean panel.

Candidate genes for PS

QTL PS4.1 in chromosome 4

One QTL for PS was found on Pv04 (PS4.1; Table 2) from the analysis on the whole panel. Within this QTL, 4 genes were identified, 3 of which (*Phvul.004G110200*, *Phvul.004G110301*, and *Phvul.004G110000*) have nonsynonymous mutations such as a stop lost, stop gained, or a frameshift mutation in our panel (Cingolani *et al.* 2012). However, the genes are uncharacterized.

Two QTLs in chromosome 5

Two QTLs were identified in Pv05: PS5.2 for the Andean panel and PS5.1 for the whole panel. PS5.2 overlaps with a previously identified QTL for seed weight, DTF, and pod weight (Arriagada *et al.* 2022; Reinprecht *et al.* 2023). However, this previous analysis with a limited number of markers did not identify a candidate gene. Based on sequence homology with *Vigna radiata*, we identified the gene *Phvul.005G077000*, which encodes a proton gradient regulation 5 (PGR5) protein (Bateman *et al.* 2023). PGR5 is involved in plant growth under different light conditions due to interactions with Photosystem I, and consequently putatively associated with differentiating PS in our panel (Munekage *et al.* 2002). The QTL PS5.1 contained 2 genes, one of which, *Phvul.005G076300*, may encode a bidirectional sugar transporter, named SWEET protein. Evidence suggests SWEET proteins have essential roles in plant development, including in reproductive organs and bud growth (Gautam *et al.* 2022).

Two QTLs in chromosome 7

Two QTLs were also identified on Pv07. PS7.1 and PS7.2, both in the Andean and the whole panel. The QTL PS7.2 contains the genes *Phvul.007G157400* and *Phvul.007G156200*. Homology with *Arabidopsis* suggests that *Phvul.007G157400* encodes a BANQUE3 BHLH161 protein. BANQUE3 is negatively regulated by *APETALA3* and *PISTILLATA* in petals and is involved in light-regulated responses and flowering time (Huala *et al.* 2001; Mara *et al.* 2010). *Phvul.007G156200* may encode the BHLH transcription factor PIF4 (Phytochrome Interacting Factor 4) based on homology with *Vigna radiata* and *Glycine soja* (Goodstein *et al.* 2012; Bateman *et al.* 2023). PIF4 is a downstream signaling component integrating environmental cues such as light (Bateman *et al.* 2023).

The QTL PS7.1 overlaps with a previously identified QTL for plant production traits (González *et al.* 2016). The QTL includes the gene *Phvul.007G117400* which encodes a putative JUMONJI domain-containing protein (Goodstein *et al.* 2012). JUMONJI proteins are involved in multiple plant developmental processes such as flowering and leaf senescence (Gan *et al.* 2014; Liu *et al.* 2019; Yamaguchi 2021; Xin *et al.* 2024). *Phvul.007G117400*s homology with a JUMONJI16 orthologue in *Vigna radiata* also supports this role (Bateman *et al.* 2023).

Two QTLs in chromosome 8

One of the QTLs found in Pv08 is PS8.1 from the whole panel. This QTL has been associated with determinacy (Campa *et al.* 2018), seed weight (Elias *et al.* 2021), DTF (Raggi *et al.* 2019), and pod

Table 2. QTLs for determinacy and photoperiod sensitivity.

Name	Chromosome	Start	End	Trait	Panel
D1.1	Chr01	6,512,000	6,521,000	Determinacy	Andean + Whole
D1.2	Chr01	11,363,000	11,372,000	Determinacy	Andean
D1.3	Chr01	42,404,000	42,413,000	Determinacy	Andean + Whole
D1.4	Chr01	44,856,000	44,847,000	Determinacy	Whole
D1.5	Chr01	44,932,000	44,941,000	Determinacy	Andean + Whole
D1.6	Chr01	45,098,000	45,107,000	Determinacy	Whole
D2.1	Chr02	24,821,000	24,830,000	Determinacy	Andean
D3.1	Chr03	25,608,000	25,617,000	Determinacy	Andean
PS4.1	Chr04	38,316,000	38,325,000	Photo sensitivity	Whole
PS5.1	Chr05	16,423,000	16,432,000	Photo sensitivity	Whole
PS5.2	Chr05	18,321,000	18,330,000	Photo sensitivity	Andean
PS7.1	Chr07	16,829,000	16,838,000	Photo sensitivity	Andean + Whole
PS7.2	Chr07	26,485,000	26,494,000	Photo sensitivity	Andean + Whole
D7.1	Chr07	36,860,000	36,869,000	Determinacy	Andean + Whole
PS8.1	Chr08	4,234,000	4,243,000	Photo sensitivity	Whole
D8.1	Chr08	7,440,000	7,449,000	Determinacy	Andean
PS8.2	Chr08	8,320,000	8,329,000	Photo sensitivity	Andean
D8.2	Chr08	47,582,000	47,591,000	Determinacy	Whole
D9.1	Chr09	20,814,000	20,823,000	Determinacy	Andean + Whole
PS9.1	Chr09	21,640,000	21,649,000	Photo sensitivity	Whole
PS9.2	Chr09	34,445,000	34,454,000	Photo sensitivity	Andean
D10.1	Chr10	43,762,000	43,771,000	Determinacy	Andean + Whole
PS11.1	Chr11	204,000	213,000	Photo sensitivity	Andean

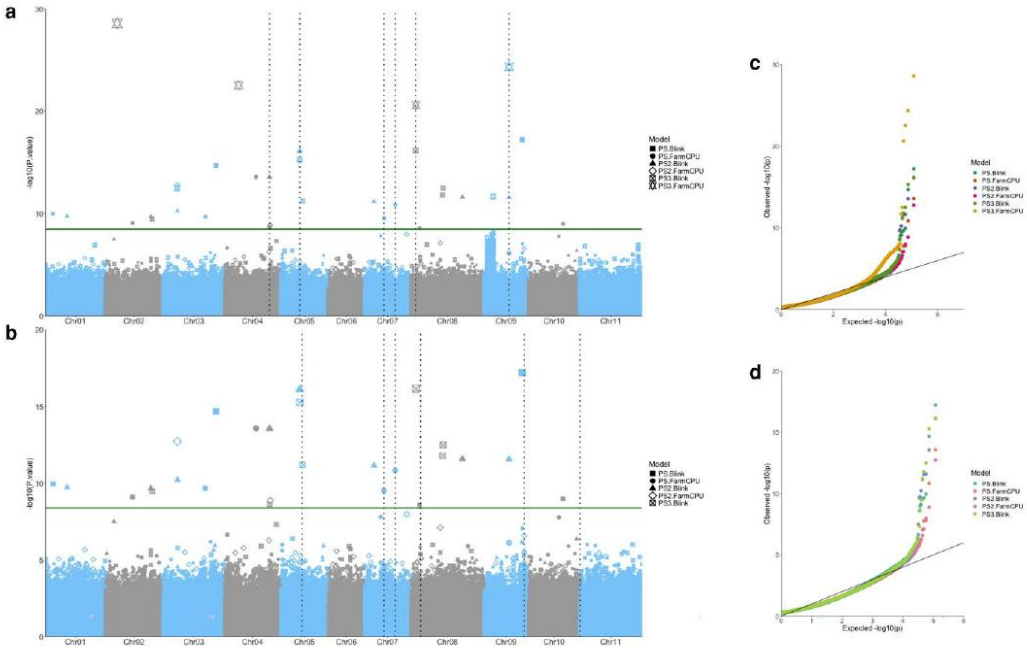


Fig. 7. Manhattan plots highlighting markers significantly associated with photoperiod insensitivity on (a) the whole panel and b) the Andean subpanel. The analyses were completed with GAPIT and the models FarmCPU, BLINK, or MLM (Zhang *et al.* 2010; Liu *et al.* 2016; Huang *et al.* 2019; Wang and Zhang 2021). The X-axis represents the genomic position of markers and the Y-axis is the −log 10 of the *P*-values for association with the phenotype. The vertical lines correspond to QTLs found by at least 2 models. Point size correlates to −log10(*P*-value). Quantile-quantile (QQ) plots are provided for c) the whole panel and d) the Andean panel.

number (Kamfwa *et al.* 2015). Due to the marker technology used, the QTL for seed weight was large so had low resolution (Elias *et al.* 2021). Our results (Fig. 4) suggest a correlation between DTF, determinacy, and PS under the same QTL. The significant MTA for this QTL was within the gene *Phvul.008G048300*. However, the function of this gene is currently unclear.

The other QTL found on Pv08 is PS8.2, which has previously been identified for seed weight (Blair *et al.* 2006). Genes within this QTL include *Phvul.008G085000*, *Phvul.008G084500*, *Phvul.008G084900*, and *Phvul.008G084100*. *Phvul.008G085000* is homologous to *gibberellin 2-oxidase 8* in *Arabidopsis* (Huala *et al.* 2001). Gibberellin oxidases may respond to light intensity, and can therefore be related to PS (Zhang *et al.* 2022). *Phvul.008G084100* is homologous to *CLAVATA3* in *Arabidopsis*, a gene that regulates shoot and floral meristem development (Clark *et al.* 1995; Hirakawa 2021). *Phvul.008G084900* is homologous to genes encoding ovate family proteins (OFPs). OFPs appear to be sensitive to light stimuli (Shahzaib *et al.* 2024). *Phvul.008G084500* has homology with *RAVEN/INDETERMINATE DOMAIN5* in *Arabidopsis*, which is linked to GA signaling pathways as well as other plant developmental pathways (Sanchez-Corrionero *et al.* 2019; Aoyanagi *et al.* 2020). *Phvul.008G085000* and *Phvul.008G084900* also both contain insertions or deletions with high-impact nonsynonymous mutations which, therefore, possibly affect function (Cingolani *et al.* 2012).

Two QTLs in chromosome 9

A QTL was identified on Pv09 in the Andean panel (PS9.1). This was near a QTL associated with grain yield (Elias *et al.* 2021), postharvest index (Sedlar *et al.* 2020), shoot biomass (Kamfwa *et al.* 2019), SS (da Silva *et al.* 2018), DTF, and yield (Blair *et al.* 2006). Genes within the QTL included *Phvul.009G229100*, *Phvul.009G229200*, *Phvul.009G229700*, and *Phvul.009G229900*. *Phvul.009G229100* is homologous to PIN3 transcription factor genes, involved in regulating root and shoot growth (Goodstein *et al.* 2012; Haga and Sakai 2012). Homology with *Arabidopsis* suggests *Phvul.009G229200* and *Phvul.009G229700* are involved in root growth (Huala *et al.* 2001), and that *Phvul.009G229900* encodes a *HAB1 (Hypersensitive To Aba1) homology to ABI (Abscisic Acid-Insensitive)1* gene involved in ABA signal transduction, which is regulated by circadian rhythm (Leitao, Santos, *et al.* 2021; Kamrani *et al.* 2022). The other QTL in PV09 (PS9.2) was found in the whole panel and included the gene *Phvul.009G145100*, which was also related to an ABA response gene in *Arabidopsis*. A nearby QTL to PS9.2 was previously identified for DTF (Keller *et al.* 2022).

QTL PS11.1 in chromosome 11

The QTL at PV11 (PS11.1) was near a QTL for seed weight (da Silva *et al.* 2018) and a QTL for disease resistance (Banoo *et al.* 2020). This may be due to pleiotropic effects or low resolution of the previous analysis with a limited number of markers. Within this QTL is the gene *Phvul.011G004000* which encodes a putative PHD finger protein. PHDs have been found to be involved in the regulation of flowering time (Zhou *et al.* 2019; Qian *et al.* 2021). Other genes within the QTL are related to root or shoot growth. For example, homology of *Phvul.011G003200* and *Phvul.011G003400* implicates them in processes involved in root meristem development (Huala *et al.* 2001). *Phvul.011G003700* is an uncharacterized gene in common bean but homology with *Arabidopsis* suggests it may be associated with phytochrome interacting factor 7 (PIF7) to regulate hypocotyl elongation (Huala *et al.* 2001; Leivar *et al.* 2008). However, there are many genes within this QTL and further research is needed to clearly distinguish a candidate gene.

Conclusion

Our common bean panel contains genetic diversity from the Andean (4 subgroups) and Mesoamerican (2 subgroups) gene pools. Including accessions from Colombia that contain introgressive hybridization and admixture diversity from the Andean and Mesoamerican gene pools. There was a systematic association between the population structure and agronomic traits such as determinacy and PS. In this study we identified genomic regions which are connected to known and novel putative candidate genes involved in developmental and reproductive pathways. We found 13 QTLs associated with determinacy and 10 QTLs associated with PS. One known QTL was the *Fin* locus on Pv01 for determinacy known for its pleiotropic effects in plant development. While other putative candidate genes were identified due to homology with *Glycine soja*, *Vigna* species and *Arabidopsis*. This includes *Phvul.008G170000* that encodes a putative FAF domain-containing protein. Consequently, GWAS are important in identifying MTAs and candidate genes, especially when accounting for population structure. By linking candidate genes to phenotypes, we hope more targeted precision breeding approaches can be adopted to improve common bean traits under climate change. Nevertheless, this current study and previous ones highlight that for some genes and genomic regions, this will be difficult due to the high proportion of pleiotropic effects in common beans.

Data availability statement

We thank CIAT's Genebank and IPK's Genebank for their generous provision of germplasm. Germplasm held in the CIAT and IPK collections is available on request. Raw reads are deposited in the SRA under accession PRJEB81566. The scripts used in this study are publicly available in Github (https://github.com/DeVegaGroup/KDJ-CBeans/).

Supplemental material available at G3 online.

Acknowledgments

The authors would like to acknowledge the support of the Norwich Bioscience Institutes Research Computing team, Horticultural Services at the John Innes Centre, and the Technical Genomics group at the Earlham Institute, as well as the Genebank personnel at the International Centre of Tropical Research (CIAT) in Colombia and IPK in Germany. All the authors contributed and approved this manuscript.

Funding

KED.-J is supported by the Biotechnology and Biological Sciences Research Council (UKRI-BBSRC) to the Norwich Research Park Doctoral Training program (#2578607). This research was partially funded by the British Council throughout the "2019 Newton Fund Institutional Links binational Bioeconomy" call in grant ID 527023146 to AJC and JJDV. This study was also partially funded by the Biotechnology and Biological Sciences Research Council (BBSRC), part of UK Research and Innovation (UKRI), to the Earlham Institute's Grant "Decoding Biodiversity" (BBX011089/1), and its constituent work package BBS/E/ER/230002B (Decode WP2 Genome Enabled Analysis of Diversity to Identify Gene Function, Biosynthetic Pathways, and Variation in Agri/Aquacultural Traits). Funding was also received from the BBSRC through the "Core Strategic Programme Grant" BB/CSP1720/1 (Genomes to Food Security) and its constituent work package BBS/E/T/000PR9818 (WP1 Signatures of Domestication and Adaptation), as well as the BBSRC-funded "Core Capability Grant" BB/CCG1720/1.

Conflicts of interest

The author(s) declare no conflict of interest.

Literature cited

Alexander DH, Lange K. 2011. Enhancements to the admixture algorithm for individual ancestry estimation. Bmc Bioinformatics. 12(1):246. doi:10.1186/1471-2105-12-246.

Ambachew D, Londono JM, Castillo NR, Asfaw A, Blair MW. 2024. Genetic diversity, linkage disequilibrium, and population structure in a common bean reference collection. Agronomy. 14(5): 985. doi:10.3390/agronomy14050985.

Aoyanagi T, Ikeya S, Kobayashi A, Kozaki A. 2020. Gene regulation via the combination of transcription factors in the indeterminate domain and gras families. Genes (Basel). 11(6):613. doi:10.3390/genes11060613.

Arriagada O, Arévalo B, Cabeza RA, Carrasco B, Schwember AR. 2022. Meta-qtl analysis for yield components in common bean (*phaseolus vulgaris* l.). Plants (Basel). 12(1):117. doi:10.3390/plants12010117.

Assefa T, Mahama AA, Brown AV, Cannon EKS, Rubyogo JC, Rao IM, Blair MW, Cannon SB. 2019. A review of breeding objectives, genomic resources, and marker-assisted methods in common bean (*phaseolus vulgaris* l.). Mol Breed. 39(2):1–23. doi:10.1007/s11032-018-0920-0.

Banoo A, Nabi A, Rasool RS, Mahiya-Farooq, Shah MHD, Ahmad M, Sofi PA, Aasiya-Nabi, Itoo H, Sharma PN, *et al.* 2020. North-western himalayan common beans: population structure and mapping of quantitative anthracnose resistance through genome wide association study. Front Plant Sci. 11:571618. doi:10.3389/fpls.2020.571618.

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. 2011. Bamtools: a c++ api and toolkit for analyzing and managing bam files. Bioinformatics. 27(12):1691–1692. doi:10.1093/bioinformatics/btr174.

Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Cukura A, Denny P, *et al.* 2023. Uniprot: the universal protein knowledgebase in 2023. Nucleic Acids Res. 51(D1):D523–D531. doi:10.1093/nar/gkac1052.

Beans and herbs. 2020. [accessed 2020]. https://www.beansandherbs.co.uk/.

Beebe S, Rao I, Mukankusi C, Buruchara R. 2012. Improving resource use efficiency and reducing risk of common bean production in Africa, Latin America, and the Caribbean. In: Hershey Clair H, editor. Eco-Efficiency: From vision to reality. Vol. 8. Centro Internacional de Agricultura Tropical (CIAT). p. 1–18. https://cgspace.cgiar.org/items/57ddcb4c-35a3-486a-ad28-eddd57443d41.

