bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Polygenic architecture of flowering time and its relationship with local environments in the grass *Brachypodium distachyon.*

Nikolaos Minadakis[1], Lars Kaderli[1], Robert Horvath[1], Yann Bourgeois[2], Wenbo Xu[1], Michael Thieme[1], Daniel P. Woods[3,4] and Anne C. Roulin[1*]

[1] Dept. of Plant and Microbial Biology, University of Zürich, Zollikerstr. 107, 8008 Zürich, Switzerland.
[2] DIADE, University of Montpellier, CIRAD, IRD, Montpellier, France.
[3] Department of Plant Sciences, University of California-Davis, 104 Robbins Hall Davis, CA 95616.
[4] Howard Hughes Medical Institute, Chevy Chase, Maryland, United States of America.

[*] **Corresponding author:** anne.roulin@botinst.uzh.ch

Abstract

Synchronizing the timing of reproduction with the environment is crucial in the wild. Among the multiple mechanisms annual plants evolved to sense their environment, the requirement of cold-mediated vernalization is a major process that prevents individuals from flowering during winter. In many annual plants including crops, both a long and short vernalization requirement can be observed within species, resulting in so-called early-(spring) and late (winter)-flowering genotypes. Here, using the grass model *Brachypodium distachyon*, we explored the link between vernalization requirement, flowering time, environmental variation, and diversity at flowering genes by combining measurements under greenhouse and outdoor conditions. These experiments confirmed that *B. distachyon* natural accessions display large differences regarding vernalization requirements and ultimately flowering time. We underline significant, albeit quantitative effects of current environmental conditions on flowering time. Population genomics in 332 natural accessions revealed that eight well-characterized flowering-time genes contribute significantly to flowering time variation and display signs of polygenic selection. Flowering-time genes, however, do not colocalize with GWAs peaks obtained with outdoor measurements, indicating that flowering-time genes may not largely contribute to flowering time variation in the wild. Altogether, our study fosters our understanding of the polygenic architecture of flowering time in a natural grass system and opens new avenues of research to investigate the gene-by-environment interaction at play for this trait.

Key words: Flowering time, adaptation, vernalization, *B. distachyon*, grasses, polygenic selection.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Introduction

The induction of reproduction is a critical fitness-related trait in the wild (Gaudinier and Blackman 2020), as a failure to produce offspring during a favorable growing season may lead to the extinction of the individual's genotype. In many annual plant species adapted to temperate climates, plantlets establish themselves in the fall and overwinter before flowering and producing seeds in more favorable spring conditions (Chouard 1960; Blackman 2017). Vernalization, the prolonged exposure to cold necessary to render plants competent to flower (Chouard 1960), is hence a key component of plant reproduction as it prevents individuals from flowering prior to winter. The adaptive potential and the genetic architecture of flowering time has been studied in an unrivaled manner in *Arabidopsis thaliana* (for review Andrés and Coupland 2012; Blümel et al. 2015; Takou et al. 2019) due to the broad geographical distribution of the species and large genomic resources developed by the community (but see Hall et al. 2006; Monnahan and Kelly 2017; Yan et al. 2021 for works on other *Brassicaceae* and *Mimulus guttatus*). Vernalization is yet controlled by different genes in different plant groups and likely evolved independently multiple times during flowering plant diversification (Ream et al. 2012; Bouché et al., 2017; Raissig and Woods 2022). Moreover, specific crop flowering-time genes (e.g. *ID1* in maize) are for instance lacking homologs in *A. thaliana* (Blümel et al. 2015). As a wild monocot, the model for the temperate grasses *Brachypodium distachyon* constitute a prime system to study the evolution of flowering-time genes. In this context, we made use of the diversity panel developed for this species (Gordon et al. 2017; Gordon et al. 2020; Skalska et al. 2020; Stritt et al. 2022, Minadakis et al. 2023) to expand our knowledge on the adaptive potential and polygenic architecture of flowering time in grasses.

Initially established as a model for bioenergy crops (International Brachypodium Initiative 2010), the grass species *B. distachyon* has more recently become a prime model for developmental biology (Woods et al. 2017[a]; Woods et al. 2019; Nunes et al. 2020; Hasterok et al. 2022; Raissig and Woods 2022; Zhang et al. 2022; Slawinska et al.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

2023), evolutionary genomics (Eichten et al. 2016; Gordon, Contreras-moreira, et al. 2017; Bourgeois et al. 2018; Stritt et al. 2018; Gordon et al. 2020; Stritt et al. 2020) and molecular ecology (Del'Acqua et al. 2014; Wilson et al. 2019; Skalska et al. 2020; Stritt et al. 2022; Minadakis et al. 2023). In addition to a near base-perfect reference genome (International Brachypodium Initiative 2010), a diversity panel composed of 332 accessions spanning from Spain to Iraq has been sequenced (Gordon et al. 2017; Gordon et al. 2020; Skalska et al. 2020; Stritt et al. 2022, Minadakis et al. 2023), opening new avenues of research in this system (Minadakis et al. 2023). We previously showed that *B. distachyon* accessions cluster into three main genetic lineages (A, B and C), which further divide into five main genetic clades: the ancestral C clade in Italy and Balkans, the B_West clade in Spain and France, the B_East clade spanning from Turkey to Caucasus and Iraq, the A_Italia clade in Italy as well as the A_East clade in Turkey and Greece (Stritt et al. 2022, Minadakis et al. 2023). These natural accessions are found in diverse habitats (Bourgeois et al. 2018; Minadakis et al. 2023) making *B. distachyon* an ideal model to investigate how genetic and environmental factors interact to shape traits.