Beebe S, Skroch PW, Tohme J, Duque MC, Pedraza F, Nienhuis J. 2000. Structure of genetic diversity among common bean landraces of middle American origin based on correspondence analysis of rapd. Crop Sci. 40(1):264–273. doi:10.2135/cropsci2000.401264x.

Bellucci E, Benazzo A, Xu CM, Bitocchi E, Rodriguez M, Alseekh S, Di Vittori V, Gioia T, Neumann K, Cortinovis G, *et al.* 2023. Selection and adaptive introgression guided the complex evolutionary history of the European common bean. Nat Commun. 14(1):1908. doi:10.1038/s41467-023-37332-z.

Bernardi C, Macri G, Biagi M, Miraldi E, Finetti F, Trabalzini L. 2023. In vitro digestion of *phaseolus vulgaris* l. Cooked beans induces autophagy in colon cancer cells. Foods. 12(4):839. doi:10.3390/foods12040839.

Bhakta MS, Jones VA, Vallejos CE. 2015. Punctuated distribution of recombination hotspots and demarcation of pericentromeric regions in *phaseolus vulgaris* l. PLoS One. 10(1):e0116822. doi:10.1371/journal.pone.0116822.

Bitocchi E, Nanni L, Bellucci E, Rossi M, Giardini A, Zeuli PS, Logozzo G, Stougaard J, McClean P, Attene G, *et al.* 2012. Mesoamerican origin of the common bean (*phaseolus vulgaris* l.) is revealed by sequence data. Proc Natl Acad Sci U S A. 109(14):E788–E796. doi:10.1073/pnas.1108973109.

Bitocchi E, Rau D, Bellucci E, Rodriguez M, Murgia ML, Gioia T, Santo D, Nanni L, Attene G, Papa R. 2017. Beans (*phaseolus* ssp.) as a model for understanding crop evolution. Front Plant Sci. 8:722. doi:10.3389/fpls.2017.00722.

Blair MW, Chaves A, Tofino A, Calderon JF, Palacio JD. 2010. Extensive diversity and inter-genepool introgression in a world-wide collection of indeterminate snap bean accessions. Theor Appl Genet. 120(7):1381–1391. doi:10.1007/s00122-010-1262-4.

Blair MW, Cortes AJ, Penmetsa RV, Farmer A, Carrasquilla-Garcia N, Cook DR. 2013. A high-throughput snp marker system for parental polymorphism screening, and diversity analysis in common bean (*phaseolus vulgaris* l.). Theor Appl Genet. 126(2):535–548. doi:10.1007/s00122-012-1999-z.

Blair MW, Diaz LM, Buendia HF, Duque MC. 2009. Genetic diversity, seed size associations and population structure of a core collection of common beans (*phaseolus vulgaris* l.). Theor Appl Genet. 119(6):955–972. doi:10.1007/s00122-009-1064-8.

Blair MW, Diaz JM, Hidalgo R, Diaz LM, Duque MC. 2007. Microsatellite characterization of andean races of common bean (*phaseolus vulgaris* l.). Theor Appl Genet. 116(1):29–43. doi:10.1007/s00122-007-0644-8.

Blair MW, Iriarte G, Beebe S. 2006. Qtl analysis of yield traits in an advanced backcross population derived from a cultivated andean × wild common bean (*phaseolus vulgaris* l.) cross. Theor Appl Genet. 112(6):1149–1163. doi:10.1007/s00122-006-0217-2.

Blair MW, Izquierdo P, Astudillo C, Grusak MA. 2013. A legume biofortification quandary: variability and genetic control of seed coat micronutrient accumulation in common beans. Front Plant Sci. 4:275. doi:10.3389/fpls.2013.00275.

Botero H, Barnes AP. 2022. The effect of enso on common bean production in Colombia: a time series approach. Food Secur. 14(6): 1417–1430. doi:10.1007/s12571-022-01290-z.

Botero H, Barnes AP, Perez L, Rios D, Ramirez-Villegas J. 2021. The determinants of common bean variety selection and diversification in Colombia. Ecol Econ. 190:107181. doi:10.1016/j.ecolecon.2021.107181.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. Tassel: software for association mapping of complex traits in diverse samples. Bioinformatics. 23(19):2633–2635. doi:10.1093/bioinformatics/btm308.

Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J. 2003. Beans (*phaseolus* spp.)—model food legumes. Plant Soil. 252(1):55–128. doi:10.1023/A:1024146710611.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. Blast+: architecture and applications. BMC Bioinformatics. 10(421). doi:10.1186/1471-2105-10-421.

Campa A, Murube E, Ferreira JJ. 2018. Genetic diversity, population structure, and linkage disequilibrium in a Spanish common bean diversity panel revealed through genotyping-by-sequencing. Genes (Basel). 9(11):518. doi:10.3390/genes9110518.

Caproni L, Raggi L, Talsma EF, Wenzl P, Negri V. 2020. European landrace diversity for common bean biofortification: a genome-wide association study. Sci Rep. 10(1):19775. doi:10.1038/s41598-020-76417-3.

Castro-Guerrero NA, Isidra-Arellano M, Mendoza-Cozatl DG, Valdes-Lopez O. 2016. Common bean: a legume model on the rise for unraveling responses and adaptations to iron, zinc, and phosphate deficiencies. Front Plant Sci. 7:600. doi:10.3389/fpls.2016.00600.

Cavalcante AG, Lemos LB, Meirelles FC, Cavalcante ACP, de Aquino LA. 2020. Thermal sum and phenological descriptions of growth

stages of the common bean according to the bbch scale. Ann Appl Biol. 176(3):342–349. doi:10.1111/aab.12571.
Cebeci Z, Bayraktar M, Gökçe G. 2023. Comparison of the statistical methods for genome-wide association studies on simulated quantitative traits of domesticated goats (*capra hircus* l.). Small Rumin Res. 227:107053. doi:10.1016/j.smallrumres.2023.107053.
Chacon-Sanchez MI, Martinez-Castillo J, Duitama J, Debouck DG. 2021. Gene flow in *phaseolus* beans and its role as a plausible driver of ecological fitness and expansion of cultigens. Front Ecol Evol. 9:618709. doi:10.3389/fevo.2021.618709.
Chacon-Sanchez MI, Pickersgill B, Debouck DG, Arias JS. 2007. Phylogeographic analysis of the chloroplast dna variation in wild common bean (*phaseolus vulgaris* l.) in the americas. Plant Systematics and Evolution. 266(3–4):175–195. doi:10.1007/s00606-007-0536-z.
Chen F, Yang YZ, Luo XF, Zhou WG, Dai YJ, Zheng C, Liu WG, Yang WY, Shu K. 2019. Genome-wide identification of grf transcription factors in soybean and expression analysis of gmgrf family under shade stress. BMC Plant Biol. 19(1):269. doi:10.1186/s12870-019-1861-4.
Cichy KA, Porch TG, Beaver JS, Cregan P, Fourie D, Glahn RP, Grusak MA, Kamfwa K, Katuuramu DN, McClean P, *et al.* 2015. A diversity panel for andean bean improvement. Crop Sci. 55(5):2149–2160. doi:10.2135/cropsci2014.09.0653.
Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu XY, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff. Fly (Austin). 6(2):80–92. doi:10.4161/fly.19695.
Clark SE, Running MP, Meyerowitz EM. 1995. Clavata3 is a specific regulator of shoot and floral meristem development affecting the same processes as clavata1. Development. 121(7):2057–2067. doi:10.1242/dev.121.7.2057.
Coelho AP, de Faria RT, Lemos LB, dos Reis MAM, Filla VA, Bertino AMP. 2023. Irrigation management of common bean cultivars with contrasting growth habits. Scientia Agricola. 80:e20220038. doi:10.1590/1678-992x-2022-0038.
Cusworth G, Garnett T, Lorimer J. 2021. Agroecological break out: legumes, crop diversification and the regenerative futures of UK agriculture. J Rural Stud. 88:126–137. doi:10.1016/j.jrurstud.2021.10.005.
Daba K, Warkentin TD, Bueckert R, Todd CD, Tar'an B. 2016. Determination of photoperiod-sensitive phase in chickpea (*cicer arietinum* l.). Front Plant Sci. 7:478. doi:10.3389/fpls.2016.00478.
Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, *et al.* 2021. Twelve years of samtools and bcftools. Gigascience. 10(2):giab008. doi:10.1093/gigascience/giab008.
da Silva LC, de Souza TLPO, Cruz CD, Carneiro PCS, Silva FFE, de Barros EG, Vianello RP, da Fonseca CEL, Song QJ, Cregan PB, *et al.* 2018. Linkage fine-mapping and qtls affecting morpho-agronomic traits of a mesoamericanxandean ril common bean population. Euphytica. 214(12):1–15. doi:10.1007/s10681-018-2299-8.
de Almeida CP, Paulino JFD, Carbonell SAM, Chiorato AF, Song QJ, Di Vittori V, Rodriguez M, Papa R, Benchimol-Reis LL. 2020. Genetic diversity, population structure, and andean introgression in Brazilian common bean cultivars after half a century of genetic breeding. Genes (Basel). 11(11):1298. doi:10.3390/genes11111298.
Debouck DG, Toro O, Paredes OM, Johnson WC, Gepts P. 1993. Genetic diversity and ecological distribution of *phaseolus-vulgaris (Fabaceae)* in northwestern south-America. Econ Bot. 47(4):408–423. doi:10.1007/BF02907356.
Delfini J, Moda-Cirino V, Neto JD, Zeffa DM, Nogueira AF, Ribeiro LAB, Ruas PM, Gepts P, Gonçalves LSA. 2021. Genome-wide association study identifies genomic regions for important morpho-agronomic traits in mesoamerican common bean. Front Plant Sci. 12:748829. doi:10.3389/fpls.2021.748829.
de Souza LP, Bitocchi E, Papa R, Tohge T, Fernie AR. 2023. Decreased metabolic diversity in common beans associated with domestication revealed by untargeted metabolomics, information theory, and molecular networking. Plant J. 115(4):1021–1036. doi:10.1111/tpj.16277.
Diesh C, Stevens GJ, Xie PT, Martinez TD, Hershberg EA, Leung A, Guo E, Dider S, Zhang JJ, Bridge C *et al.* 2023. Jbrowse 2: a modular genome browser with views of synteny and structural variation. Genome Biol. 24(1):74. doi:10.1186/s13059-023-02914-z.
Diniz AL, Giordani W, Costa ZP, Margarido GRA, Perseguini JMKC, Benchimol-Reis LL, Chiorato AF, Garcia AAF, Vieira MLC. 2018. Evidence for strong kinship influence on the extent of linkage disequilibrium in cultivated common beans. Genes (Basel). 10(1):5. doi:10.3390/genes10010005.
Elias JCF, Gonçalves-Vidigal MC, Bisneta MV, Valentini G, Vidigal PS, Gilio TAS, Moda-Cirino V, Song QJ. 2021. Genetic mapping for agronomic traits in iapar 81/lp97–28 population of common bean (phaseolus vulgaris l.) under drought conditions. Plants (Basel). 10(8):1568. doi:10.3390/plants10081568.
Ewels P, Magnusson M, Lundin S, Kaller M. 2016. Multiqc: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 32(19):3047–3048. doi:10.1093/bioinformatics/btw354.
Foyer CH, Lam HM, Nguyen HT, Siddique KHM, Varshney RK, Colmer TD, Cowling W, Bramley H, Mori TA, Hodgson JM, *et al.* 2016. Neglecting legumes has compromised human health and sustainable food production. Nat Plants. 2(8):16112. doi:10.1038/nplants.2016.112.
Gan ES, Xu YF, Wong JY, Goh JG, Sun B, Wee WY, Huang JB, Ito T. 2014. Jumonji demethylases moderate precocious flowering at elevated temperature via regulation of *flc* in *arabidopsis*. Nat Commun. 5(1):5098. doi:10.1038/ncomms6098.
Ganesan K, Xu BJ. 2017. Polyphenol-rich dry common beans (*phaseolus vulgaris* l.) and their health benefits. Int J Mol Sci. 18(11):2331. doi:10.3390/ijms18112331.
García-Fernández C, Campa A, Garzón AS, Miklas P, Ferreira JJ. 2021. Gwas of pod morphological and color characters in common bean. BMC Plant Biol. 21(1):184. doi:10.1186/s12870-021-02967-x.
Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv 1207.3907. https://doi.org/10.48550/arXiv.1207.3907, preprint: not peer reviewed.
Gautam T, Dutta M, Jaiswal V, Zinta G, Gahlaut V, Kumar S. 2022. Emerging roles of sweet sugar transporters in plant development and abiotic stress responses. Cells. 11(8):1303. doi:10.3390/cells11081303.
Gepts P, Debouck D. 1991. Origin, domestication, and evolution of the common bean (*phaseolus vulgaris* l.). In: van Schoonhoven A, Voysest O, editors. Common Beans: Research for Crop Improvement. C.A.B. Int.; Wallingford, UK: CIAT; Cali, Colombia. p. 7–53.
Ginestet C. 2011. Ggplot2: elegant graphics for data analysis. Journal of the Royal Statistical Society Series a-Statistics in Society. 174(1):245–246. doi:10.1111/j.1467-985X.2010.00676_9.x.
Giordani W, Gama HC, Chiorato AF, Garcia AAF, Vieira MLC. 2022. Genome-wide association studies dissect the genetic architecture of seed shape and size in common bean. G3 (Bethesda). 12(4):jkac048. doi:10.1093/g3journal/jkac048.
González AM, Yuste-Lisbona FJ, Saburido S, Bretones S, De Ron AM, Lozano R, Santalla M. 2016. Major contribution of flowering time and vegetative growth to plant production in common