B. distachyon accessions display large phenotypic variation with regard to flowering time (e.g. Ream et al. 2014; Gordon et al. 2017; Woods et al., 2019; Sharma et al. 2017), with some accessions requiring little to no vernalization to flower rapidly (early flowering accessions) in certain photoperiods, while other accessions require a few weeks to several months of vernalization in order to flower (late flowering accessions) (Ream et al., 2014; Gordon et al., 2017; Woods et al., 2019). These flowering differences have been described as potentially adaptive and responsible for population distribution according to climate variation (Gordon et al. 2017; Woods et al. 2019; Skalska et al. 2020). However, the extent to which variation in vernalization requirement and ultimately flowering time correlates with local environmental conditions has yet to be formally tested in this species. For instance, whether late-flowering genotypes, which require long vernalization treatments, have been selected to complete their life cycle at a slower rate to overcome harsher or longer winter, as observed in Swedish

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

populations of *A. thaliana* (Ågren et al. 2017), remains an open question. Hence, we combined flowering-time measurements under greenhouse and outdoor conditions and asked i) Does flowering time correlate with environmental variables and show signs of local adaptation? ii) What is the respective contribution of flowering time genes to flowering time variation? iii) Are known flowering-time genes contributing to flowering time variation in the wild?

Results

Flowering time measurements under greenhouse conditions

We selected 61 accessions (Figure 1A) from the *B. distachyon* diversity panel (Minadakis et al. 2023) for our flowering time experiment. Those accessions were chosen to represent all five genetic clades and occur, when possible, along latitudinal gradients. Briefly, we submitted plants to five vernalization treatments (2, 4, 6, 8 or 10 weeks at 4°C) with three replicates each and measured how long plants took to flower after the return to warm conditions (Figure S1, Table S1 for the raw data). For five out of the 61 accessions (Veg12 from A_East; Ren4, Lb1, Lb23 and Cm7 from A_Italia), none of the replicates flowered by the end of the experiment (Figure 1A, Figure S1) despite normal growth. All subsequent analyses were hence performed on the 56 accessions for which flowering time data were collected. We aggregated the results per accession and vernalization treatment (see methods) to estimate the fastest time a given accession took to flower from i) the day plants were moved to the vernalization chamber (time to flower including vernalization time) and ii) the day after the return to warm conditions (time to flower after vernalization, Figure S1). Both variables were highly correlated ($R2 = 0.92$, P-value $< 2.2e-16$). As more representative of the life cycle of plants, we nonetheless used for rest of the study time to flower including vernalization as the measure of flowering time.

We observed a strong partitioning of the phenotypes per genetic clades. Accessions from the C, B_East and B_West clades flowered significantly earlier than accessions

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

from the A_East and A_Italia clades, while no significant differences were observed within the A and B lineages (Figure 1B; Table S2 for P-values). This reflects that accessions from the A lineage display overall a significantly longer life cycle (late-flowering genotypes) than the ones from the B and C lineages (early-flowering genotypes) due to longer vernalization time requirements as well as longer time to flower after the return to warm conditions (Figure S1).

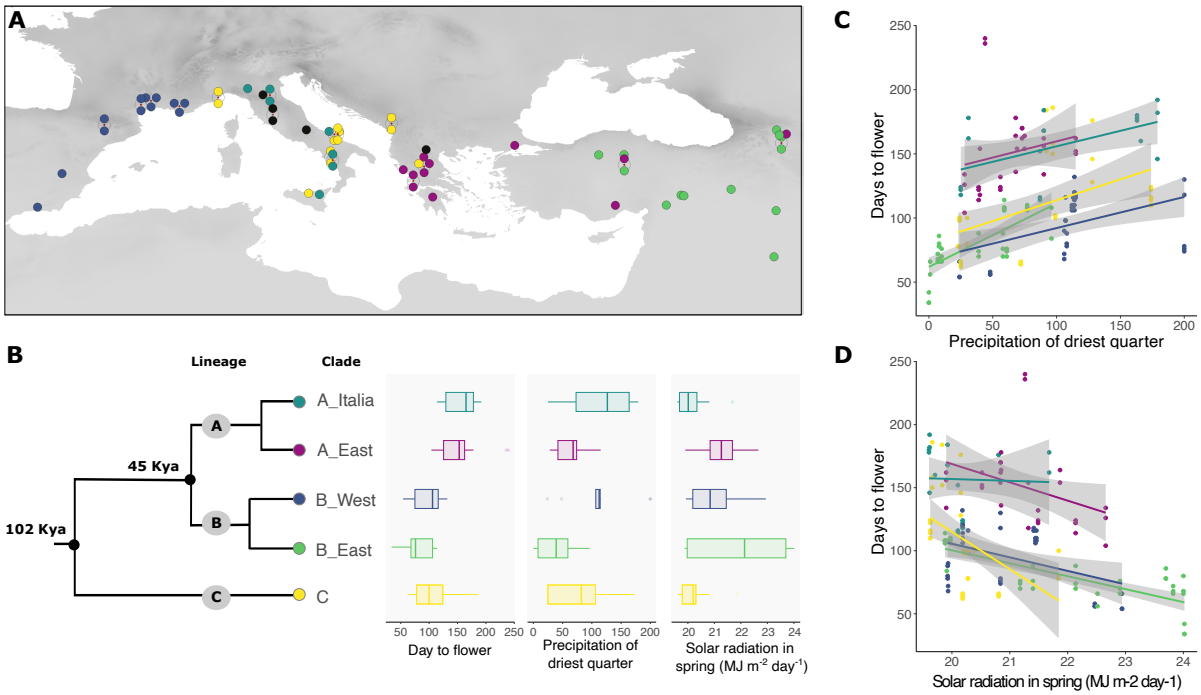

Fig. 1: Geographical origins of the samples and flowering time variation (greenhouse experiment) A) Map displaying the location of a given accession as well as their genetic clade of origin (C in yellow, A_East in magenta; A_Italy in turquoise; B_East in green and B_West in dark blue). Accessions that did not flower by the end of the experiment are depicted in black B) Schematic phylogeny and distribution of flowering time-, precipitation of driest quarter- (bio17) and solar radiation in spring per cluster C) Association between flowering time and Precipitation of driest quarter (bio17) and, D) Association between flowering time and solar radiation in spring. For panels C and D, individual replicates are plotted.

Association between flowering time and bioclimatic variables.