bean as deduced from a comparative genetic mapping. Front Plant Sci. 7:1940. doi:10.3389/fpls.2016.01940.
Goodstein DM, Shu SQ, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, *et al.* 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40(D1):D1178–D1186. doi:10.1093/nar/gkr944.
Gori B, Ulian T, Bernal HY, Diazgranados M. 2022. Understanding the diversity and biogeography of Colombian edible plants. Sci Rep. 12(1):7835. doi:10.1038/s41598-022-11600-2.
Haga K, Sakai T. 2012. Pin auxin efflux carriers are necessary for pulse-induced but not continuous light-induced phototropism in *arabidopsis*. Plant Physiol. 160(2):763–776. doi:10.1104/pp.112.202432.
Hill WG, Robertson A. 2007. The effect of linkage on limits to artificial selection (reprinted). Genet Res (Camb). 89(5–6):311–336. doi:10.1017/S001667230800949X.
Hirakawa Y. 2021. Clavata3, a plant peptide controlling stem cell fate in the meristem. Peptides. 142:170579. doi:10.1016/j.peptides.2021.170579.
Hoyos-Villegas V, Song QJ, Kelly JD. 2017. Genome-wide association analysis for drought tolerance and associated traits in common bean. Plant Genome. 10(1):2015.12.0122. doi:10.3835/plantgenome2015.12.0122.
Hu TS, Chitnis N, Monos D, Dinh A. 2021. Next-generation sequencing technologies: an overview. Hum Immunol. 82(11):801–811. doi:10.1016/j.humimm.2021.02.012.
Huala E, Dickerman AW, Garcia-Hernandez M, Weems D, Reiser L, LaFond F, Hanley D, Kiphart D, Zhuang MZ, Huang W, *et al.* 2001. The arabidopsis information resource (tair): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res. 29(1):102–105. doi:10.1093/nar/29.1.102.
Huang M, Liu XL, Zhou Y, Summers RM, Zhang ZW. 2019. Blink: a package for the next level of genome-wide association studies with both individuals and markers in the millions. Gigascience. 8(2):giy154. doi:10.1093/gigascience/giy154.
Humann J, Jung S, Cheng C-H, Lee T, Zheng P, Frank M, McGaughey D, Scott K, Buble K, Yu J, *et al.* 2019. Cool Season Food Legume Genome Database: A Resource for pea, Lentil, faba Bean and Chickpea Genetics, Genomics and Breeding. Paper Presented at: International Plant and Animal Genome Conference; San Diego, CA, USA.
Jha AB, Ashokkumar K, Diapari M, Ambrose SJ, Zhang HX, Tar'an B, Bett KE, Vandenberg A, Warkentin TD, Purves RW. 2015. Genetic diversity of folate profiles in seeds of common bean, lentil, chickpea and pea. J Food Compost Anal. 42:134–140. doi:10.1016/j.jfca.2015.03.006.
Jungleseeds. 2020. Seed catalogue [accessed 2020 Jun]. https://jungleseeds.co.uk/
Kachinski WD, Avila FW, Dos Reis AR, Muller MML, Mendes MC, Petranski PH. 2022. Agronomic biofortification increases concentrations of zinc and storage proteins in common bean (*phaseolus vulgaris* l.) grains. Food Res Int. 155:111105. doi:10.1016/j.foodres.2022.111105.
Kamfwa K, Cichy KA, Kelly JD. 2015. Genome-wide association study of agronomic traits in common bean. Plant Genome. 8(2):2014.09.0059. doi:10.3835/plantgenome2014.09.0059.
Kamfwa K, Cichy KA, Kelly JD. 2019. Identification of quantitative trait loci for symbiotic nitrogen fixation in common bean. Theor Appl Genet. 132(5):1375–1387. doi:10.1007/s00122-019-03284-6.
Kami J, Velasquez VB, Debouck DG, Gepts P. 1995. Identification of presumed ancestral dna-sequences of phaseolin in *phaseolus-vulgaris*. Proc Natl Acad Sci U S A. 92(4):1101–1104. doi:10.1073/pnas.92.4.1101.
Kamrani YY, Shomali A, Aliniaeifard S, Lastochkina O, Moosavi-Nezhad M, Hajinajaf N, Talar U. 2022. Regulatory role of circadian clocks on aba production and signaling, stomatal responses, and water-use efficiency under water-deficit conditions. Cells. 11(7):1154. doi:10.3390/cells11071154.
Keller B, Ariza-Suarez D, Portilla-Benavides AE, Buendia HF, Aparicio JS, Amongi W, Mbiu J, Msolla SN, Miklas P, Porch TG, *et al.* 2022. Improving association studies and genomic predictions for climbing beans with data from bush bean populations. Front Plant Sci. 13:830896. doi:10.3389/fpls.2022.830896.
Kim J, Kang H, Park J, Kim W, Yoo J, Lee N, Kim J, Yoon TY, Choi G. 2016. Pif1-interacting transcription factors and their binding sequence elements determine the in vivo targeting sites of pif1. Plant Cell. 28(6):1388–1405. doi:10.1105/tpc.16.00125.
Kirkwood M, Vohra P, Bawn M, Thilliez G, Pye H, Tanner J, Chintoan-Uta C, Branchu P, Petrovska L, Dallman T, *et al.* 2021. Ecological niche adaptation of *salmonella* typhimurium u288 is associated with altered pathogenicity and reduced zoonotic potential. Commun Biol. 4(1):498. doi:10.1038/s42003-021-02013-4.
Kwak M, Toro O, Debouck DG, Gepts P. 2012. Multiple origins of the determinate growth habit in domesticated common bean (*phaseolus vulgaris*). Ann Bot. 110(8):1573–1580. doi:10.1093/aob/mcs207.
Kwak M, Velasco D, Gepts P. 2008. Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (*phaseolus vulgaris*). J Hered. 99(3):283–291. doi:10.1093/jhered/esn005.
Leitao ST, Bicho MC, Pereira P, Paulo MJ, Malosetti M, Araujo SD, van Eeuwijk F, Patto MCV. 2021. Common bean snp alleles and candidate genes affecting photosynthesis under contrasting water regimes. Hortic Res. 8(1):4. doi:10.1038/s41438-020-00434-6.
Leitao ST, Santos C, Araújo SD, Rubiales D, Vaz Patto MC. 2021. Shared and tailored common bean transcriptomic responses to combined fusarium wilt and water deficit. Hortic Res. 8(1):149. doi:10.1038/s41438-021-00583-2.
Leivar P, Monte E, Al-Sady B, Carle C, Storer A, Alonso JM, Ecker JR, Quail PH. 2008. The *arabidopsis* phytochrome-interacting factor pif7, together with pif3 and pif4, regulates responses to prolonged red light by modulating phyb levels. Plant Cell. 20(2):337–352. doi:10.1105/tpc.107.052142.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25(14):1754–1760. doi:10.1093/bioinformatics/btp324.
Liu P, Zhang SB, Zhou B, Luo X, Zhou XF, Cai B, Jin YH, Niu D, Lin JX, Cao XF, *et al.* 2019. The histone h3k4 demethylase jmj16 represses leaf senescence in arabidopsis. Plant Cell. 31(2):430–443. doi:10.1105/tpc.18.00693.
Liu X, Huang M, Fan B, Buckler ES, Zhang Z. 2016. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies (vol 12, e1005767, 2016). PLoS One. 11(3):e1005767. doi:10.1371/journal.pgen.1005767.
Mamidi S, Rossi M, Moghaddam SM, Annam D, Lee R, Papa R, McClean PE. 2013. Demographic factors shaped diversity in the two gene pools of wild common bean *phaseolus vulgaris* l. Heredity (Edinb). 110(3):267–276. doi:10.1038/hdy.2012.82.
Mara CD, Huang TB, Irish VF. 2010. The floral homeotic proteins apetala3 and pistillata negatively regulate the genes implicated in light signaling. Plant Cell. 22(3):690–702. doi:10.1105/tpc.109.065946.
Merrick LF, Burke AB, Zhang ZW, Carter AH. 2022. Comparison of single-trait and multi-trait genome-wide association models and inclusion of correlated traits in the dissection of the genetic architecture of a complex trait in a breeding program. Front Plant Sci. 12:772907. doi:10.3389/fpls.2021.772907.

Miklas PN, Johnson WC, Delorme R, Gepts P. 2001. Qtl conditioning physiological resistance and avoidance to white mold in dry bean. Crop Sci. 41(2):309–315. doi:10.2135/cropsci2001.412309x.
Moghaddam SM, Mamidi S, Osorno JM, Lee R, Brick M, Kelly J, Miklas P, Urrea C, Song QJ, Cregan P, *et al.* 2016. Genome-wide association study identifies candidate loci underlying agronomic traits in a middle American diversity panel of common bean. Plant Genome. 9(3):2016.02.0012. doi:10.3835/plantgenome2016.02.0012.
Munekage Y, Hojo M, Meurer J, Endo T, Tasaka M, Shikanai T. 2002. *Pgr5* is involved in cyclic electron flow around photosystem i and is essential for photoprotection in *arabidopsis*. Cell. 110(3): 361–371. doi:10.1016/S0092-8674(02)00867-X.
Mupangwa W, Nyagumbo I, Liben F, Chipindu L, Craufurd P, Mkuhlani S. 2021. Maize yields from rotation and intercropping systems with different legumes under conservation agriculture in contrasting agro-ecologies. Agricul Ecosys Environ. 306: 107170. doi:10.1016/j.agee.2020.107170.
Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature. 403(6772):853–858. doi:10.1038/35002501.
Mylona P, Pawlowski K, Bisseling T. 1995. Symbiotic nitrogen-fixation. Plant Cell. 7(7):869–885. doi:10.2307/3870043.
Patto MCV, Amarowicz R, Aryee ANA, Boye JI, Chung HJ, Martin-Cabrejas MA, Domoney C. 2015. Achievements and challenges in improving the nutritional quality of food legumes. CRC Crit Rev Plant Sci. 34(1–3):105–143. doi:10.1080/07352689.2014.897907.
Pérez-Vega E, Pañeda A, Rodríguez-Suárez C, Campa A, Giraldez R, Ferreira JJ. 2010. Mapping of qtls for morpho-agronomic and seed quality traits in a ril population of common bean (*phaseolus vulgaris* l.). Theor Appl Genet. 120(7):1367–1380. doi:10.1007/s00122-010-1261-5.
Perez L, Rios DA, Giraldo DC, Twyman J, Blundo-Canto G, Prager SD, Ramirez-Villegas J. 2020. Determinants of vulnerability of bean growing households to climate variability in Colombia. Clim Dev. 12(8):730–742. doi:10.1080/17565529.2019.1685931.
Phiri AT, Njira KOW. 2023. Grain legume-based cropping systems' effects on soil organic carbon and nutrient dynamics. Agric Res. 12(1):45–52. doi:10.1007/s40003-022-00619-6.
Pironon S, Borrell JS, Ondo I, Douglas R, Phillips C, Khoury CK, Kantar MB, Fumia N, Gomez MS, Viruel J, *et al.* 2020. Toward unifying global hotspots of wild and domesticated biodiversity. Plants-Basel. 9(9):1128. doi:10.3390/plants9091128.
Qian F, Zhao QY, Zhang TN, Li YL, Su YN, Li L, Sui JH, Chen S, He XJ. 2021. A histone h3k27me3 reader cooperates with a family of phd finger-containing proteins to regulate flowering time in arabidopsis. J Integr Plant Biol. 63(4):787–802. doi:10.1111/jipb.13067.
Raggi L, Caproni L, Carboni A, Negri V. 2019. Genome-wide association study reveals candidate genes for flowering time variation in common bean (*phaseolus vulgaris* l.). Front Plant Sci. 10:962. doi:10.3389/fpls.2019.00962.
Rawal V, Navarro DK. 2019. The Global Economy of Pulses. Rome: FAO.
Reinprecht Y, Schram L, Perry GE, Morneau E, Smith TH, Pauls KP. 2023. Mapping yield and yield-related traits using diverse common bean germplasm. Front Genet. 14:1246904. doi:10.3389/fgene.2023.1246904.
Rendon-Anaya M, Herrera-Estrella A, Gepts P, Delgado-Salinas A. 2017. A new species of *phaseolus* (leguminosae, papilionoideae) sister to *phaseolus vulgaris*, the common bean. Phytotaxa. 313(3): 259–266. doi:10.11646/phytotaxa.313.3.3.
Repinski SL, Kwak M, Gepts P. 2012. The common bean growth habit gene pvtfl1y is a functional homolog of *arabidopsis* TFL1. Theor Appl Genet. 124(8):1539–1547. doi:10.1007/s00122-012-1808-8.
Sablowski R. 2007. Flowering and determinacy in *arabidopsis*. J Exp Bot. 58(5):899–907. doi:10.1093/jxb/erm002.
Sanchez-Corrionero A, Perez-Garcia P, Cabrera J, Silva-Navas J, Perianez-Rodriguez J, Gude I, del Pozo JC, Moreno-Risueno MA. 2019. Root patterning and regeneration are mediated by the quiescent center and involve bluejay, jackdaw and scarecrow regulation of vasculature factors. bioRxiv 803973. https://doi.org/10.1101/803973, preprint: not peer reviewed.
Santalla M, Menendez-Sevillano MC, Monteagudo AB, De Ron AM. 2004. Genetic diversity of argentinean common bean and its evolution during domestication. Euphytica. 135(1):75–87. doi:10.1023/B:EUPH.0000009543.46471.72.
Schier HE, Eliot KA, Herron SA, Landfried LK, Migicovsky Z, Rubin MJ, Miller AJ. 2019. Comparative analysis of perennial and annual *phaseolus* seed nutrient concentrations. Sustainability. 11(10): 2787. doi:10.3390/su11102787.
Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu SQ, Song QJ, Chavarro C, *et al.* 2014. A reference genome for common bean and genome-wide analysis of dual domestications. Nat Genet. 46(7):707–713. doi:10.1038/ng.3008.
Krueger F, James F, Ewels P, Afyounian E, Weinstein M, Schuster-Boeckler B, Hulselmans G, Sclamons. 2023. Felixkrueger/trimgalore: V0.6.10—add default decompression path (0.6.10). Zenodo.
Sedlar A, Zupin M, Maras M, Razinger J, Sustar-Vozlic J, Pipan B, Meglic V. 2020. Qtl mapping for drought-responsive agronomic traits associated with physiology, phenology, and yield in an andean intra-gene pool common bean population. Agronomy. 10(2): 225. doi:10.3390/agronomy10020225.
Shahzaib M, Khan UM, Azhar MT, Atif RM, Khan SH, Zaman QU, Rana IA. 2024. Phylogenomic curation of ovate family proteins (ofps) in the u's triangle of *brassica* l. Indicates stress-induced growth modulation. PLoS One. 19(1):e0297473. doi:10.1371/journal.pone.0297473.
Shinohara H. 2021. Root meristem growth factor rgf, a sulfated peptide hormone in plants. Peptides. 142:170556. doi:10.1016/j.peptides.2021.170556.
Singh SP, Schwartz HF. 2010. Breeding common bean for resistance to diseases: a review. Crop Sci. 50(6):2199–2223. doi:10.2135/cropsci2009.03.0163.
Sitaraman J, Bui M, Liu ZC. 2008. *Leunig_homolog* and *leunig* perform partially redundant functions during arabidopsis embryo and floral development. Plant Physiol. 147(2):672–681. doi:10.1104/pp.108.115923.
Soler-Garzón A, Mulube M, Kamfwa K, Lungu DM, Hamabwe S, Roy J, Salegua V, Fourie D, Porch TG, McClean PE, *et al.* 2024. Gwas of resistance to three bacterial diseases in the andean common bean diversity panel. Front Plant Sci. 15:1469381. doi:10.3389/fpls.2024.1469381.
Suarez-Martinez SE, Ferriz-Martinez RA, Campos-Vega R, Elton-Puente JE, Carbot KD, Garcia-Gasca T. 2016. Bean seeds: leading nutraceutical source for human health. Cyta-Journal of Food. 14(1):131–137. doi:10.1080/19476337.2015.1063548.
Suarez JC, Polania JA, Anzola JA, Contreras AT, Mendez DL, Vanegas JI, Noriega JE, Rodriguez L, Urban MO, Beebe S, *et al.* 2021. Influence of nitrogen supply on gas exchange, chlorophyll fluorescence and grain yield of breeding lines of common bean evaluated in the Amazon region of Colombia. Acta Physiologiae Plantarum. 43(4):1–15. doi:10.1007/s11738-021-03233-1.
Suarez JC, Urban MO, Contreras AT, Grajales MA, Cajiao C, Beebe SE, Rao IM. 2021. Adaptation of interspecific mesoamerican common bean lines to acid soils and high temperature in the Amazon region of Colombia. Plants (Basel). 10(11):2412. doi:10.3390/plants10112412.

Tai APK, Martin MV, Heald CL. 2014. Threat to future global food security from climate change and ozone air pollution. Nat Clim Chang. 4(9):817–821. doi:10.1038/nclimate2317.

Tar'an B, Michaels TE, Pauls KP. 2002. Genetic mapping of agronomic traits in common bean. Crop Sci. 42(2):544–556. doi:10.2135/cropsci2002.5440.

Tohme J, Gonzalez DO, Beebe S, Duque MC. 1996. Aflp analysis of gene pools of a wild bean core collection. Crop Sci. 36(5):1375–1384. doi:10.2135/cropsci1996.0011183X003600050048x.

Tseng TS, Salomé PA, McClung CR, Olszewski NE. 2004. Spindly and gigantea interact and act in *arabidopsis thaliana* pathways involved in light responses, flowering, and rhythms in cotyledon movements. Plant Cell. 16(6):1550–1563. doi:10.1105/tpc.019224.

Ugwuanyi S, Udengwu OS, Snowdon RJ, Obermeier C. 2022. Novel candidate loci for morpho-agronomic and seed quality traits detected by targeted genotyping-by-sequencing in common bean. Front Plant Sci. 13:1014282. doi:10.3389/fpls.2022.1014282.

Valdisser PAMR, Pereira WJ, Almeida JE, Müller BSF, Coelho GRC, de Menezes IPP, Vianna JPG, Zucchi MI, Lanna AC, Coelho ASG, *et al.* 2017. In-depth genome characterization of a Brazilian common bean core collection using dartseq high-density snp genotyping. Bmc Genomics. 18(1):423. doi:10.1186/s12864-017-3805-4.

Van der Auwera G, O'Connor B. 2020. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media.

Vargas Y, Mayor-Duran VM, Buendia HF, Ruiz-Guzman H, Raatz B. 2021. Physiological and genetic characterization of heat stress effects in a common bean RIL population. PLoS One. 16(4):e0249859. doi:10.1371/journal.pone.0249859.

Wahl V, Brand LH, Guo YL, Schmid M. 2010. The fantastic four proteins influence shoot meristem size in *Arabidopsis thaliana*. BMC Plant Biol. 10(1):285. doi:10.1186/1471-2229-10-285.

Wang JB, Zhang ZW. 2021. Gapit version 3: boosting power and accuracy for genomic association and prediction. Genom Proteomic Bioinformat. 19(4):629–640. doi:10.1016/j.gpb.2021.08.005.

Wei T, Simko V. 2021. R package 'Corrplot': Visualization of a Correlation Matrix. 0.92 ed. [accessed 2024 Jun]. https://github.com/taiyun/corrplot

Weller JL, Schoor JKV, Perez-Wright EC, Hecht V, Gonzalez AM, Capel C, Yuste-Lisbona FJ, Lozano R, Santalla M. 2019. Parallel origins of photoperiod adaptation following dual domestications of common bean. J Exp Bot. 70(4):1209–1219. doi:10.1093/jxb/ery455.

Wingett SW, Andrews S. 2018. FastQ Screen: A tool for multi-genome mapping and quality control. F1000Research. 7:1338. doi:10.12688/f1000research.15931.2.

Wu L, Chang YJ, Wang LF, Wang SM, Wu J. 2024. Genome-wide association study dissecting drought resistance- associated loci based on physiological traits in common bean. J Integr Agric. 23(11):3657–3671. doi:10.1016/j.jia.2024.03.079.

Wu J, Wang LF, Fu JJ, Chen JB, Wei SH, Zhang SL, Zhang J, Tang YS, Chen ML, Zhu JF, *et al.* 2020. Resequencing of 683 common bean genotypes identifies yield component trait associations across a north-south cline. Nat Genet. 52(1):118–125. doi:10.1038/s41588-019-0546-0.

Wu JF, Wang Y, Wu SH. 2008. Two new clock proteins, lwd1 and lwd2, regulate arabidopsis photoperiodic flowering. Plant Physiol. 148(2):948–959. doi:10.1104/pp.108.124917.

Xin XY, Li PR, Zhao XY, Yu YJ, Wang WH, Jin GH, Wang J, Sun LL, Zhang DS, Zhang FL, *et al.* 2024. Temperature-dependent jumonji demethylase modulates flowering time by targeting h3k36me2/3 in. Nat Commun. 15(1):5470. doi:10.1038/s41467-024-49721-z.

Yamaguchi N. 2021. Removal of h3k27me3 by jmj proteins controls plant development and environmental responses in. Front Plant Sci. 12:687416. doi:10.3389/fpls.2021.687416.

Zhang XY, Blair MW, Wang SM. 2008. Genetic diversity of Chinese common bean (*phaseolus vulgaris* l.) landraces assessed with simple sequence repeat markers. Theor Appl Genet. 117(4):629–640. doi:10.1007/s00122-008-0807-2.