To test whether flowering time correlates with environmental variation, we extracted the 19 classical worldclim bioclimatic variables (Bio1 to Bio19) as well as solar radiation in spring, global aridity index in spring and altitude for each locality and performed a PCA with the resulting 22 variables. The first two axes of the PCA explained together about 58% of the variation among our samples, indicating that our set of selected accessions occur in different environmental conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

We ran linear mixed effect models (LMM) where flowering time was entered as the response variable, the 22 environmental variables mentioned above as well as the first two PCA axes entered separately as fixed factors and the clade of origin as random factors to account for population structure. We found that environmental variables linked to warm months were all significantly associated with flowering time variation while environmental variables linked to cold month showed no association (Table 1).

Table 1: Output of the LMM between flowering time and environmental variables

Bioclim	Pval	Fixed_effect	Random_effect	residuals	Name_bioclim
bio17	0.0002	0.117	0.829	0.055	Precipitation of Driest Quarter
srad spring	0.0005	0.117	0.826	0.057	srad spring (mean)
PC1 + PC2	0.004	0.088	0.856	0.056	PC1 + PC2
bio18	0.0007	0.081	0.864	0.055	Precipitation of Warmest Quarter
bio12	0.001	0.079	0.864	0.057	Annual Precipitation
bio14	0.001	0.081	0.861	0.057	Precipitation of Driest Month
aridity spring	0.0017	0.095	0.846	0.059	aridity spring mean)
bio13	0.002	0.066	0.879	0.055	Precipitation of Wettest Month
bio16	0.006	0.057	0.888	0.056	Precipitation of Wettest Quarter
bio7	0.01	0.096	0.849	0.055	Temperature Annual Range (BIO5-BIO6)
bio5	0.04	0.042	0.902	0.055	Max Temperature of Warmest Month
bio2	0.04	0.06	0.886	0.055	Mean Diurnal Range
bio9	0.07	0.023	0.925	0.052	Mean Temperature of Driest Quarter
bio8	0.13	0.019	0.928	0.052	Mean Temperature of Wettest Quarter
bio19	0.13	0.017	0.928	0.054	Precipitation of Coldest Quarter
elevation	0.14	0.016	0.93	0.054	elevation
bio10	0.17	0.016	0.93	0.054	Mean Temperature of Warmest Quarter
bio1	0.19	0.013	0.934	0.053	Annual Mean Temperature
bio4	0.19	0.029	0.917	0.054	Temperature Seasonality
bio15	0.31	0.011	0.935	0.053	Precipitation Seasonality
bio11	0.67	0.002	0.946	0.052	Mean Temperature of Coldest Quarter
bio6	0.68	0.001	0.946	0.053	Min Temperature of Coldest Month
bio3	1.00	0	0.948	0.052	Isothermality

Fixed, random and residual effects indicate the part of the variance explained by these factors.

Precipitation in the driest quarter (Bio17) or solar radiation in spring explained the largest part of the variance within clades, and accessions from more arid environments tend to flower faster (Table 1, Figure 1C and D). Kendall rank correlation between flowering time and Bio17 or solar radiation in spring show significant (P-value

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

= 5.804e-10 and 2.391e-08) but only partial correlations (r = 0.34 and -0.31 respectively), implying that while bioclimatic variables like Bio17 and solar radiation (srad) in spring are driving within clade flowering time variation (Figure 1C and D), they only partly explain the early/late-flowering partitioning of the phenotypes we observe among genetic lineages (Figure 1B). This is in line with the fact that even when significant, environmental variables explain a relatively small part of the variance in flowering time (maximum 11.7% for Bio17 and srad in spring) compared to the clade of origin (Table 1).

However, the *B. distachyon* genetic clades are occurring in different ecological niches (Minadakis et al. 2023). Hence, by correcting for population structure, we may also correct for the confounding effect of the environment. We indeed found significant and larger effects of the environment on flowering time when not accounting for population structure and running classical linear model analyses (max R^2 = 0.22, Table S3). Regardless of the correction method, these results indicate that flowering time, when measured under greenhouse conditions, is significantly but only partially shaped by the current environmental conditions we tested in *B. distachyon*.

Contribution of flowering time genes to flowering time variation

The large proportion of the variance explained by the genetic clade (Table 1) indicates that genetic factors might be playing a larger role in flowering time variation than the environment. The genetic basis of flowering time has been extensively characterized at the molecular level in *B. distachyon* (e.g. Raissig and Woods, 2022; Woods et al., 2023). Due to our relatively small sample size (56 accessions with phenotypes), we opted for a targeted approach rather than a classical GWAs and selected 22 flowering-time genes (Table S4) molecularly characterized and described as impacting flowering time in our study system (Higgins et al. 2010; Wu et al. 2013; Woods et al. 2014; Sharma

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

et al. 2017; Woods et al. 2017[a,b]; Lomax et al. 2018; Qin et al. 2019; Woods et al. 2019; Cao et al. 2020; Kennedy and Geuten 2020; Woods et al. 2020).

Using the SNP calling performed by Minadakis et al. (2023), we extracted 695 SNPs across these 22 flowering-time genes and 56 accessions. We then ran LMM where flowering time was used as the response variable, SNPs at flowering-time genes as fixed factors and the genetic clade of origin as random factors. We found SNPs significantly associated with flowering-time for ten (hereafter AFT-genes) out of the 22 flowering-time genes (Table S5). Due to strong linkage disequilibrium (LD) over short distances (Figure S2), many SNPs gave precisely the same signal of association with flowering time within a given gene (Table S5). For each AFT-gene, we thus only kept one SNP with the largest effect on flowering time as a focal SNP for further analyses. The resulting ten SNPs individually explained as much as 47% of the phenotypic variation (Table 2, Figure 2A) while their additive effect explained 58% of the variance (Table 2). *VRN1*, *FTL10*, *POLD3*, and, *FTL1* showed the strongest association with flowering time. We did not detect epistatic effects most likely because genes explaining a large part of the variance in flowering time are almost fixed within genetic clades and therefore display identical or near identical allele frequencies.