Zhang C, Dong SS, Xu JY, He WM, Yang TL. 2019. Poplddecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics. 35(10):1786–1788. doi:10.1093/bioinformatics/bty875.

Zhang ZW, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu JM, Arnett DK, Ordovas JM, *et al.* 2010. Mixed linear model approach adapted for genome-wide association studies. Nat Genet. 42(4):355–360. doi:10.1038/ng.546.

Zhang CH, Nie X, Kong WL, Deng XX, Sun T, Liu XH, Li YS. 2022. Genome-wide identification and evolution analysis of the gibberellin oxidase gene family in six gramineae crops. Genes (Basel). 13(5):863. doi:10.3390/genes13050863.

Zhou AM, Sun HW, Dai SY, Feng S, Zhang JZ, Gong SF, Wang JG. 2019. Identification of transcription factors involved in the regulation of flowering in *Adonis amurensis* through combined RNA-seq transcriptomics and iTRAQ proteomics. Genes (Basel). 10(4):305. doi:10.3390/genes10040305.

Editor: S. Pearce

Glossary

CWR	Crop wild relatives
PI	Photoperiod insensitive
PS	Photoperiod sensitive
D	Determinacy
GH	Growth habit
ENSO	El Niño-Southern Oscillation
GWAS	Genome-wide association mapping
QTL	Quantitative trait loci
MTA	Marker trait association
LD	Linkage disequilibrium
DTF	Days to flowering
SS	Seed size
E100 SW	Estimated seed weight
BLINK	Bayesian-information and Linkage- disequilibrium Iteratively Nested Keyway
FarmCPU	Fixed and random model Circulating Probability Unification
MLM	Mixed Linear Model
MLMM	Multiple Locus Mixed Linear Model
A	Andean
M	Mesoamerican
Admx	Admixed
DAS	Days after sowing

HOMR	Homozygous reference
HOMR	Homozygous alternative
SG	Stay-green
g_{sw}	Stomatal conductance to water vapour
***E*_apparent**	Transpiration
VPDleaf	Leaf vapour pressure deficit
rh_s	relative humidity of the sample
Tleaf	Leaf temperature
ETR	Electron transport rate
Fs	Steady-state fluorescence
Fm'	Maximum fluorescence
PhiPS2	Quantum yield of fluorescence
WC	Water content
Bulk EC	Bulk electrical conductivity
MP	Soil matric potential
ST	Soil temperature

References

Aguilar OM, Collavino MM, Mancini U. 2022. Nodulation competitiveness and diversification of symbiosis genes in common beans from the american centers of domestication. Scientific Reports. 12(1).

Akbudak MA, Filiz E, Kontbay K. 2018. *Dreb2* (dehydration-responsive element-binding protein 2) type transcription factor in sorghum (*sorghum bicolor*): Genome-wide identification, characterization and expression profiles under cadmium and salt stresses. 3 Biotech. 8(10).

Akhtar M, Jaiswal A, Taj G, Jaiswal JP, Qureshi MI, Singh NK. 2012. Dreb1/cbf transcription factors: Their structure, function and role in abiotic stress tolerance in plants. Journal of Genetics. 91(3):385-395.

Alexander DH, Lange K. 2011. Enhancements to the admixture algorithm for individual ancestry estimation. Bmc Bioinformatics. 12.

Alexandratos N, Bruinsma J. 2012. World agriculture towards 2030/2050: The 2012 revision.

Alza JO, Gutierrez N, Cabezas JM, Torres AM, Lorite IJ. 2024. Interactive effects of atmospheric co2 concentration and water stress on the phenology and physiology of faba bean (*vicia faba* l.). Environmental and Experimental Botany. 218.

Ambachew D, Londono JM, Castillo NR, Asfaw A, Blair MW. 2024. Genetic diversity, linkage disequilibrium, and population structure in a common bean reference collection. Agronomy-Basel. 14(5).

Andrews S. 2010. Fastqc: A quality control tool for high throughput sequence data [online]. 0.11.5 ed.

Angioi SA, Rau D, Attene G, Nanni L, Bellucci E, Logozzo G, Negri V, Zeuli PLS, Papa R. 2010. Beans in europe: Origin and structure of the european landraces of phaseolus vulgaris l. Theoretical and Applied Genetics. 121(5):829-843.

Aoyanagi T, Ikeya S, Kobayashi A, Kozaki A. 2020. Gene regulation via the combination of transcription factors in the indeterminate domain and gras families. Genes. 11(6).

Araujo SS, Beebe S, Crespi M, Delbreil B, Gonzalez EM, Gruber V, Lejeune-Henaut I, Link W, Monteros MJ, Prats E et al. 2015. Abiotic stress responses in legumes: Strategies used to cope with environmental challenges. Critical Reviews in Plant Sciences. 34(1-3):237-280.

Arnold JB. 2024. Ggthemes: Extra themes, scales and geoms for 'ggplot2'. R package version 5.1.0 ed.

Arregocés HA, Rojano R, Castellanos ML. 2025. Trends and variability of drought in northern south america between 1982 and 2022. Earth Systems and Environment.

Arriagada O, Arévalo B, Cabeza RA, Carrasco B, Schwember AR. 2023. Meta-qtl analysis for yield components in common bean (phaseolus vulgaris l.). Plants-Basel. 12(1).

Arriagada O, Schwember AR, Greve MJ, Urban MO, Cabeza RA, Carrasco B. 2021. Morphological and molecular characterization of selected chilean runner bean (phaseolus coccineus l.) genotypes shows moderate agronomic and genetic variability. Plants-Basel. 10(8).

Aslam MM, Waseem M, Jakada BH, Okal EJ, Lei ZL, Saqib HSA, Yuan W, Xu WF, Zhang Q. 2022. Mechanisms of abscisic acid-mediated drought stress responses in plants. International Journal of Molecular Sciences. 23(3).

Assefa T, Mahama AA, Brown AV, Cannon EKS, Rubyogo JC, Rao IM, Blair MW, Cannon SB. 2019. A review of breeding objectives, genomic resources, and marker-assisted methods in common bean (phaseolus vulgaris l.). Molecular Breeding. 39(2).

Azani N, Babineau M, Bailey CD, Banks H, Barbosa AR, Pinto RB, Boatwright JS, Borges LM, Brown GK, Bruneau A et al. 2017. A new subfamily classification of the leguminosae based on a taxonomically comprehensive phylogeny. Taxon. 66(1):44-77.

Azizi A, Verchot J, Moieni A, Shams-bakhsh M. 2020. Efficient silencing gene construct for resistance to multiple common bean (phaseolus vulgaris l.) viruses. 3 Biotech. 10(6).

Bandurska H. 2022. Drought stress responses: Coping strategy and resistance. Plants-Basel. 11(7).

Banoo A, Nabi A, Rasool RS, Mahiya-Farooq, Shah MHD, Ahmad M, Sofi PA, Aasiya-Nabi, Itoo H, Sharma PN et al. 2020. North-western himalayan common beans: Population structure and mapping of quantitative anthracnose resistance through genome wide association study. Frontiers in Plant Science. 11.

Barezzi M, Sanginario A, Canone D, Gisolo D, Gentile A, Nari L, Pettiti F, Pettiti F. 2024. Wappfruit—an automatic system for drip irrigation in orchards based on real-time soil matric potential data. IEEE Transactions on AgriFood Electronics. 2:293-305.

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. 2011. Bamtools: A c++ api and toolkit for analyzing and managing bam files. Bioinformatics. 27(12):1691-1692.

Barrett LW, Fletcher S, Wilton SD. 2012. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. Cellular and Molecular Life Sciences. 69(21):3613-3634.

Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Cukura A, Denny P et al. 2023. Uniprot: The universal protein knowledgebase in 2023. Nucleic Acids Research. 51(D1):D523-D531.

Becerra VV, Paredes CM, Rojo MC, Diaz LM, Blair MW. 2010. Microsatellite marker characterization of chilean common bean (phaseolus vulgaris l.) germplasm. Crop Science. 50(5):1932-1941.

Beebe S, Rao I, Mukankusi C, Buruchara R. 2012. Improving resource use efficiency and reducing risk of common bean production in africa, latin america, and the caribbean. Centro Internacional de Agricultura Tropical (CIAT).

Beebe S, Rengifo J, Gaitan E, Duque MC, Tohme J. 2001. Diversity and origin of andean landraces of common bean. Crop Science. 41(3):854-862.

Beebe S, Skroch PW, Tohme J, Duque MC, Pedraza F, Nienhuis J. 2000. Structure of genetic diversity among common bean landraces of middle american origin based on correspondence analysis of rapd. Crop Science. 40(1):264-273.

Beebe SE, Rao IM, Blair MW, Acosta-Gallegos JA. 2013. Phenotyping common beans for adaptation to drought. Frontiers in Physiology. 4.

Beebe SE, Rao IM, Cajiao C, Grajales M. 2008. Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. Crop Science. 48(2):582-592.

Bellucci E, Benazzo A, Xu CM, Bitocchi E, Rodriguez M, Alseekh S, Di Vittori V, Gioia T, Neumann K, Cortinovis G et al. 2023. Selection and adaptive introgression guided the complex evolutionary history of the european common bean. Nature Communications. 14(1).

Bernardi C, Macri G, Biagi M, Miraldi E, Finetti F, Trabalzini L. 2023. In vitro digestion of *phaseolus vulgaris* l. Cooked beans induces autophagy in colon cancer cells. Foods. 12(4).

Bhakta MS, Gezan SA, Michelangeli JAC, Carvalho M, Zhang L, Jones JW, Boote KJ, Correll MJ, Beaver J, Osorno JM et al. 2017. A predictive model for time-to-flowering in the common bean based on qtl and environmental variables. G3-Genes Genomes Genetics. 7(12):3901-3912.

Bista R, Thakuri S, Koirala P. 2024. Assessing water deficit induced water stress of households and their welfare in small cities of nepal. Economic Journal of Environmental Studies.

Bitocchi E, Bellucci E, Giardini A, Rau D, Rodriguez M, Biagetti E, Santilocchi R, Zeuli PS, Gioia T, Logozzo G et al. 2013. Molecular analysis of the parallel domestication of the common bean (phaseolus vulgaris) in mesoamerica and the andes. New Phytologist. 197(1):300-313.

Bitocchi E, Nanni L, Bellucci E, Rossi M, Giardini A, Zeuli PS, Logozzo G, Stougaard J, McClean P, Attene G et al. 2012. Mesoamerican origin of the common bean (phaseolus vulgaris l.) is revealed by sequence data. Proceedings of the National Academy of Sciences of the United States of America. 109(14):E788-E796.

Bitocchi E, Rau D, Bellucci E, Rodriguez M, Murgia ML, Gioia T, Santo D, Nanni L, Attene G, Papa R. 2017. Beans (phaseolus ssp.) as a model for understanding crop evolution. Frontiers in Plant Science. 8.

Blair MW, Chaves A, Tofino A, Calderon JF, Palacio JD. 2010a. Extensive diversity and inter-genepool introgression in a world-wide collection of indeterminate snap bean accessions. Theoretical and Applied Genetics. 120(7):1381-1391.

Blair MW, Cortes AJ, Penmetsa RV, Farmer A, Carrasquilla-Garcia N, Cook DR. 2013a. A high-throughput snp marker system for parental polymorphism screening, and diversity analysis in common bean (*phaseolus vulgaris* l.). Theoretical and Applied Genetics. 126(2):535-548.

Blair MW, Diaz JM, Hidalgo R, Diaz LM, Duque MC. 2007. Microsatellite characterization of andean races of common bean (phaseolus vulgaris l.). Theoretical and Applied Genetics. 116(1):29-43.

Blair MW, Diaz LM, Buendia HF, Duque MC. 2009. Genetic diversity, seed size associations and population structure of a core collection of common beans (phaseolus vulgaris l.). Theoretical and Applied Genetics. 119(6):955-972.

Blair MW, Gonzalez LF, Kimani PM, Butare L. 2010b. Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (phaseolus vulgaris l.) from central africa. Theoretical and Applied Genetics. 121(2):237-248.

Blair MW, Iriarte G, Beebe S. 2006. Qtl analysis of yield traits in an advanced backcross population derived from a cultivated andean × wild common bean (phaseolus vulgaris l.) cross. Theoretical and Applied Genetics. 112(6):1149-1163.

Blair MW, Izquierdo P, Astudillo C, Grusak MA. 2013b. A legume biofortification quandary: Variability and genetic control of seed coat micronutrient accumulation in common beans. Frontiers in Plant Science. 4.

Blair MW, Soler A, Cortés AJ. 2012. Diversification and population structure in common beans (*phaseolus vulgaris* l.). Plos One. 7(11).

Blum A. 2015. Towards a conceptual aba ideotype in plant breeding for water limited environments. Functional Plant Biology. 42(6):502-513.

Bohra A, Choudhary M, Bennett D, Joshi R, Mir RR, Varshney RK. 2024. Drought-tolerant wheat for enhancing global food security. Functional & Integrative Genomics. 24(6).

Borkiewicz L, Polkowska-Kowalczyk L, Ciesla J, Sowinski P, Jonczyk M, Rymaszewski W, Szymanska KP, Jazwiec R, Muszynska G, Szczegielniak J. 2020. Expression of maize calcium-dependent protein kinase (zmcpk11) improves salt tolerance in transgenic arabidopsis plants by regulating sodium and potassium homeostasis and stabilizing photosystem ii. Physiologia Plantarum. 168(1):38-57.

Borrell AK, Wong ACS, George-Jaeggli B, van Oosterom EJ, Mace ES, Godwin ID, Liu GQ, Mullet JE, Klein PE, Hammer GL et al. 2022. Genetic modification of pin genes induces causal mechanisms of stay-green drought adaptation phenotype. Journal of Experimental Botany. 73(19):6711-6726.

Botero H, Barnes AP. 2022. The effect of enso on common bean production in colombia: A time series approach. Food Security. 14(6):1417-1430.

Botero H, Barnes AP, Perez L, Rios D, Ramirez-Villegas J. 2021. The determinants of common bean variety selection and diversification in colombia. Ecological Economics. 190.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. Tassel: Software for association mapping of complex traits in diverse samples. Bioinformatics. 23(19):2633-2635.

Brookfield H, Stocking M. 1999. Agrodiversity: Definition, description and design. Global Environmental Change-Human and Policy Dimensions. 9(2):77-80.

Broughton WJ, Hernández G, Blair M, Beebe S, Gepts P, Vanderleyden J. 2003. Beans (phaseolus spp.) – model food legumes. Plant and Soil. 252(1):55-128.

Buitrago-Bitar MA, Cortes AJ, Lopez-Hernandez F, Londono-Caicedo JM, Munoz-Florez JE, Munoz LC, Blair MW. 2021. Allelic diversity at abiotic stress responsive genes in relationship to ecological drought indices for cultivated tepary bean, phaseolus acutifolius a. Gray, and its wild relatives. Genes. 12(4).

Burle ML, Fonseca JR, Kami JA, Gepts P. 2010. Microsatellite diversity and genetic structure among common bean (phaseolus vulgaris l.) landraces in brazil, a secondary center of diversity. Theoretical and Applied Genetics. 121(5):801-813.

Cai X, Liu CJ, Tang LY, Zhang SJ, Li XH, Wang HT, Zhang JH. 2025. A cysteine-rich transmembrane module peptide *ghcystm9* is involved in cold stress response. Bmc Plant Biology. 25(1).

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. Blast plus : Architecture and applications. BMC Bioinformatics. 10.

Campa A, Murube E, Ferreira JJ. 2018. Genetic diversity, population structure, and linkage disequilibrium in a spanish common bean diversity panel revealed through genotyping-by-sequencing. Genes. 9(11).

Campos K, Schwember AR, Machado D, Ozores-Hampton M, Gil PM. 2021. Physiological and yield responses of green-shelled beans (*phaseolus vulgaris* l.) grown under restricted irrigation. Agronomy-Basel. 11(3).

Cantalapiedra C, Hernandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. Eggnog-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Molecular Biology and Evolution.

Caproni L, Raggi L, Talsma EF, Wenzl P, Negri V. 2020. European landrace diversity for common bean biofortification: A genome-wide association study. Scientific Reports. 10(1).

Casanas F, Simo J, Casals J, Prohens J. 2017. Toward an evolved concept of landrace. Frontiers in Plant Science. 8.

Castro-Guerrero NA, Isidra-Arellano M, Mendoza-Cozatl DG, Valdes-Lopez O. 2016. Common bean: A legume model on the rise for unraveling responses and adaptations to iron, zinc, and phosphate deficiencies. Frontiers in Plant Science. 7.

Catarcione G, Paolacci AR, Alicandri E, Gramiccia E, Taviani P, Rea R, Costanza MT, De Lorenzis G, Puccio G, Mercati F et al. 2023. Genetic diversity and population structure of common bean (phaseolus vulgaris l.) landraces in the lazio region of italy. Plants (Basel). 12(4).

Cavalcante AG, Lemos LB, Meirelles FC, Cavalcante ACP, de Aquino LA. 2020. Thermal sum and phenological descriptions of growth stages of the common bean according to the bbch scale. Annals of Applied Biology. 176(3):342-349.

Cebeci Z, Bayraktar M, Gökçe G. 2023. Comparison of the statistical methods for genome-wide association studies on simulated quantitative traits of domesticated goats (*capra hircus* l.). Small Ruminant Research. 227.

Cerutti PH, Melo RC, Carbonari LTS, Guidolin AF, Coimbra JLM. 2023. Improvement of common bean root system phenotyping for identification of superior genotypes. Genetics and Molecular Research. 22(2).

Chacon MI, Pickersgill B, Debouck DG. 2005. Domestication patterns in common bean (phaseolus vulgaris l.) and the origin of the mesoamerican and andean cultivated races. Theoretical and Applied Genetics. 110(3):432-444.

Chacon S MI, Pickersgill B, Debouck DG, Arias JS. 2007. Phylogeographic analysis of the chloroplast dna variation in wild common bean (phaseolus vulgaris l.) in the americas. Plant Systematics and Evolution. 266(3-4):175-195.

Chacon-Sanchez MI, Martinez-Castillo J, Duitama J, Debouck DG. 2021. Gene flow in *phaseolus* beans and its role as a plausible driver of ecological fitness and expansion of cultigens. Frontiers in Ecology and Evolution. 9.

Chen F, Yang YZ, Luo XF, Zhou WG, Dai YJ, Zheng C, Liu WG, Yang WY, Shu K. 2019. Genome-wide identification of grf transcription factors in soybean and expression analysis of gmgrf family under shade stress. Bmc Plant Biology. 19.

Chen YH, Lu J, Yang X, Huang LC, Zhang CQ, Liu QQ, Li QF. 2023. Gene editing of non-coding regulatory dna and its application in crop improvement. Journal of Experimental Botany. 74(19):6158-6175.

Chu XT, Su HA, Hayashi S, Gresshoff PM, Ferguson BJ. 2022. Spatiotemporal changes in gibberellin content are required for soybean nodulation. New Phytologist. 234(2):479-493.

Cichy KA, Porch TG, Beaver JS, Cregan P, Fourie D, Glahn RP, Grusak MA, Kamfwa K, Katuuramu DN, McClean P et al. 2015. A diversity panel for andean bean improvement. Crop Science. 55(5):2149-2160.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu XY, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff. Fly. 6(2):80-92.

Clark SE, Running MP, Meyerowitz EM. 1995. Clavata3 is a specific regulator of shoot and floral meristem development affecting the same processes as clavata1. Development. 121(7):2057-2067.

Coelho AP, de Faria RT, Lemos LB, dos Reis MAM, Filla VA, Bertino AMP. 2023. Irrigation management of common bean cultivars with contrasting growth habits. Scientia Agricola. 80.

Cominelli S, Rivera LD, Brown WG, Ochsner TE, Patrignani A. 2024. Calibration of teros 10 and teros 12 electromagnetic soil moisture sensors. Soil Science Society of America Journal. 88(6):2104-2122.

Cortes AJ, Chavarro MC, Blair MW. 2011. Snp marker diversity in common bean (phaseolus vulgaris l.). Theoretical and Applied Genetics. 123(5):827-845.

Cortinovis G, Frascarelli G, Di Vittori V, Papa R. 2020. Current state and perspectives in population genomics of the common bean. Plants-Basel. 9(3).

Cortés AJ, Monserrate FA, Ramírez-Villegas J, Madriñán S, Blair MW. 2013. Drought tolerance in wild plant populations: The case of common beans (*phaseolus vulgaris* l.). Plos One. 8(5).

Cusworth G, Garnett T, Lorimer J. 2021. Agroecological break out: Legumes, crop diversification and the regenerative futures of uk agriculture. Journal of Rural Studies. 88:126-137.

da Silva HAP, Caetano VS, Pessôa DDV, Pacheco RS, Meneses CHSG, Simoes-Araújo JL. 2024. Unraveling the drought-responsive transcriptomes in nodules of two common bean genotypes during biological nitrogen fixation. Frontiers in Plant Science. 15.

da Silva LC, de Souza TLPO, Cruz CD, Carneiro PCS, Silva FFE, de Barros EG, Vianello RP, da Fonseca CEL, Song QJ, Cregan PB et al. 2018. Linkage fine-mapping and qtls affecting morpho-agronomic traits of a mesoamericanxandean ril common bean population. Euphytica. 214(12).

Daba K, Warkentin TD, Bueckert R, Todd CD, Tar'an B. 2016. Determination of photoperiod-sensitive phase in chickpea (cicer arietinum l.). Frontiers in Plant Science. 7.

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM et al. 2021. Twelve years of samtools and bcftools. Gigascience. 10(2).

Darkwa K, Ambachew D, Mohammed H, Asfaw A, Blair MW. 2016. Evaluation of common bean (*phaseolus vulgaris* l.) genotypes for drought stress adaptation in ethiopia. Crop Journal. 4(5):367-376.

de Almeida CP, Paulino JFD, Carbonell SAM, Chiorato AF, Song QJ, Di Vittori V, Rodriguez M, Papa R, Benchimol-Reis LL. 2020. Genetic diversity, population structure, and andean introgression in brazilian common bean cultivars after half a century of genetic breeding. Genes. 11(11).

De Ron AM, Papa R, Bitocchi E, González AM, Debouck DG, Brick MA, Fourie D, Marsolais F, Beaver J, Geffroy V et al. 2015. Common beans. In: De Ron AM, editor. Grain legumes, handbook of plant breeding. New York: Springer.

de Souza IP, de Azevedo BR, Coelho ASG, de Souza TLPO, Valdisser PAMR, Gomes-Messias LM, Funicheli BO, Brondani C, Vianello RP. 2023a. Whole-genome resequencing of common bean elite breeding lines. Scientific Reports. 13(1).

de Souza LP, Bitocchi E, Papa R, Tohge T, Fernie AR. 2023b. Decreased metabolic diversity in common beans associated with domestication revealed by untargeted metabolomics, information theory, and molecular networking. Plant Journal. 115(4):1021-1036.

Debouck DG, Toro O, Paredes OM, Johnson WC, Gepts P. 1993. Genetic diversity and ecological distribution of phaseolus-vulgaris (fabaceae) in northwestern south-america. Economic Botany. 47(4):408-423.

Delfin EF, Drobnitch ST, Comas LH. 2021. Plant strategies for maximizing growth during water stress and subsequent recovery in *solanum melongena* l. (eggplant). Plos One. 16(9).

Delfini J, Moda-Cirino V, Neto JD, Zeffa DM, Nogueira AF, Ribeiro LAB, Ruas PM, Gepts P, Gonçalves LSA. 2021. Genome-wide association study identifies genomic regions for important morpho-agronomic traits in mesoamerican common bean. Frontiers in Plant Science. 12.

Denning-James K, Chater CCC, Cortes AJ, Blair M, Pelaez D, Hall A, De Vega JJ. 2025. Genome-wide association mapping dissects the selective breeding of determinacy and photoperiod sensitivity in common bean (phaseolus vulgaris l.). G3 Genes|Genomes|Genetics.

Derek H O, Jason C D, A PW, Alexis D. 2025. Fsa: Simple fisheries stock assessment methods. R package version 0.10.0 ed.

Di Vittori V, Bellucci E, Bitocchi E, Rau D, Rodriguez M, Leonarda Murgia M, Nanni L, Attene G, Papa R. 2017. Domestication and crop history. In: de la Vega MP, Santalla M, Marsolais F, editors. The common bean genome. Cham: Springer. p. 21-55.

Diesh C, Stevens GJ, Xie PT, Martinez TD, Hershberg EA, Leung A, Guo E, Dider S, Zhang JJ, Bridge C et al. 2023. Jbrowse 2: A modular genome browser with views of synteny and structural variation. Genome Biology. 24(1).

Diniz AL, Giordani W, Costa ZP, Margarido GRA, Perseguini JMKC, Benchimol-Reis LL, Chiorato AF, Garcia AAF, Vieira MLC. 2019. Evidence for strong kinship influence

on the extent of linkage disequilibrium in cultivated common beans. Genes. 10(1).

Phaseolus vulgaris labor ovalle v1.1. 2021. [accessed 2021]. https://phytozome-next.jgi.doe.gov/info/PvulgarisLaborOvalle_v1_1.

Dramadri IO, Nkalubo ST, Kelly JD. 2019. Identification of qtl associated with drought tolerance in andean common bean. Crop Science. 59(3):1007-1020.

Driever SM, Mossink L, Ocaña DN, Kaiser E. 2023. A simple system for phenotyping of plant transpiration and stomatal conductance response to drought. Plant Science. 329.

Du H, Chang Y, Huang F, Xiong LZ. 2015. Gid1 modulates stomatal response and submergence tolerance involving abscisic acid and gibberellic acid signaling in rice. Journal of Integrative Plant Biology. 57(11):954-968.

Dunn OJ. 1964. Multiple comparisons using rank sums. Technometrics. 6:241-252.

Duran LA, Blair MW, Giraldo MC, Macchiavelli R, Prophete E, Nin JC, Beaver JS. 2005. Morphological and molecular characterization of common bean landraces and cultivars from the caribbean. Crop Science. 45(4):1320-1328.

Egesa AO, Vallejos CE, Begcy K. 2024. Cell size differences affect photosynthetic capacity in a mesoamerican and an andean genotype of *phaseolus vulgaris* l. Frontiers in Plant Science. 15.

El Bey N, Maazoun A, Nahdi O, Krima N. 2024. Water stress indicators in citrus, olive and apple trees: A review. Journal of Applied Horticulture.

Elias JCF, Gonçalves-Vidigal MC, Bisneta MV, Valentini G, Vidigal PS, Gilio TAS, Moda-Cirino V, Song QJ. 2021. Genetic mapping for agronomic traits in iapar 81/lp97-28 population of common bean (phaseolus vulgaris l.) under drought conditions. Plants (Basel). 10(8).

Ewels P, Magnusson M, Lundin S, Kaller M. 2016. Multiqc: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 32(19):3047-3048.

FAO. 2009. International treaty on plant genetic resources for food and agriculture. Food and Agriculture Organization.

FAO. 2024. The state of food and agriculture 2024 – value-driven transformation of agrifood systems. Rome.

Http://faostat.Fao.Org/. 2023. [accessed 05 05 2025].

Farooq M, Gogoi N, Barthakur S, Baroowa B, Bharadwaj N, Alghamdi SS, Siddique KHM. 2017. Drought stress in grain legumes during reproduction and grain filling. Journal of Agronomy and Crop Science. 203(2):81-102.

Fenta BA, Beebe SE, Kunert KJ. 2020. Role of fixing nitrogen in common bean growth under water deficit conditions. Food and Energy Security. 9(1).

Fioravanti G, Toreti A, Cammalleri C, Muñoz CA, Bavera D, De Jager A, Essenfelder AH, Di Ciollo C, Masante D, Magni D et al. 2025. A dataset for monitoring agricultural drought in europe. Scientific Data. 12(1).

Foyer CH, Lam HM, Nguyen HT, Siddique KHM, Varshney RK, Colmer TD, Cowling W, Bramley H, Mori TA, Hodgson JM et al. 2016. Neglecting legumes has compromised human health and sustainable food production. Nature Plants. 2(8).

Frascarelli G, Galise TR, D'Agostino N, Cafasso D, Cozzolino S, Cortinovis G, Sparvoli F, Bellucci E, Di Vittori V, Nanni L et al. 2025. The evolutionary history of the common bean (*phaseolus vulgaris*) revealed by chloroplast and nuclear genomes analysis. Theoretical and Applied Genetics. 138(2).

Freytag GF, Debouck DG. 2002. Taxonomy, distribution, and ecology of the genus phaseolus (leguminosae-papilionoideae) in north america, mexico and central america. Botanical Research Institute of Texas (BRIT), Forth Worth, TX, USA. 298 p. (Sida, botanical miscellany no. 23).

Gan ES, Xu YF, Wong JY, Goh JG, Sun B, Wee WY, Huang JB, Ito T. 2014. Jumonji demethylases moderate precocious flowering at elevated temperature via regulation of flc in arabidopsis. Nature Communications. 5.

Ganesan K, Xu BJ. 2017. Polyphenol-rich dry common beans (*phaseolus vulgaris* l.) and their health benefits. International Journal of Molecular Sciences. 18(11).

Ganesan SP, Bordoloi S, Cai WL, Garg A, Sekharan S, Sahoo L. 2024. Effect of soil type on tipping point hydrological requirements for *axonopus compressus* grass under extreme drought stress. Science of the Total Environment. 953.

García-Fernández C, Campa A, Garzón AS, Miklas P, Ferreira JJ. 2021. Gwas of pod morphological and color characters in common bean. Bmc Plant Biology. 21(1).

Garratt MPD, Coston DJ, Truslove CL, Lappage MG, Polce C, Dean R, Biesmeijer JC, Potts SG. 2014. The identity of crop pollinators helps target conservation for improved ecosystem services. Biological Conservation. 169:128-135.

Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv preprint. (arXiv:1207.3907).

Gasc C, Peyretaillade E, Peyret P. 2016. Sequence capture by hybridization to explore modern and ancient genomic diversity in model and nonmodel organisms. Nucleic Acids Research. 44(10):4504-4518.

Gautam T, Dutta M, Jaiswal V, Zinta G, Gahlaut V, Kumar S. 2022. Emerging roles of sweet sugar transporters in plant development and abiotic stress responses. Cells. 11(8).

Gebregiorgis MF, Savage MJ. 2006. Field, laboratory and estimated soil-water content limits. Water Sa. 32(2):155-161.

Gepts P, Bliss FA. 1985. F1-hybrid weakness in the common bean - differential geographic origin suggests 2 gene pools in cultivated bean germplasm. Journal of Heredity. 76(6):447-450.

Gepts P, Debouck D. 1991. Origin, domestication, and evolution of the common bean (*phaseolus vulgaris* l.). In: van Schoonhoven A, Voysest O, editors. Common beans: Research for crop improvement. C.A.B. Int.; Wallingford, UK: CIAT; Cali, Colombia. p. 7-53.

Ginestet C. 2011. Ggplot2: Elegant graphics for data analysis. Journal of the Royal Statistical Society Series a-Statistics in Society. 174:245-245.

Giordani W, Gama HC, Chiorato AF, Garcia AAF, Vieira MLC. 2022. Genome-wide association studies dissect the genetic architecture of seed shape and size in common bean. G3-Genes Genomes Genetics. 12(4).

Girón-Ramírez A, Peña-Rodríguez LM, Escalante-Erosa F, Fuentes G, Santamaría JM. 2021. Identification of the shine clade of ap2/erf domain transcription factors genes in *carica papaya*; their gene expression and their possible role in wax accumulation and water deficit stress tolerance in a wild and a commercial papaya genotypes. Environmental and Experimental Botany. 183.

González AM, Yuste-Lisbona FJ, Saburido S, Bretones S, De Ron AM, Lozano R, Santalla M. 2016. Major contribution of flowering time and vegetative growth to plant

production in common bean as deduced from a comparative genetic mapping. Frontiers in Plant Science. 7.

González-Delgado A, Martínez-Rivas FJ, Jiménez-Gómez JM. 2025. Photoperiod insensitivity in crops. Journal of Experimental Botany.

Goodstein DM, Shu SQ, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N et al. 2012. Phytozome: A comparative platform for green plant genomics. Nucleic Acids Research. 40(D1):D1178-D1186.

Gori B, Ulian T, Bernal HY, Diazgranados M. 2022. Understanding the diversity and biogeography of colombian edible plants. Scientific Reports. 12(1).

Guidi L, Lo Piccolo E, Landi M. 2019. Chlorophyll fluorescence, photoinhibition and abiotic stress: Does it make any difference the fact to be a c3 or c4 species? Frontiers in Plant Science. 10.

Gutierrez N, Torres AM. 2025. Combination of qtl mapping and gwas for fine mapping and gene mining of drought tolerance and seed yield components in faba bean (vicia faba l.). Bmc Plant Biology. 25(1).

Guttikonda SK, Valliyodan B, Neelakandan AK, Tran LSP, Kumar R, Quach TN, Voothuluru P, Gutierrez-Gonzalez JJ, Aldrich DL, Pallardy SG et al. 2014. Overexpression of atdreb1d transcription factor improves drought tolerance in soybean. Molecular Biology Reports. 41(12):7995-8008.

Habus Jercic I, Baric M, Keresa S, Bosnjak Mihovilovic A, Poljak M, Lazarevic B. 2018. Effect of terminal drought on yield and some physiological traits of winter wheat. Genetika-Belgrade. 50(2):747-753.

Haga K, Sakai T. 2012. Pin auxin efflux carriers are necessary for pulse-induced but not continuous light-induced phototropism in arabidopsis. Plant Physiology. 160(2):763-776.

Hageman AN, Urban MO, Van Volkenburgh E. 2020. Sensitivity of leaflet growth rate to drought predicts yield in common bean (phaseolus vulgaris). Functional Plant Biology. 47(9):792-802.

Hamabwe SM, Otieno NA, Odhiambo JA, Parker T, Kamfwa K. 2024. Agronomic and morpho-physiological response of andean genotypes of common bean to terminal drought. Crop Science. 64(6):3521-3532.

Han DL, Yu ZB, Lai JB, Yang CW. 2022a. Post-translational modification: A strategic response to high temperature in plants. Abiotech. 3(1):49-64.

Han JP, Xie XX, Zhang Y, Yu XF, He GY, Li Y, Yang GX. 2022b. Evolution of the dehydration-responsive element-binding protein subfamily in green plants. Plant Physiology. 190(1):421-440.

Harlan J, Wet Jd. 1971. Toward a rational classification of cultivated plants. Taxon. 20(4):509-517.

He SL, Li B, Zahurancik WJ, Arthur HC, Sidharthan V, Gopalan V, Wang L, Jang JC. 2024. Overexpression of stress granule protein tzf1 enhances salt stress tolerance by targeting *aca11* mrna for degradation in arabidopsis. Frontiers in Plant Science. 15.