Table 2: Output of the linear mixed model analyses between flowering time and SNP at AFT genes

CHR	BP	Pval	Fixed effect	Random effect	residual effect	gene
Bd1	5'866'489	0.0004	0.47	0.465	0.066	VRN1
Bd2	17'324'316	0.0006	0.463	0.476	0.061	FTL10
Bd2	4'928'157	0.0006	0.463	0.476	0.061	POLD3
Bd2	5'415'945	0.0006	0.438	0.498	0.064	FTL1
Bd1	41'631'615	0.001	0.367	0.565	0.067	CO1
Bd1	5'923'377	0.02	0.342	0.605	0.053	PHYC
Bd3	8'185'926	0.001	0.218	0.713	0.069	VRN2
Bd3	55'978'083	0.04	0.094	0.837	0.069	CO2
Bd2	39'854'115	0.01	0.075	0.876	0.049	RVR1
Bd2	12'881'238	0.04	0.037	0.911	0.052	ELF3
Additive model		<0.001	0.58	0.353	0.053	All

For each gene, we only display the SNP with the strongest association with flowering time. Fixed, Random and residual effect indicate the part of the variance explained by these factors.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Differentiation of flowering time genes

The fact that the AFT-genes gave identical (e.g. *FTL10* and *POLD3*) or very similar signals with regard to their association with flowering time (Table 2) suggests, as displayed in Figure 2A, that SNPs at AFT-genes harbor similar allele frequencies. This could result from two processes. On the one hand, *B. distachyon* populations underwent bottlenecks during the last glaciation (Minadakis et al. 2023), which may have led to reduced genetic diversity and highly differentiated alleles among genetic lineages/clades genome-wide. On the other hand, AFT-genes might have been co-selected (polygenic selection) and remain at similar allele frequencies (Zan and Carlborg 2019; Gupta et al. 2023).

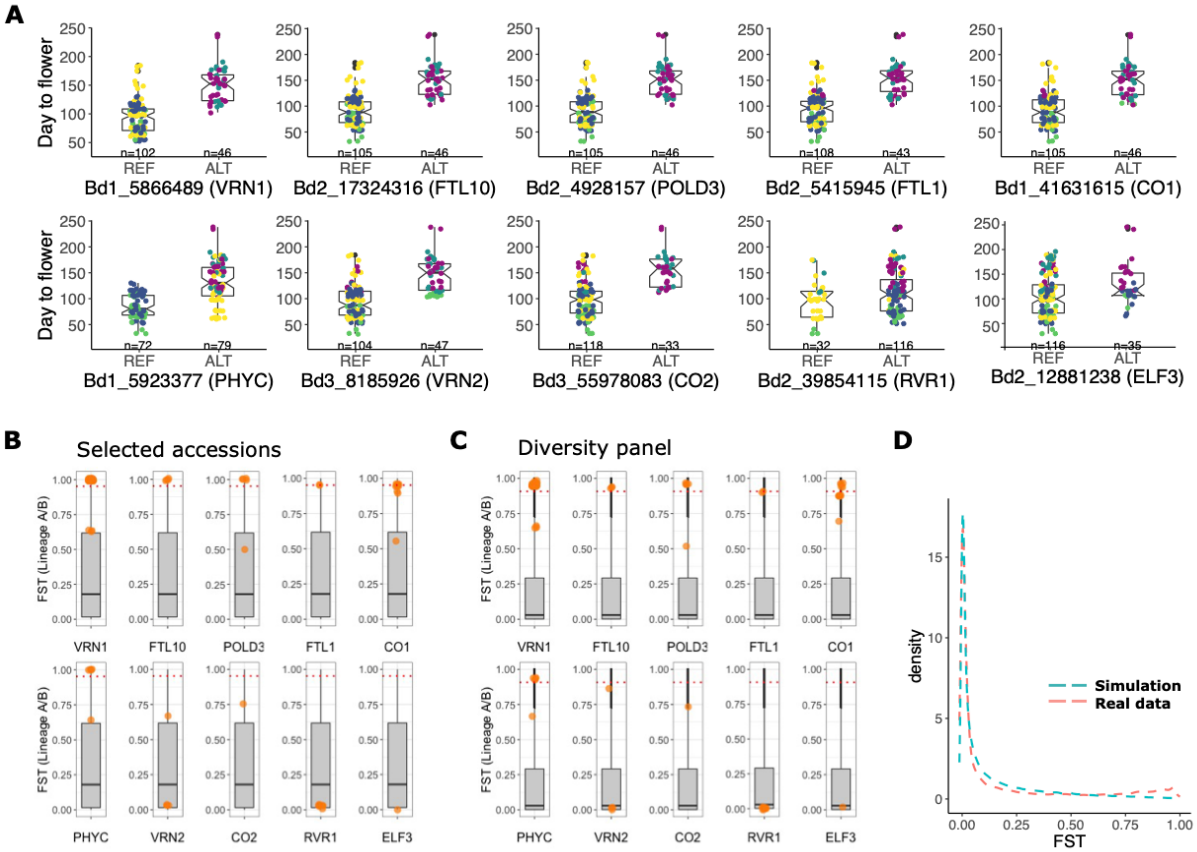

Fig. 2: Association between the ten AFT-genes and flowering time. A) For each gene, the boxplots contrast the effect of the reference (Bd21) and alternative alleles on flowering time. The color code is the same as in Figures 1 B) Single SNP FST computed for the 56 accessions of the A and B lineages. C) Single SNP FST computed for the 332 accessions of the A and B lineages. For the B and C panels, orange dots display FST for SNPs significantly associated with flowering time variation. Boxes display genome-wide distribution. The dashed line indicates the 95 percentiles. D) Distribution of FST calculated between the A and B lineage with real and forward-simulated data under a neutral scenario

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

To assess the extent to which our SNPs of interest are more differentiated compared to genome-wide levels, we computed single SNP F_{ST}. To limit the number of comparison and considering the partitioning of the reference and alternative alleles among accessions in the ten AFT-genes (Figure 2A), we only computed F_{ST} between accession of the A and B lineages. We found that the large majority of SNPs significantly associated with flowering time are above the 3rd quartile of the genome-wide distribution (Figure 2B). Significantly associated SNPs in *VERNALIZATION 1* (*VRN1*), *FTL10*, *POLD3*, *FTL1*, *CONSTANS* (*CO1*), and *PHYTOCHROME C* (*PHYC*) belong to the top 5% outliers. Hence, with the exception of SNPs in *REPRESSOR OF VERNALIZATION1* (*RVR1*) and *EARLY FLOWERING 3* (*ELF3*), which only contribute marginally to flowering time variation (Table 2), SNPs significantly associated with flowering time tend to be more differentiated than the rest of the genome.

To gain power for our analyses, we further made use of the *B. distachyon* diversity panel composed of 332 natural accessions (Minadakis et al. 2023, Figure 3A) and found very similar levels of allele differentiation at AFT-genes (Figure 2C). These results are in line with a forward simulation we performed under a neutral scenario using the demographic estimates computed by Minadakis et al. (2023). The F_{ST} distribution calculated between the simulated A and B lineages, as the distribution obtained with the real data, is indeed largely shifted towards low values. These latter results demonstrate that bottlenecks did not lead to elevated genetic differentiation at the genome-wide level (Figure 2D).

Long-range linkage disequilibrium among flowering genes

The similar allele frequencies described above also indicate that AFT-genes are in strong linkage disequilibrium despite the large physical distances that separate them. LD computed among pairs of focal SNPs at AFT-genes in the 56 selected accessions as well as in the entire diversity panel show a similar pattern (Figure 3B): with the exception of pairs of AFT-genes involving *ELF3* and *RVR1*, we found that AFT-genes

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

located on the same chromosome are in much higher LD than expected considering the 95% confidence interval (CI) of the genome-wide LD-decay (Figure 3C).

Strikingly, we also found very high LD between AFT-genes located on different chromosomes both in the selected accessions and entire diversity panel (Figure 3B). Located on chromosome 1, *VRN1* for instance, is in strong LD with five other genes located on chromosome 2 (*POLD3*, LD = 0.84; *FTL1*, LD = 0.76; *FTL10*, LD = 0.83) and chromosome 3 (*VRN2*, LD = 0.65; *CO2*, LD = 0.55). To test to what extent this long-range LD deviates from genome-wide patterns, we computed LD for a subset of 100,000 random pairs of genic loci located on different chromosomes across the genome. This allowed us to estimate the distribution and CI of inter-chromosomal LD between two loci as well as the average inter-chromosomal LD among five and six loci (Figure 3D). For each combination, the inter-chromosomal LD we observed between pairs of focal AFT-genes (two loci) or between five (for AFT-genes located on

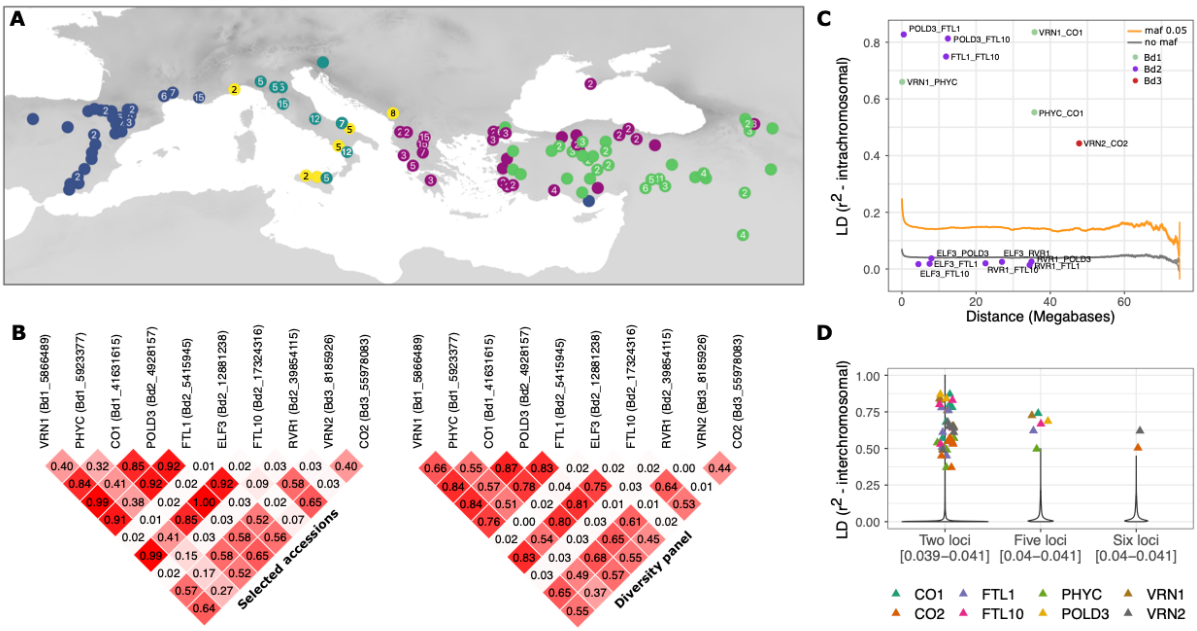

Fig. 3: Linkage disequilibrium (LD) in the flowering time pathway. A) Map displaying the locality of the 332 sequenced accessions composing the diversity panel B) LD among the 10 AFT-genes associated with flowering time C) LD decay computed over the entire diversity panel. The orange and grey lines display the 95% CI of the LD decay calculated with or without filtering for a minimum allele frequency (maf) respectively. Dots display LD between pairs of flowering-time genes located on the same chromosome D) Distribution of LD between loci located on different chromosomes. The violin plots display the means of LD calculated between two, five or six loci. Numbers in square brackets indicate the confidence interval around the mean

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

chromosomes 1 and 2) and six AFT-genes (for AFT-genes located on chromosome 3) was largely outside the CI calculated with random genic loci (Figure 3D).