Beans and herbs. 2020. [accessed 2020]. https://www.beansandherbs.co.uk/.

Hernández-López JA, Puerta-Cortés DX, Andrade HJ. 2024. Predictive analysis of adaptation to drought of farmers in the central zone of colombia. Sustainability. 16(16).

Hill WG, Robertson A. 2007. The effect of linkage on limits to artificial selection (reprinted). Genetics Research. 89(5-6):311-336.

Hirakawa Y. 2021. Clavata3, a plant peptide controlling stem cell fate in the meristem. Peptides. 142.

Hothorn T, Bretz F, Westfall P. 2008. Simultaneous inference in general parametric models. Biometrical Journal. 50:346-363.

Hoyos-Villegas V, Song QJ, Kelly JD. 2017. Genome-wide association analysis for drought tolerance and associated traits in common bean. Plant Genome. 10(1).

Hu TS, Chitnis N, Monos D, Dinh A. 2021a. Next-generation sequencing technologies: An overview. Human Immunology. 82(11):801-811.

Hu W, Ji CM, Liang Z, Ye JQ, Ou WJ, Ding ZH, Zhou G, Tie WW, Yan Y, Yang JH et al. 2021b. Resequencing of 388 cassava accessions identifies valuable loci and selection for variation in heterozygosity. Genome Biology. 22(1).

Huala E, Dickerman AW, Garcia-Hernandez M, Weems D, Reiser L, LaFond F, Hanley D, Kiphart D, Zhuang MZ, Huang W et al. 2001. The arabidopsis information resource (tair): A comprehensive database and web-based information retrieval,

analysis, and visualization system for a model plant. Nucleic Acids Research. 29(1):102-105.

Huang G, Zeng Y. 2024. Increased stomatal conductance and leaf biochemical capacity, not mesophyll conductance, contributing to the enhanced photosynthesis in *oryza* plants during domestication. Planta. 259.

Huang M, Liu XL, Zhou Y, Summers RM, Zhang ZW. 2019. Blink: A package for the next level of genome-wide association studies with both individuals and markers in the millions. Gigascience. 8(2).

Humann J, Jung S, Cheng C-H, Lee T, Zheng P, Frank M, McGaughey D, Scott K, Buble K, Yu J et al. 2019. Cool season food legume genome database: A resource for pea, lentil, faba bean and chickpea genetics, genomics and breeding. Paper presented at: International Plant and Animal Genome Conference. San Diego, CA, USA.

Hussaini SMB, Sidle RC, Kazimi Z, Khan AA, Rezaei AQ, Ghulami Z, Buda T, Rastagar R, Fatimi AA, Muhmmadi Z. 2021. Drought tolerant varieties of common beans (*phaseolus vulgaris*) in central afghanistan. Agronomy-Basel. 11(11).

Hvitfeldt E. 2021. Paletteer: Comprehensive collection of color palettes. R package version 1.3.0 ed.

IPCC. 2021. Summary for policymakers. In: Climate change 2021: The physical science basis. Contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change. Masson-Delmotte V, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L., Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou editor. Cambridge University Press.

IPCC. 2023. Climate change 2023: Synthesis report. Contribution of working groups i, ii and iii to the sixth assessment report of the intergovernmental panel on climate change. IPCC, Geneva, Switzerland.

Irshad A, Ahmad H, Muhammad I, Khan S. 2024. The role of water stress and soil texture on plant roots anatomy, architecture, and senescence. Frontiers in Plant Science.

Jacques C, Salon C, Barnard RL, Vernoud V, Prudent M. 2021. Drought stress memory at the plant cycle level: A review. Plants-Basel. 10(9).

Jahan E, Sharwood RE, Tissue DT. 2023. Effects of leaf age during drought and recovery on photosynthesis, mesophyll conductance and leaf anatomy in wheat leaves. Frontiers in Plant Science. 14.

Jaramillo-Correa JP, Verdú M, González-Martínez SC. 2010. The contribution of recombination to heterozygosity differs among plant evolutionary lineages and life-forms. Bmc Evolutionary Biology. 10.

Jat M, Ray M, Ahmad MA, Prakash P. 2024. Unravelling the photosynthetic dynamics and fluorescence parameters under ameliorative effects of 24-epibrassinolide in wheat (*triticum aestivum* l.) grown under heat stress regime. Scientific Reports. 14(1).

Jha AB, Ashokkumar K, Diapari M, Ambrose SJ, Zhang HX, Tar'an B, Bett KE, Vandenberg A, Warkentin TD, Purves RW. 2015. Genetic diversity of folate profiles in seeds of common bean, lentil, chickpea and pea. Journal of Food Composition and Analysis. 42:134-140.

2020. [accessed 2020]. https://jungleseeds.co.uk/.

Kachinski WD, Avila FW, Dos Reis AR, Muller MML, Mendes MC, Petranski PH. 2022. Agronomic biofortification increases concentrations of zinc and storage proteins in common bean (*phaseolus vulgaris* l.) grains. Food Res Int. 155:111105.

Kalyta K, Stelmaszczyk W, Szczesniak D, Kotula L, Dobosz P, Mroczek M. 2023. The spectrum of the heterozygous effect in biallelic mendelian diseases-the symptomatic heterozygote issue. Genes. 14(8).

Kamal NM, Gorafi YSA, Abdelrahman M, Abdellatef E, Tsujimoto H. 2019. Stay-green trait: A prospective approach for yield potential, and drought and heat stress adaptation in globally important cereals. International Journal of Molecular Sciences. 20(23).

Kamfwa K, Cichy KA, Kelly JD. 2015. Genome-wide association study of agronomic traits in common bean. Plant Genome. 8(2).

Kamfwa K, Cichy KA, Kelly JD. 2019. Identification of quantitative trait loci for symbiotic nitrogen fixation in common bean. Theoretical and Applied Genetics. 132(5):1375-1387.

Kami J, Velasquez VB, Debouck DG, Gepts P. 1995. Identification of presumed ancestral dna-sequences of phaseolin in phaseolus-vulgaris. Proceedings of the National Academy of Sciences of the United States of America. 92(4):1101-1104.

Kamrani YY, Shomali A, Aliniaeifard S, Lastochkina O, Moosavi-Nezhad M, Hajinajaf N, Talar U. 2022. Regulatory role of circadian clocks on aba production and signaling, stomatal responses, and water-use efficiency under water-deficit conditions. Cells. 11(7).

Kassambara A. 2023. Rstatix: Pipe-friendly framework for basic statistical tests. R package version 0.7.2.

Kavas M, Abdulla MF, Mostafa K, Seçgin Z, Yerlikaya BA, Otur C, Gökdemir G, Kizildogan AK, Al-Khayri JM, Jain SM. 2022. Investigation and expression analysis of r2r3-mybs and anthocyanin biosynthesis-related genes during seed color development of common bean (*phaseolus vulgaris*). Plants-Basel. 11(23).

Keller B, Ariza-Suarez D, Portilla-Benavides AE, Buendia HF, Aparicio JS, Amongi W, Mbiu J, Msolla SN, Miklas P, Porch TG et al. 2022. Improving association studies and genomic predictions for climbing beans with data from bush bean populations. Frontiers in Plant Science. 13.

Khoudi H. 2023. Shine clade of erf transcription factors: A significant player in abiotic and biotic stress tolerance in plants. Plant Physiology and Biochemistry. 195:77-88.

Kim J, Kang H, Park J, Kim W, Yoo J, Lee N, Kim J, Yoon TY, Choi G. 2016. Pif1-interacting transcription factors and their binding sequence elements determine the in vivo targeting sites of pif1. Plant Cell. 28(6):1388-1405.

Kirkwood M, Vohra P, Bawn M, Thilliez G, Pye H, Tanner J, Chintoan-Uta C, Branchu P, Petrovska L, Dallman T et al. 2021. Ecological niche adaptation of salmonella typhimurium u288 is associated with altered pathogenicity and reduced zoonotic potential. Commun Biol. 4(1):498.

Knaus BJ, Grünwald NJ. 2017. Vcfr: A package to manipulate and visualize variant call format data in r. Molecular Ecology Resources. 17(1):44-53.

Koinange EMK, Gepts P. 1992. Hybrid weakness in wild phaseolus-vulgaris l. Journal of Heredity. 83(2):135-139.

Kolde R. 2025. Pheatmap: Pretty heatmaps. 1.0.13 ed. R package.

Krueger F, James F, Ewels P, Afyounian E, Weinstein M, Schuster-Boeckler B, Hulselmans G, Sclamons. 2023. Felixkrueger/trimgalore: V0.6.10 - add default decompression path (0.6.10). Zenodo.

Kruskal WH, Wallis WA. 1987. Citation-classic - use of ranks in one-criterion variance analysis. Current Contents/Social & Behavioral Sciences. (40):20-20.

Kumar R, Harikrishna, Barman D, Ghimire OP, Gurumurthy S, Singh PK, Chinnusamy V, Padaria JC, Arora A. 2022. Stay-green trait serves as yield stability attribute under combined heat and drought stress in wheat (*triticum aestivum* l.). Plant Growth Regulation. 96(1):67-78.

Kwak M, Gepts P. 2009. Structure of genetic diversity in the two major gene pools of common bean (phaseolus vulgaris l., fabaceae). Theor Appl Genet. 118(5):979-992.

Kwak M, Kami JA, Gepts P. 2009. The putative mesoamerican domestication center of phaseolus vulgaris is located in the lerma-santiago basin of mexico. Crop Science. 49(2):554-563.

Kwak M, Toro O, Debouck DG, Gepts P. 2012. Multiple origins of the determinate growth habit in domesticated common bean (phaseolus vulgaris). Annals of Botany. 110(8):1573-1580.

Kwak M, Velasco D, Gepts P. 2008. Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (phaseolus vulgaris). Journal of Heredity. 99(3):283-291.

Labastida D, Ingvarsson PK, Rendón-Anaya M. 2023. Dissecting the genetic basis of drought responses in common bean using natural variation. Frontiers in Plant Science. 14.

Lei ZY, He Y, Li XF, He ZQ, Zhang YJ, Zhang WF, Liu F, Zhang YL. 2023. Domestication has reduced leaf water use efficiency associated with the anatomy of abaxial stomata in cotton. Journal of Experimental Botany. 74(3):878-888.

Leitao ST, Bicho MC, Pereira P, Paulo MJ, Malosetti M, Araujo SD, van Eeuwijk F, Patto MCV. 2021a. Common bean snp alleles and candidate genes affecting photosynthesis under contrasting water regimes. Horticulture Research. 8(1).

Leitao ST, Santos C, Araújo SD, Rubiales D, Vaz Patto MC. 2021b. Shared and tailored common bean transcriptomic responses to combined fusarium wilt and water deficit. Horticulture Research. 8(1).

Leivar P, Monte E, Al-Sady B, Carle C, Storer A, Alonso JM, Ecker JR, Quail PH. 2008. The arabidopsis phytochrome-interacting factor pif7, together with pif3 and pif4, regulates responses to prolonged red light by modulating phyb levels. Plant Cell. 20(2):337-352.

Lenth R. 2025. Emmeans: Estimated marginal means, aka least-squares means. R package version 1.11.2 ed.

Li B, Byrt C, Qiu J, Baumann U, Hrmova M, Evrard A, Johnson AAT, Birnbaum KD, Mayo GM, Jha D et al. 2016. Identification of a stelar-localized transport protein that facilitates root-to-shoot transfer of chloride in arabidopsis. Plant Physiology. 170(2):1014-1029.

Li DN, Gu BX, Huang CX, Shen JY, Wang X, Guo JA, Yu RQ, Mou SR, Guan QJ. 2023a. Functional study of *amorpha fruticosa wrky20* gene in response to drought stress. International Journal of Molecular Sciences. 24(15).

Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics. 25(14):1754-1760.

Li JY, Fu YL, Pike SM, Bao J, Tian W, Zhang Y, Chen CZ, Zhang Y, Li HM, Huang J et al. 2010. The *arabidopsis* nitrate transporter nrt1.8 functions in nitrate removal from the xylem sap and mediates cadmium tolerance. Plant Cell. 22(5):1633-1646.

Li M, Xu YW, Fu Q, Singh VP, Liu D, Li TX. 2020. Efficient irrigation water allocation and its impact on agricultural sustainability and water scarcity under uncertainty. Journal of Hydrology. 586.

Li P, Xiao L, Du QZ, Quan MY, Song YP, He YL, Huang WX, Xie JB, Lv CF, Wang D et al. 2023b. Genomic insights into selection for heterozygous alleles and woody traits in *populus tomentosa*. Plant Biotechnology Journal. 21(10):2002-2018.

Lisciani S, Marconi S, Le Donne C, Camilli E, Aguzzi A, Gabrielli P, Gambelli L, Kunert K, Marais D, Vorster BJ et al. 2024. Legumes and common beans in sustainable diets: Nutritional quality, environmental benefits, spread and use in food preparations. Frontiers in Nutrition. 11.

Liu C, Yang XX, Yan ZS, Fan YJ, Feng GJ, Liu DJ. 2019a. Analysis of differential gene expression in cold-tolerant vs. Cold-sensitive varieties of snap bean (*phaseolus vulgaris* l.) in response to low temperature stress. Genes & Genomics. 41(12):1445-1455.

Liu P, Zhang SB, Zhou B, Luo X, Zhou XF, Cai B, Jin YH, Niu D, Lin JX, Cao XF et al. 2019b. The histone h3k4 demethylase jmj16 represses leaf senescence in arabidopsis. Plant Cell. 31(2):430-443.

Liu X, Huang M, Fan B, Buckler ES, Zhang Z. 2016. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies (vol 12, e1005767, 2016). Plos One. 11(3).

Liu X, Wang DR, Chen GL, Wang X, Hao SY, Qu MS, Liu JY, Wang XF, You CX. 2024. *Mdtpr16,* an apple tetratricopeptide repeat (tpr)-like superfamily gene, positively regulates drought stress in apple. Plant Physiology and Biochemistry. 210.

Liu XS, Liang CC, Hou SG, Wang X, Chen DH, Shen JL, Zhang W, Wang M. 2020. The lrr-rlk protein hsl3 regulates stomatal closure and the drought stress response by modulating hydrogen peroxide homeostasis. Frontiers in Plant Science. 11.

Lobaton JD, Miller T, Gil J, Ariza D, de la Hoz JF, Soler A, Beebe S, Duitama J, Gepts P, Raatz B. 2018. Resequencing of common bean identifies regions of inter-gene pool introgression and provides comprehensive resources for molecular breeding. Plant Genome. 11(2).

Lozano-Arce D, García T, Gonzalez-Garcia LN, Guyot R, Chacón-Sánchez MI, Duitama J. 2023. Selection signatures and population dynamics of transposable elements in lima bean. Communications Biology. 6(1).

Lui H, Mihara M. 2024. Effects of adding coconut charcoal on soil physical properties and maize performance. International Journal of Environmental & Rural Development.

Luo X, Bai X, Sun XL, Zhu D, Liu BH, Ji W, Cai H, Cao L, Wu J, Hu MR et al. 2013. Expression of wild soybean *wrky20* in *arabidopsis* enhances drought tolerance and regulates aba signalling. Journal of Experimental Botany. 64(8):2155-2169.

Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. Science. 290(5494):1151-1155.

Lysenko EA, Kozuleva MA, Klaus AA, Pshybytko NL, Kusnetsov VV. 2023. Lower air humidity reduced both the plant growth and activities of photosystems i and ii under prolonged heat stress. Plant Physiology and Biochemistry. 194:246-262.

López-Hernández F, Burbano-Erazo E, León-Pacheco RI, Cordero-Cordero CC, Villanueva-Mejía DF, Tofiño-Rivera AP, Cortés AJ. 2023. Multi-environment genome-wide association studies of yield traits in common bean (*phaseolus vulgaris* l.) x tepary bean (*p. Acutifolius* a. Gray) interspecific advanced lines in humid and dry colombian caribbean subregions. Agronomy-Basel. 13(5).

Maciel FL, Echeverrigaray S, Gerald LTS, Grazziotin FG. 2003. Genetic relationships and diversity among brazilian cultivars and landraces of common beans (phaseolus vulgaris l.) revealed by aflp markers. Genetic Resources and Crop Evolution. 50(8):887-893.

Mackay IJ, Cockram J, Howell P, Powell W. 2021. Understanding the classics: The unifying concepts of transgressive segregation, inbreeding depression and heterosis and their central relevance for crop breeding. Plant Biotechnology Journal. 19(1):26-34.

Magrin GO, J.A. Marengo, J.-P. Boulanger, M.S. Buckeridge, E. Castellanos, G. Poveda, F.R. Scarano, Vicuña S. 2014. Central and south america. In: Climate change 2014: Impacts, adaptation, and vulnerability. Part b: Regional aspects. Contribution of working group ii to the fifth assessment report of the intergovernmental panel on climate change. Barros VR, C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L.White, editor. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Majumder M. 2015. Impact of urbanization on water shortage in face of climatic aberrations. Springer.

Mamidi S, Rossi M, Moghaddam SM, Annam D, Lee R, Papa R, McClean PE. 2013. Demographic factors shaped diversity in the two gene pools of wild common bean *phaseolus vulgaris* l. Heredity. 110(3):267-276.

Mangiafico. SS. 2025. Rcompanion: Functions to support extension education program evaluation. 2.5.0 ed. New Brunswick, New Jersey: Rutgers Cooperative Extension.

Mara CD, Huang TB, Irish VF. 2010. The floral homeotic proteins apetala3 and pistillata negatively regulate the genes implicated in light signaling. Plant Cell. 22(3):690-702.