Note that only two flowering-time genes but non AFT-genes, *FUL2* and *PPD1*, show putative footprint of insertion polymorphisms in the diversity panel (Table S4). Furthermore, in a selfing species like *B. distachyon*, heterozygous SNPs are hallmarks of gene duplication (Stritt et al. 2022; Jaegle et al. 2023). Using a vcf file not filtered for heterozygous SNPs, we found that only 0.6% of the SNPs located in flowering-time genes are heterozygous across the 332 accessions. These results imply that flowering-time genes and specifically AFT-genes are occurring in single copy and that structural rearrangements or duplications did not alter their position in non-reference genomes.

Taken together, these results indicate that the long-range LD among AFT-genes as well as their differentiation levels deviate from genome-wide patterns and are difficult to explain by the demographic history of the population alone. We hence conclude that *VRN1*, *PHYC*, *CO1*, *POLD3*, *FTL1*, *FTL10*, *VRN2*, and *CO2* are co-evolving and undergo polygenic selection.

Genome-wide scans of positive selection

To properly test for positive selection while accounting for the structure and demographic history of our populations, we computed X^TX statistics, a measure comparable to single SNP F_{ST} that accounts for the neutral covariance structure across populations. In brief, we computed X^TX with our actual SNP dataset over the entire diversity panel using the five genetic clades as focal populations. We then simulated a pseudo-observed dataset (POD) of 100,000 SNPs under the demographic model inferred from the covariance matrix of the actual SNP dataset. X^TX statistics were then computed for the POD to determine the probability of neutrality for each SNP. The threshold of significance was thus set to 11.2, a value slightly lower than the 1% outlier threshold (13.6). With the exception of *ELF3*, all AFT-genes display SNPs more differentiated than expected under a neutral scenario (Figure 4A) suggesting that AFT-genes have evolved under positive selection. We also tested for extended haplotypes

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

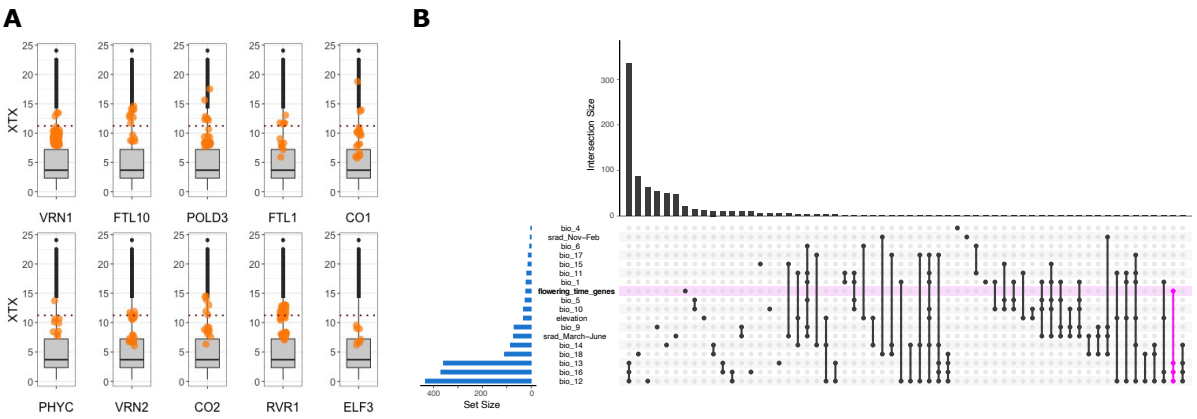

Fig. 4: Positive selection at AFT-genes and association with bioclimatic variables A) XTX analyses performed in the entire diversity panel using the five genetic clades as focal populations. The boxplots display genome-wide XTX levels. The orange dots display XTX values for any SNP located in a given AFT-gene. The dash-line indicate the threshold of significance inferred from the POD. B) Upset plot displaying the overlap between gene sets identified with the GEAs performed by Minadakis et al. (2023) for 23 bioclimatic variables and flowering-time genes.

and footprint of selective sweeps with the integrated Haplotype Score (iHS). However, apart from *CO2* in the A lineage, none of the AFT genes are located in regions displaying significantly longer haplotypes (Figure S3).

We ran SnpEff to test which SNPs in flowering-time genes are more likely to have a functional impact. Only two SNPs in *CO2* were categorized as variants with high impact while the large majority of SNPs in other AFT genes were categorized as variants with moderate and low effect (Figure S4), rendering the identification of the potential targets of selection challenging for most of our genes of interest.

In order to characterize which environmental factors might have shaped diversity at AFT-genes, we eventually made use of genotype-environment association (GEA) analyses performed by Minadakis et al. (2023) with bioclimatic variables associated to precipitation levels, temperature or elevation. We found that only one of the 22 flowering-time genes (Bradi2g59119, *ODDSOC2*-like) showed an overlap with the gene sets significantly associated with current bioclimatic variables (Figure 4B).

Here again, the GEAs performed by Minadakis et al. corrected for population structure and can hence result in false negative since the genetic clades occupy different ecological niches. We indeed found significant associations between SNPs at AFT-genes (with the exception of *PHYC*) and bioclimatic variables such as Bio14

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

(Precipitation of driest quarter) and aridity levels in spring, when not accounting for population structure (Kendall correlation; Figure S5A), indicating that the confounding effect of population structure and adaptation at a regional scale may mask the significant effect of the environment on AFT-genes diversity.

Potential adaptation to past climatic conditions

We estimated the age of AFT alleles and found that those arose relatively recently between 9,000 to 38,000 years ago (Table S6) as most alleles in our system (Minadakis et al. 2023). We speculated that variation in flowering time could also reflect adaptation to recent past conditions, potentially to the last glacial maximum (LGM) 22,000 years ago. We therefore tested whether the delay in flowering time we observed in accessions from the A_East and A_Italia clades could result from an adaptation to past colder climates. To do so, we used the niche suitability projections under LGM conditions computed by Minadakis et al (2023). We selected a set of 200 random points per clade in highly suitable habitats and extracted the 19 LGM bioclimatic variables for the corresponding sites. The PCA performed with these 19 LGM bioclimatic variables do not allow us to separate the five genetic clades. Altogether, accessions from the A_East and A_Italia clades neither occurred in colder nor in wetter environments than the B_East, B_West and C accessions (Figure S6) which suggests that the extended vernalization requirement and delay in flowering we observed for the A_East and A_Italia accessions were not selected during LGM.