Marchin RM, Ossola A, Leishman MR, Ellsworth DS. 2020. A simple method for simulating drought effects on plants. Frontiers in Plant Science. 10.

Martin G, Cottin A, Baurens FC, Labadie K, Hervouet C, Salmon F, Paulo-de-la-Reberdiere N, van den Houwe I, Sardos J, Aury JM et al. 2023. Interspecific introgression patterns reveal the origins of worldwide cultivated bananas in new guinea. Plant Journal. 113(4):802-818.

Maxted N, Ford-Lloyd BV, Jury S, Kell S, Scholten M. 2006. Towards a definition of a crop wild relative. Biodiversity and Conservation. 15(8):2673-2685.

McCouch S, Navabi ZK, Abberton M, Anglin NL, Barbieri RL, Baum M, Bett K, Booker H, Brown GL, Bryan GJ et al. 2020. Mobilizing crop biodiversity. Molecular Plant. 13(10):1341-1344.

Medynska-Juraszek A, Latawiec A, Królczyk J, Bogacz A, Kawalko D, Bednik M, Dudek M. 2021. Biochar improves maize growth but has a limited effect on soil properties: Evidence from a three-year field experiment. Sustainability. 13(7).

Mei FM, Chen B, Du LY, Li SM, Zhu DH, Chen N, Zhang YF, Li FF, Wang ZX, Cheng XX et al. 2022. A gain-of-function allele of a dreb transcription factor gene ameliorates drought tolerance in wheat. Plant Cell. 34(11):4472-4494.

Merrick LF, Burke AB, Zhang ZW, Carter AH. 2022. Comparison of single-trait and multi-trait genome-wide association models and inclusion of correlated traits in the dissection of the genetic architecture of a complex trait in a breeding program. Frontiers in Plant Science. 12.

Meuwissen THE, Sonesson AK, Gebregiwergis G, Woolliams JA. 2020. Management of genetic diversity in the era of genomics. Frontiers in Genetics. 11.

Meyer RS, DuVal AE, Jensen HR. 2012. Patterns and processes in crop domestication: An historical review and quantitative analysis of 203 global food crops. New Phytologist. 196(1):29-48.

Miklas PN, Johnson WC, Delorme R, Gepts P. 2001. Qtl conditioning physiological resistance and avoidance to white mold in dry bean. Crop Science. 41(2):309-315.

Moghaddam SM, Mamidi S, Osorno JM, Lee R, Brick M, Kelly J, Miklas P, Urrea C, Song QJ, Cregan P et al. 2016. Genome-wide association study identifies candidate loci underlying agronomic traits in a middle american diversity panel of common bean. Plant Genome. 9(3).

Munekage Y, Hojo M, Meurer J, Endo T, Tasaka M, Shikanai T. 2002. Pgr5 is involved in cyclic electron flow around photosystem i and is essential for photoprotection in arabidopsis. Cell. 110(3):361-371.

Munoz-Perea CG, Teran H, Allen RG, Wright JL, Westermann DT, Singh SP. 2006. Selection for drought resistance in dry bean landraces and cultivars. Crop Science. 46(5):2111-2120.

Mupangwa W, Nyagumbo I, Liben F, Chipindu L, Craufurd P, Mkuhlani S. 2021. Maize yields from rotation and intercropping systems with different legumes under conservation agriculture in contrasting agro-ecologies. Agriculture Ecosystems & Environment. 306.

Myers JR, Formiga AK, Janick J. 2022. Iconography of beans and related legumes following the columbian exchange. Frontiers in Plant Science. 13.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature. 403(6772):853-858.

Mylona P, Pawlowski K, Bisseling T. 1995. Symbiotic nitrogen-fixation. Plant Cell. 7(7):869-885.

Müllers Y, Postma JA, Poorter H, van Dusschoten D. 2022. Stomatal conductance tracks soil-to-leaf hydraulic conductance in faba bean and maize during soil drying. Plant Physiology. 190(4):2279-2294.

Nadeem MA, Yeken MZ, Shahid MQ, Habyarimana E, Yilmaz H, Alsaleh A, Hatipoglu R, Cilesiz Y, Khawar KM, Ludidi N et al. 2021. Common bean as a potential crop for future food security: An overview of past, current and future contributions in genomics, transcriptomics, transgenics and proteomics. Biotechnology & Biotechnological Equipment. 35(1):758-786.

Naumann G, Alfieri L, Wyser K, Mentaschi L, Betts RA, Carrao H, Spinoni J, Vogt J, Feyen L. 2018. Global changes in drought conditions under different levels of warming. Geophysical Research Letters. 45(7):3285-3296.

Nesporová T, Vitámvás P, Kosová K, Hynek R, Planchon S, Renaut J. 2024. Water-saving and water-spending strategy: The physiological, proteomic and metabolomic investigation of wheat response to drought and the following recovery. Plant Stress. 13.

Neumann PM. 2008. Coping mechanisms for crop plants in drought-prone environments. Annals of Botany. 101(7):901-907.

Nunes C, Moreira R, Pais I, Semedo J, Simoes F, Veloso MM, Scotti-Campos P. 2022. Cowpea physiological responses to terminal drought-comparison between four landraces and a commercial variety. Plants-Basel. 11(5).

Padilla-Chacón D, Peña Valdivia CB, García-Esteva A, Cayetano-Marcial MI, Kohashi Shibata J. 2019. Phenotypic variation and biomass partitioning during post-flowering in two common bean cultivars (*phaseolus vulgaris* l.) under water restriction. South African Journal of Botany. 121:98-104.

Papa R, Gepts P. 2003. Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (phaseolus vulgaris l.) from mesoamerica. Theoretical and Applied Genetics. 106(2):239-250.

Papathanasiou F, Ninou E, Mylonas I, Baxevanos D, Papadopoulou F, Avdikos I, Sistanis I, Koskosidis A, Vlachostergios DN, Stefanou S et al. 2022. The evaluation of common bean (*phaseolus vulgaris* l.) genotypes under water stress based on physiological and agronomic parameters. Plants-Basel. 11(18).

Patto MCV, Amarowicz R, Aryee ANA, Boye JI, Chung HJ, Martin-Cabrejas MA, Domoney C. 2015. Achievements and challenges in improving the nutritional quality of food legumes. Critical Reviews in Plant Sciences. 34(1-3):105-143.

Perez L, Rios DA, Giraldo DC, Twyman J, Blundo-Canto G, Prager SD, Ramirez-Villegas J. 2020. Determinants of vulnerability of bean growing households to climate variability in colombia. Climate and Development. 12(8):730-742.

Phiri AT, Njira KOW. 2022. Grain legume-based cropping systems' effects on soil organic carbon and nutrient dynamics. Agric Res.

Pinon MGT, Moghaddam SM, Lee RK, Merida JVC, DeYoung DJ, Reyes BA, Maredia M, Osorno JM, McClean PE. 2021. Genetic diversity of guatemalan climbing bean collections. Genetic Resources and Crop Evolution. 68(2):639-656.

Pironon S, Borrell JS, Ondo I, Douglas R, Phillips C, Khoury CK, Kantar MB, Fumia N, Gomez MS, Viruel J et al. 2020. Toward unifying global hotspots of wild and domesticated biodiversity. Plants-Basel. 9(9).

Plestenjak E, Meglic V, Sinkovic L, Pipan B. 2024. Factors influencing the emergence of heterogeneous populations of common bean (*phaseolus vulgaris* l.) and their potential for intercropping. Plants-Basel. 13(8).

Polania J, Poschenrieder C, Rao I, Beebe S. 2017a. Root traits and their potential links to plant ideotypes to improve drought resistance in common bean. Theoretical and Experimental Plant Physiology. 29(3):143-154.

Polania J, Rao IM, Cajiao C, Grajales M, Rivera M, Velasquez F, Raatz B, Beebe SE. 2017b. Shoot and root traits contribute to drought resistance in recombinant inbred lines of md 23-24 x sea 5 of common bean. Frontiers in Plant Science. 8.

Polania JA, Chater CCC, Covarrubias AA, Rao IM. 2020. Phaseolus species responses and tolerance to drought. In: Hasanuzzaman M, Araújo S, Gill SS, editors. The plant family fabaceae: Biology and physiological responses to environmental stresses. Singapore: Springer Singapore. p. 319-336.

Polania JA, Poschenrieder C, Beebe S, Rao IM. 2016. Effective use of water and increased dry matter partitioned to grain contribute to yield of common bean improved for drought resistance. Frontiers in Plant Science. 7.

Polania JA, Salazar-Chavarria V, Gonzalez-Lemes I, Acosta-Maspons A, Chater CCC, Covarrubias AA. 2022. Contrasting *phaseolus* crop water use patterns and stomatal dynamics in response to terminal drought. Frontiers in Plant Science. 13.

Power AG. 2010. Ecosystem services and agriculture: Tradeoffs and synergies. Philosophical Transactions of the Royal Society B-Biological Sciences. 365(1554):2959-2971.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics. 155(2):945-959.

Pérez-Vega E, Pañeda A, Rodríguez-Suárez C, Campa A, Giraldez R, Ferreira JJ. 2010. Mapping of qtls for morpho-agronomic and seed quality traits in a ril population of common bean (phaseolus vulgaris l.). Theoretical and Applied Genetics. 120(7):1367-1380.

Qian F, Zhao QY, Zhang TN, Li YL, Su YN, Li L, Sui JH, Chen S, He XJ. 2021. A histone h3k27me3 reader cooperates with a family of phd finger-containing proteins to regulate flowering time in arabidopsis. J Integr Plant Biol. 63(4):787-802.

Qin XH, Duan ZK, Zheng Y, Liu WC, Guo SY, Botella JR, Song CP. 2020. Abc1k10a, an atypical kinase, functions in plant salt stress tolerance. Bmc Plant Biology. 20(1).

R Core Team. 2024. R: A language and environment for statistical computing. R Foundation for Statistical Computing. R version 4.4.1 (2024-06-14) ed. Vienna, Austria.

Raggi L, Caproni L, Carboni A, Negri V. 2019. Genome-wide association study reveals candidate genes for flowering time variation in common bean (phaseolus vulgaris l.). Frontiers in Plant Science. 10.

Raman H, Raman R, Qiu Y, Yadav AS, Sureshkumar S, Borg L, Rohan M, Wheeler D, Owen O, Menz I et al. 2019. Gwas hints at pleiotropic roles for flowering locus t in flowering time and yield-related traits in canola. Bmc Genomics. 20(1).

Ramírez-Estrada CA, Sánchez E, Flores-Córdova MA, Pérez-Álvarez S, Noperi-Mosqueda LC, Chávez-Mendoza C. 2023. Photosynthetic efficiency in green bean plants through the application of omeprazole and melatonin at low doses. International Journal of Plant Biology. 14:864-878.

Rao MJ, Zheng BS. 2025. The role of polyphenols in abiotic stress tolerance and their antioxidant properties to scavenge reactive oxygen species and free radicals. Antioxidants. 14(1).

Rawal V, Navarro DK. 2019. The global economy of pulses. Rome: FAO.

Rehling F, Sandner TM, Matthies D. 2021. Biomass partitioning in response to intraspecific competition depends on nutrients and species characteristics: A study of 43 plant species. Journal of Ecology. 109(5):2219-2233.

Reinprecht Y, Schram L, Perry GE, Morneau E, Smith TH, Pauls KP. 2024. Mapping yield and yield-related traits using diverse common bean germplasm. Frontiers in Genetics. 14.

Renard D, Mahaut L, Noack F. 2023. Crop diversity buffers the impact of droughts and high temperatures on food production. Environmental Research Letters. 18(4).

Rendon-Anaya M, Herrera-Estrella A, Gepts P, Delgado-Salinas A. 2017a. A new species of phaseolus (leguminosae, papilionoideae) sister to phaseolus vulgaris, the common bean. Phytotaxa. 313(3):259-266.

Rendon-Anaya M, Montero-Vargas JM, Saburido-Alvarez S, Vlasova A, Capella-Gutierrez S, Ordaz-Ortiz JJ, Aguilar OM, Vianello-Brondani RP, Santalla M, Delaye L et al. 2017b. Genomic history of the origin and domestication of common bean unveils its closest sister species. Genome Biology. 18.

Repinski SL, Kwak M, Gepts P. 2012. The common bean growth habit gene pvtfl1y is a functional homolog of arabidopsis tfl1. Theoretical and Applied Genetics. 124(8):1539-1547.

Roeber VM, Schmuelling T, Cortleven A. 2022. The photoperiod: Handling and causing stress in plants. Frontiers in Plant Science. 12.

Rosales MA, Ocampo E, Rodríguez-Velentín R, Olvera-Carrillo Y, Acosta-Gallegos J, Covarrubias AA. 2012. Physiological analysis of common bean (*phaseolus vulgaris* l.) cultivars uncovers characteristics related to terminal drought resistance. Plant Physiology and Biochemistry. 56:24-34.

Rosner B, Glynn RJ, Lee MLT. 2006. The wilcoxon signed rank test for paired comparisons of clustered data. Biometrics. 62(1):185-192.

Rossi M, Bitocchi E, Bellucci E, Nanni L, Rau D, Attene G, Papa R. 2009. Linkage disequilibrium and population structure in wild and domesticated populations of phaseolus vulgaris l. Evolutionary Applications. 2(4):504-522.

Rushton DL, Tripathi P, Rabara RC, Lin J, Ringler P, Boken AK, Langum TJ, Smidt L, Boomsma DD, Emme NJ et al. 2012. Wrky transcription factors: Key components in abscisic acid signalling. Plant Biotechnology Journal. 10(1):2-11.

Sablowski R. 2007. Flowering and determinacy in arabidopsis. Journal of Experimental Botany. 58(5):899-907.

Sadohara R, Izquierdo P, Alves FC, Porch T, Beaver J, Urrea CA, Cichy K. 2022. The phaseolus vulgaris l. Yellow bean collection: Genetic diversity and characterization for cooking time. Genetic Resources and Crop Evolution.

Sadohara R, Wiesinger JA, Thompson HJ, Glahn RP, Cichy K. 2025. Yellow bean (*phaseolus vulgaris* l.) germplasm with less dietary fiber have shorter cooking times and more bioavailable iron. Current Research in Food Science. 10.

Sajgalik M, Ondreickova K, Hauptvogel P, Mihalik D, Glasa M, Kraic J. 2019. Higher effectiveness of new common bean (phaseolus vulgaris l.) germplasm acquisition by collecting expeditions associated with molecular analyses. Sustainability. 11(19).

Salehin M, Li BH, Tang M, Katz E, Song L, Ecker JR, Kliebenstein DJ, Estelle M. 2019. Auxin-sensitive aux/iaa proteins mediate drought tolerance in arabidopsis by regulating glucosinolate levels. Nature Communications. 10.

Salojaervi J, Rambani A, Yu Z, Guyot R, Strickler S, Lepelley M, Wang C, Rajaraman S, Rastas P, Zheng CF et al. 2024. The genome and population genomics of allopolyploid *coffea arabica* reveal the diversification history of modern coffee cultivars. Nature Genetics. 56(4).

Sanchez D, Ben Sadoun S, Mary-Huard T, Allier A, Moreau L, Charcosset A. 2023. Improving the use of plant genetic resources to sustain breeding programs' efficiency. Proceedings of the National Academy of Sciences of the United States of America. 120(14).

Sanchez-Corrionero A, Perez-Garcia P, Cabrera J, Silva-Navas J, Perianez-Rodriguez J, Gude I, del Pozo JC, Moreno-Risueno MA. 2019. Root patterning and regeneration are mediated by the quiescent center and involve bluejay, jackdaw and scarecrow regulation of vasculature factors. bioRxiv. (803973).

Sanjak JS, Long AD, Thornton KR. 2017. A model of compound heterozygous, loss-of-function alleles is broadly consistent with observations from complex-disease gwas datasets. Plos Genetics. 13(1).

Santalla M, Menendez-Sevillano MC, Monteagudo AB, De Ron AM. 2004. Genetic diversity of argentinean common bean and its evolution during domestication. Euphytica. 135(1):75-87.

Santalla M, Rodino AP, De Ron AM. 2002. Allozyme evidence supporting southwestern europe as a secondary center of genetic diversity for the common bean. Theoretical and Applied Genetics. 104(6-7):934-944.

Sato H, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. 2024. Complex plant responses to drought and heat stress under climate change. Plant Journal. 117(6):1873-1892.

Savic A, Pipan B, Vasic M, Meglic V. 2021. Genetic diversity of common bean (phaseolus vulgaris l.) germplasm from serbia, as revealed by single sequence repeats (ssr). Scientia Horticulturae. 288.

Schier HE, Eliot KA, Herron SA, Landfried LK, Migicovsky Z, Rubin MJ, Miller AJ. 2019. Comparative analysis of perennial and annual phaseolus seed nutrient concentrations. Sustainability. 11(10).

Schmit R, de Melo RC, Trevisani N, Guidolin AF, Coimbra JLM. 2019. Screening and agronomic benefits of the stay-green trait in common bean genotypes. Bioscience Journal. 35(3):869-877.

Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, Jenkins J, Shu SQ, Song QJ, Chavarro C et al. 2014. A reference genome for common bean and genome-wide analysis of dual domestications. Nature Genetics. 46(7):707-713.

Scullion J, Neale S, Philipps L. 2002. Comparisons of earthworm populations and cast properties in conventional and organic arable rotations. Soil Use and Management. 18:293-300.

Sedlar A, Zupin M, Maras M, Razinger J, Sustar-Vozlic J, Pipan B, Meglic V. 2020. Qtl mapping for drought-responsive agronomic traits associated with physiology, phenology, and yield in an andean intra-gene pool common bean population. Agronomy-Basel. 10(2).

Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nature Genetics. 44(7):825-U144.

Shahzaib M, Khan UM, Azhar MT, Atif RM, Khan SH, Zaman QU, Rana IA. 2024. Phylogenomic curation of ovate family proteins (ofps) in the u's triangle of brassica l. Indicates stress-induced growth modulation. Plos One. 19(1).

Shavrukov Y, Kurishbayev A, Jatayev S, Shvidchenko V, Zotova L, Koekemoer F, de Groot S, Soole K, Langridge P. 2017. Early flowering as a drought escape mechanism in plants: How can it aid wheat production? Frontiers in Plant Science. 8.

Shinohara H. 2021. Root meristem growth factor rgf, a sulfated peptide hormone in plants. Peptides. 142.

Singh J, Gezan SA, Vallejos CE. 2019. Developmental pleiotropy shaped the roots of the domesticated common bean (phaseolus vulgaris). Plant Physiology. 180(3):1467-1479.

Singh SP, Gepts P, Debouck DG. 1991a. Races of common bean (phaseolus-vulgaris, fabaceae). Economic Botany. 45(3):379-396.

Singh SP, Nodari R, Gepts P. 1991b. Genetic diversity in cultivated common bean .1. Allozymes. Crop Science. 31(1):19-23.

Singh SP, Schwartz HF. 2010. Breeding common bean for resistance to diseases: A review. Crop Science. 50(6):2199-2223.

Sitaraman J, Bui M, Liu ZC. 2008. Leunig_homolog and leunig perform partially redundant functions during arabidopsis embryo and floral development. Plant Physiology. 147(2):672-681.

Smith MR, Veneklaas E, Polania J, Rao IM, Beebe SE, Merchant A. 2019. Field drought conditions impact yield but not nutritional quality of the seed in common bean (phaseolus vulgaris l.). PLoS One. 14(6):e0217099.

Sofi PA, Rehman K, Gull M, Kumari J, Djanaguiraman M, Prasad PVV. 2021. Integrating root architecture and physiological approaches for improving drought tolerance in common bean (*phaseolus vulgaris* l.). Plant Physiology Reports. 26(1):4-22.

Soler-Garzón A, Mulube M, Kamfwa K, Lungu DM, Hamabwe S, Roy J, Salegua V, Fourie D, Porch TG, McClean PE et al. 2024. Gwas of resistance to three bacterial diseases in the andean common bean diversity panel. Frontiers in Plant Science. 15.

Stopnisek N, Shade A. 2021. Persistent microbiome members in the common bean rhizosphere: An integrated analysis of space, time, and plant genotype. Isme Journal. 15(9):2708-2722.

Su HG, Li B, Song XY, Ma J, Chen J, Zhou YB, Chen M, Min DH, Xu ZS, Ma YZ. 2019. Genome-wide analysis of the dyw subgroup ppr gene family and identification of *gmppr4* responses to drought stress. International Journal of Molecular Sciences. 20(22).

Su JH, He BQ, Li PY, Yu BY, Cen QW, Xia LF, Jing Y, Wu FB, Karnik R, Xue DW et al. 2024. Overexpression of tonoplast ca2+-atpase in guard cells synergistically enhances

stomatal opening and drought tolerance. Journal of Integrative Plant Biology. 66(8):1587-1602.

Suarez JC, Polania JA, Anzola JA, Contreras AT, Mendez DL, Vanegas JI, Noriega JE, Rodriguez L, Urban MO, Beebe S et al. 2021a. Influence of nitrogen supply on gas exchange, chlorophyll fluorescence and grain yield of breeding lines of common bean evaluated in the amazon region of colombia. Acta Physiologiae Plantarum. 43(4).

Suarez JC, Urban MO, Contreras AT, Grajales MA, Cajiao C, Beebe SE, Rao IM. 2021b. Adaptation of interspecific mesoamerican common bean lines to acid soils and high temperature in the amazon region of colombia. Plants (Basel). 10(11).

Suarez-Gonzalez A, Lexer C, Cronk QCB. 2018. Adaptive introgression: A plant perspective. Biology Letters. 14(3).

Suarez-Martinez SE, Ferriz-Martinez RA, Campos-Vega R, Elton-Puente JE, Carbot KD, Garcia-Gasca T. 2016. Bean seeds: Leading nutraceutical source for human health. Cyta-Journal of Food. 14(1):131-137.

Sánchez-Reinoso AD, Ligarreto-Moreno GA, Trepo-Díaz HR. 2020. Evaluation of drought indices to identify tolerant genotypes in common bean bush (*phaseolus vulgaris* l.). Journal of Integrative Agriculture. 19(1):99-107.

Tabassum T, Farooq M, Ahmad R, Zohaib A, Wahid A, Shahid M. 2018. Terminal drought and seed priming improves drought tolerance in wheat. Physiology and Molecular Biology of Plants. 24(5):845-856.

Tai APK, Martin MV, Heald CL. 2014. Threat to future global food security from climate change and ozone air pollution. Nature Climate Change. 4(9):817-821.

Tao LY, Yu GH, Chen HN, Wang B, Jiang LL, Han XR, Lin GL, Cheng XG. 2022. Gene specifically recognizing to the universal dre elements is a transcriptional activator improving drought tolerance in tomato. Scientia Horticulturae. 295.

Tar'an B, Michaels TE, Pauls KP. 2002. Genetic mapping of agronomic traits in common bean. Crop Science. 42(2):544-556.

Tardieu F. 2012. Any trait or trait-related allele can confer drought tolerance: Just design the right drought scenario. Journal of Experimental Botany. 63(1):25-31.

Teran H, Singh SP. 2002a. Comparison of sources and lines selected for drought resistance in common bean. Crop Science. 42(1):64-70.

Teran H, Singh SP. 2002b. Selection for drought resistance in early generations of common bean populations. Canadian Journal of Plant Science. 82(3):491-497.

Thomas H, Ougham H. 2014. The stay-green trait. Journal of Experimental Botany. 65(14):3889-3900.

Tohme J, Gonzalez DO, Beebe S, Duque MC. 1996. Aflp analysis of gene pools of a wild bean core collection. Crop Science. 36(5):1375-1384.

Trucchi E, Benazzo A, Lari M, Iob A, Vai S, Nanni L, Bellucci E, Bitocchi E, Raffini F, Xu CM et al. 2021. Ancient genomes reveal early andean farmers selected common beans while preserving diversity. Nature Plants. 7(2).

Tseng TS, Salomé PA, McClung CR, Olszewski NE. 2004. Spindly and gigantea interact and act in arabidopsis thaliana pathways involved in light responses, flowering, and rhythms in cotyledon movements. Plant Cell. 16(6):1550-1563.

Tuberosa R. 2012. Phenotyping for drought tolerance of crops in the genomics era. Frontiers in Physiology. 3.

Turner S. 2018. Qqman: An r package for visualizing gwas results using q-q and manhattan plots. The Journal of Open Source Software.

Ugwuanyi S, Udengwu OS, Snowdon RJ, Obermeier C. 2022. Novel candidate loci for morpho-agronomic and seed quality traits detected by targeted genotyping-by-sequencing in common bean. Frontiers in Plant Science. 13.

Umezawa T, Fujita M, Fujita Y, Yamaguchi-Shinozaki K, Shinozaki K. 2006. Engineering drought tolerance in plants: Discovering and tailoring genes to unlock the future. Current Opinion in Biotechnology. 17(2):113-122.

Urban A, Urban J. 2024. A decline in stomatal conductance is the primary reason for low photosynthesis in veteran pedunculate oak trees. Forests. 15(12).

Valdisser PAMR, Pereira WJ, Almeida JE, Müller BSF, Coelho GRC, de Menezes IPP, Vianna JPG, Zucchi MI, Lanna AC, Coelho ASG et al. 2017. In-depth genome characterization of a brazilian common bean core collection using dartseq high-density snp genotyping. Bmc Genomics. 18.

Van der Auwera G, O'Connor B. 2020. Genomics in the cloud: Using docker, gatk, and wdl in terra. O'Reilly Media.

Vargas Y, Mayor-Duran VM, Buendia HF, Ruiz-Guzman H, Raatz B. 2021. Physiological and genetic characterization of heat stress effects in a common bean ril population. Plos One. 16(4).

Velho LPS, De Melo RC, Bernardy JPF, Grigolo S, Guidolin AF, Coimbra JLM. 2018. Root distribution and its association with bean growth habit. Anais Da Academia Brasileira De Ciencias. 90(2):1837-1844.

Verma S, Negi NP, Pareek S, Mudgal G, Kumar D. 2022. Auxin response factors in plant adaptation to drought and salinity stress. Physiologia Plantarum. 174(3).

Vexler L, Leyva-Perez MD, Konkolewska A, Clot CR, Byrne S, Griffin D, Ruttink T, Hutten RCB, Engelen C, Visser RGF et al. 2024. Qtl discovery for agronomic and quality traits in diploid potato clones using potatomash amplicon sequencing. G3-Genes Genomes Genetics. 14(10).

Vicente-Serrano SM, Quiring SM, Pena-Gallardo M, Yuan SS, Dominguez-Castro F. 2020. A review of environmental droughts: Increased risk under global warming? Earth-Science Reviews. 201.

Vidak M, Lazarevic B, Javornik T, Satovic Z, Carovic-Stanko K. 2022. Seed water absorption, germination, emergence and seedling phenotypic characterization of the common bean landraces differing in seed size and color. Seeds. 1(4):324-339.

Villa T, Maxted N, Scholten M, Ford-Lloyd B. 2005. Defining and identifying crop landraces. Plant Genetic Resources. 3(3):373-384.

Villordo-Pineda E, González-Chavira MM, Giraldo-Carbajo P, Acosta-Gallegos JA, Caballero-Pérez J. 2015. Identification of novel drought-tolerant-associated snps in common bean (*phaseolus vulgaris*). Frontiers in Plant Science. 6.

Visual Crossing Corporation. 2024. Visual crossing weather 06/2023 - 09/2023. www.visualcrossing.com/.

Visual Crossing Corporation. 2025. Visual crossing weather 12/2022 - 03/2023. www.visualcrossing.com.

Wahl V, Brand LH, Guo YL, Schmid M. 2010. The fantastic four proteins influence shoot meristem size in arabidopsis thaliana. Bmc Plant Biology. 10.

Wang BM, Liu C, Zhang DF, He CM, Zhang JR, Li ZX. 2019. Effects of maize organ-specific drought stress response on yields from transcriptome analysis. Bmc Plant Biology. 19.

Wang HH, Lu S, Guan XY, Jiang Y, Wang B, Hua J, Zou BH. 2022. Dehydration-responsive element binding protein 1c, 1e, and 1g promote stress tolerance to chilling, heat, drought, and salt in rice. Frontiers in Plant Science. 13.

Wang JB, Zhang ZW. 2021. Gapit version 3: Boosting power and accuracy for genomic association and prediction. Genomics Proteomics & Bioinformatics. 19(4):629-640.

Wang Q, Shi HF, Huang RS, Ye R, Luo YR, Guo ZF, Lu SY. 2021. Air12 confers cold tolerance through regulation of the cbf cold response pathway and ascorbate homeostasis. Plant Cell and Environment. 44(5):1522-1533.

Wang XY, Li XM, Zhao W, Hou XM, Dong SK. 2024. Current views of drought research: Experimental methods, adaptation mechanisms and regulatory strategies. Frontiers in Plant Science. 15.

Wang ZY, Fei ZJ. 2025. Understanding the genomic basis to empower sweet potato breeding. Nature Plants. 11(1):14-15.

Warschefsky E, Penmetsa RV, Cook DR, von Wettberg EJB. 2014. Back to the wilds: Tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. American Journal of Botany. 101(10):1791-1800.

Wedger MJ, Xiao E, Butts TR, Chlapecka JL, Webster LC, Olsen KM. 2024. Recent crop-to-weed adaptive introgression has reshaped the genomic composition and geographical structure of us weedy rice (*oryza* spp.). Molecular Ecology.

WEF. 2024. Global risks report 2024. World Economic Forum.

Wei T, Simko V. 2021. R package 'corrplot': Visualization of a correlation matrix. 0.92 ed.

Weller JL, Vander Schoor JK, Perez-Wright EC, Hecht V, Gonzalez AM, Capel C, Yuste-Lisbona FJ, Lozano R, Santalla M. 2019. Parallel origins of photoperiod adaptation following dual domestications of common bean. Journal of Experimental Botany. 70(4):1209-1219.

White JW, Laing DR. 1989. Photoperiod response of flowering in diverse genotypes of common bean (phaseolus-vulgaris). Field Crops Research. 22(2):113-128.

Whitney KD, Randell RA, Rieseberg LH. 2006. Adaptive introgression of herbivore resistance traits in the weedy sunflower *helianthus annuus*. American Naturalist. 167(6):794-807.

Https://www.Who.Int/health-topics/nutrition#tab=tab_1. 2023. [accessed].

Wickham H. 2016. Ggplot2: Elegant graphics for data analysis. Springer-Verlag New York.

Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with kraken 2. Genome Biology. 20(1).

Highest water stressed countries. 2023. World Resource Institude; [accessed 2025]. https://www.wri.org/insights/highest-water-stressed-countries.

Wortmann CS, Kirkby RA, Eledu CA, Allen DJ. 1998. Atlas of common bean (phaseolus vulgaris l.) production in africa. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. 131 p.

Wu J, Wang LF, Fu JJ, Chen JB, Wei SH, Zhang SL, Zhang J, Tang YS, Chen ML, Zhu JF et al. 2020. Resequencing of 683 common bean genotypes identifies yield component trait associations across a north-south cline. Nature Genetics. 52(1):118-+.

Wu JF, Wang Y, Wu SH. 2008. Two new clock proteins, lwd1 and lwd2, regulate arabidopsis photoperiodic flowering. Plant Physiology. 148(2):948-959.

Wu L, Chang YJ, Wang LF, Wang SM, Wu J. 2021. Genome-wide association analysis of drought resistance based on seed germination vigor and germination rate at the bud stage in common bean. Agronomy Journal. 113(4).

Wu L, Chang YJ, Wang LF, Wang SM, Wu J. 2024. Genome-wide association study dissecting drought resistance- associated loci based on physiological traits in common bean. Journal of Integrative Agriculture. 23(11):3657-3671.

Xu S, Chen M, Feng T, Zhan L, Zhou L, Yu G. 2021. Use ggbreak to effectively utilize plotting space to deal with large datasets and outliers. Frontiers in Genetics. 12:774846.

Xu Y, Yu ZP, Zhang D, Huang JG, Wu CG, Yang GD, Yan K, Zhang SZ, Zheng CC. 2018. Cystm, a novel non-secreted cysteine-rich peptide family, involved in environmental stresses in *arabidopsis thaliana*. Plant and Cell Physiology. 59(2):423-438.

Yaneva TG, Wiczkowski W, Marchev AS, Iserliyska D, Georgiev MI, Tomlekova NB. 2024. Evaluation of polyphenols synthesized in mature seeds of common bean (*phaseolus vulgairs* l.) advanced mutant lines. International Journal of Molecular Sciences. 25(11).

Yang H, Shukla MK, Gonzalez A, Yuan YS. 2024. Determination of crop coefficients for flood-irrigated winter wheat in southern new mexico using three eto estimation methods. Water. 16(17).

Yin LM, Liu C, Liang ZC, Liu DJ, Feng GJ, Yan ZS, Yang XX. 2023. Transcriptome analysis of molecular mechanisms underlying phenotypic variation in *phaseolus vulgaris* mutant 'nts'. Phyton-International Journal of Experimental Botany. 92(11):2981-2998.

Yu MH, Ding GD, Gao GL, Zhao YY, Yan L, Sai K. 2015. Using plant temperature to evaluate the response of stomatal conductance to soil moisture deficit. Forests. 6(10):3748-3762.

Yu Y, Wang H, You CJ, Chen XM. 2026. Plant microrna maturation and function. Nature Reviews Molecular Cell Biology. 27(1):55-70.

Zeven AC. 1998. Landraces: A review of definitions and classifications. Euphytica. 104(2):127-139.

Zhang CH, Nie X, Kong WL, Deng XX, Sun T, Liu XH, Li YS. 2022. Genome-wide identification and evolution analysis of the gibberellin oxidase gene family in six gramineae crops. Genes. 13(5).

Zhang H, Zhu JK. 2025. Epigenetic gene regulation in plants and its potential applications in crop improvement. Nature Reviews Molecular Cell Biology. 26(1):51-67.

Zhang XY, Blair MW, Wang SM. 2008. Genetic diversity of chinese common bean (phaseolus vulgaris l.) landraces assessed with simple sequence repeat markers. Theoretical and Applied Genetics. 117(4):629-640.

Zhang ZW, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK, Gore MA, Bradbury PJ, Yu JM, Arnett DK, Ordovas JM et al. 2010. Mixed linear model approach adapted for genome-wide association studies. Nature Genetics. 42(4):355-U118.

Zhou AM, Sun HW, Dai SY, Feng S, Zhang JZ, Gong SF, Wang JG. 2019. Identification of transcription factors involved in the regulation of flowering in adonis amurensis through combined rna-seq transcriptomics and itraq proteomics. Genes. 10(4).

Zhou C, Wang HY, Zhu XB, Li YQ, Zhang B, Tadege M, Wu SH, Qi ZM, Xia ZJ. 2025. Functional genomics: From soybean to legume. International Journal of Molecular Sciences. 26(13).