Genome-wide association for flowering time variation measured in outdoor conditions

Greenhouse conditions are far from any natural optimum and may obscure association with relevant environmental cues, while taking an environment closer to the natural conditions encountered by of one or several clade acts as a useful pivot to contrast clades. Hence, we eventually made use of flowering time data we collected for a subset of 131 accessions (Stritt et al. 2022) in Zürich, Switzerland, to assess flowering time

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

variation in semi-natural conditions. While *B. distachyon* do not occur at such latitudes (Minadakis et al. 2023), a PCA performed with the 22 bioclimatic variables (see above) on our 332 natural accessions and including a site in Zurich shows that the environmental conditions in Zürich are not different from the ones encountered by ABR9 for instance (Figure 5A). While other biotic (vegetation type and density) and abiotic (soil characteristic) factors are certainly preventing the species from occurring in northern latitude, Zürich harbours climatic conditions similar to ones encountered by the species at the northern margin of its natural distribution, making it a valid experimental site to study flowering time in *B. distachyon*.

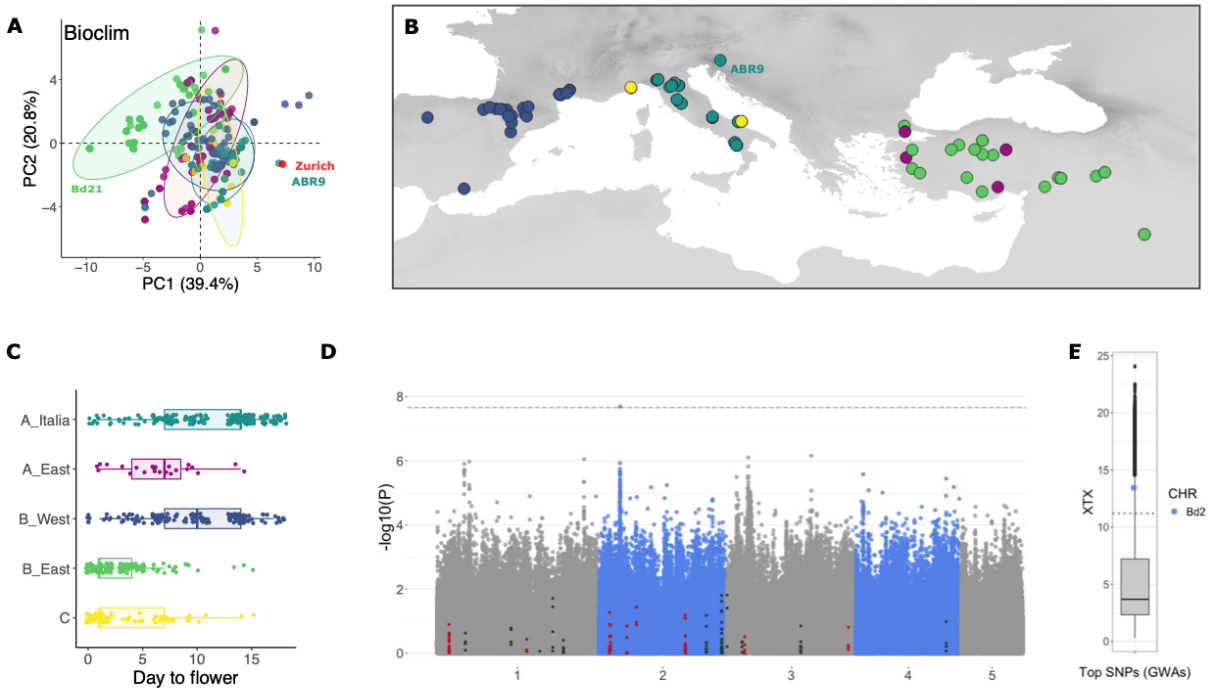

Fig. 5: Genome-wide association analysis with flowering time (outdoor experiment). A) PCA performed with 19 classical wordlclim variables combined to solar radiation in spring, global aridity in spring and altitude for the 332 accessions of the diversity panel and a locality in Zurich B) Geographical origin of the 131 accessions used for the experiment C) Flowering time per cluster (displaying replicates) D) Manhattan plot displaying the association between SNPs and flowering time. The red and dark dots display SNPs located into AFT- or other flowering-time genes respectively. The dashed line corresponds to FDR threshold of significance E) XTX values for the top GWAs SNP on chromosome 2 (Bd2).

In brief, we planted seeds for 131 accessions (Figure 5B) outdoors in November 2017, in the Botanical Garden of Zurich, Switzerland, with six replicates per accession and recorded flowering time in spring. All plants flowered within 20 days in April (Figure 5C). Flowering time for plants from the B_East and C clades were not significantly different (Kruskal-Wallis test, P-value = 0.96) but those accessions

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

flowered significantly faster than accessions from the other clades (Kruskal-Wallis test, all P-values < 0.01). Plants from the A_Italia and B_West flowered significantly later than plants from the other clades (Kruskal-Wallis test, all P-values < 0.02) but were not significantly different from each other (P-value = 0.06). As such, the data collected outdoor contrasts with the ones collected in the greenhouse. As for the experiment in the greenhouse, linear-mixed model analyses showed that Bio17 (Precipitation of driest quarter) is the bioclimatic variable explaining the largest part of flowering time variation. In this case, however, it explains as much as 21% of the variance, and hence substantially more than with the data collected under greenhouse conditions (Table 3). These results are in line with those obtained with a simple linear model, where Bio17 explains more than 44% of the variance (Table S7).

Table 3: Output of the LMM analyses between flowering time and environmental variables (outdoor conditions)

Bioclim	Pval	Fixed effect	Random effect	residuals	Name_bioclim
bio17	<0.00001	0.21	0.191	0.599	Precipitation of Driest Quarter
bio12	<0.00001	0.147	0.297	0.556	Annual Precipitation
bio18	<0.00001	0.115	0.272	0.613	Precipitation of Warmest Quarter
bio13	<0.00001	0.097	0.379	0.524	Precipitation of Wettest Month
bio14	<0.00001	0.094	0.292	0.614	Precipitation of Driest Month
bio16	<0.00001	0.089	0.382	0.529	Precipitation of Wettest Quarter
bio3	<0.00001	0.081	0.409	0.51	Isothermality
bio19	<0.00001	0.059	0.437	0.504	Precipitation of Coldest Quarter
srad spring	0.00012	0.058	0.369	0.573	srad spring mean
bio8	<0.00001	0.045	0.521	0.435	Mean Temperature of Wettest Quarter
bio4	0.00021	0.036	0.51	0.454	Temperature Seasonality
aridity spring	1.00E-05	0.034	0.402	0.564	aridity spring mean
bio1	<0.00001	0.025	0.436	0.54	Annual Mean Temperature
bio11	<0.00001	0.022	0.469	0.509	Mean Temperature of Coldest Quarter
bio2	0.008	0.017	0.472	0.511	Mean Diurnal Range
bio6	0.0001	0.016	0.476	0.508	Min Temperature of Coldest Month
bio10	0.005	0.01	0.44	0.55	Mean Temperature of Warmest Quarter
bio5	0.01	0.009	0.441	0.55	Max Temperature of Warmest Month
bio7	0.45	0.002	0.468	0.53	Temperature Annual Range
bio15	0.9	0	0.462	0.538	Precipitation Seasonality
bio9	0.9	0	0.456	0.544	Mean Temperature of Driest Quarter

Fixed, Random and residual effect indicate the part of the variance explained by these factors.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

A genome-wide association analysis (GWAs) performed with GEMMA (Zhou and Stephens 2012) on the 131 accessions identified one significant peak which does not overlap with any of the 22 flowering-time genes we studied above (Figure 5D). We also made use of the GEAs performed by Minadakis et al. (2023) and found that the gene underlying the peak (Bradi2g11490, a carbohydrate-binding-like gene) does not overlap with the gene sets significantly associated with current bioclimatic variables. Yet, similarly to what we found for the AFT-genes, we found significant associations (Figure S5B) between SNPs at the GWAs peak and bioclimatic variables, especially with Bio14 (Precipitation_of_Driest_Month), when not taking into account population structure. Consistent with selection by the environment, the top GWAs SNP constitutes also an X^TX outlier (Figure 5E).

Discussion

Flowering time and adaptation to local climate

Flowering time has been shown to play an important role in local adaptation across many plant systems including crops (Izawa 2007; Anderson et al. 2012; Anderson, et al. 2013; Ågren et al. 2017; Navarro et al. 2017; Wadgymar et al. 2017; Takou et al. 2019; Qian et al. 2020; Yan et al. 2021). Here, we show a significant association between flowering time and bioclimatic variables, especially precipitation levels in warm months (e.g. Bio17), both under greenhouse and outdoor conditions. The part of the variance explained by bioclimatic variables was yet much higher with the outdoor (21%) than with the greenhouse experiment (11.7%). Yet, because the five *B. distachyon* genetic clades occupy different ecological niches, accounting for population structure might also lead to an underestimation of the environmental effect. In fact, the part of the variance in flowering time explained by the environment is much larger when not accounting for population structure (22% and 40% respectively). While estimating the effect of the environment remains a challenge in such structured populations, these results show that flowering time variation is a locally adapted trait in our system as

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

well. Interestingly, precipitation levels and hence water limitation seem especially important for flowering time as already shown for broader fitness effects in our species (Des Marais et al. 2017).

The greenhouse experiment further confirms that *B. distachyon* natural accessions vary greatly regarding their vernalization requirement and ultimately flowering time and can be broadly grouped into early- and late-flowering genotypes (Figure S1 and Figure 1B, but see Ream et al., 2014; Gordon, et al. 2017 for a finer classification scheme). Despite within clade variation regarding flowering time (Figure 1B), accessions from the A lineage (A_East and A_Italia) display a large delay in flowering compared to accessions from the B (B_East and B_West) and C lineage, which typically result from longer vernalization requirements. This pattern has been suggested to be a sign of adaptation at a regional scale in Turkey (Skalska et al. 2020) and could be interpreted as a footprint of diversifying selection at the genetic clade level. Based on this hypothesis, accessions from the A_East and A_Italia clades may display a delay in flowering as globally adapted to colder and less arid environmental conditions, as shown for instance in *A. thaliana* (Ågren and Schemske 2012). Yet, our results do not fully support this scenario, as we found a large discrepancy between the partitioning of the phenotypes and the bioclimatic variables associated with flowering time within the genetic clades (Figure 1B, C, D). Our niche modelling projections for the last glacial maximum (Minadakis et al. 2023), do not further support an adaptation to the last glacial maximum conditions (Figure S6). As such, the partitioning of the early vs. late flowering accessions remains difficult to explain with an "adaptive" scenario. In summary, we conclude that local environmental conditions are partly driving flowering time variation within genetic clades but not the early- and late-flowering partitioning of the phenotypes we observed among clades in the greenhouse.

As already shown by Stritt et al. (2022) with a smaller subset of accessions, variation in flowering time is largely attenuated in outdoor conditions and although significant differences are observed among genetic clades, all plants flowered within 20 days

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

when grown outdoors in Zürich (Figure 5C). This suggests that vernalization is only one important known environmental cue that shapes a given accessions flowering behavior. However, in the wild many other additional environmental cues, some of which are not currently well understood in *B. distachyon*, may also play an important role in impacting the flowering behavior of a given accession. Hence, such experiments, largely performed in *B. distachyon* (Ream et al. 2014; Gordon, et al. 2017; Sharma, et al. 2017) may only weakly capture how accessions behave in the wild and most likely reflect factors necessary for the plant to acquire the competence to flower rather than "true" flowering time in the wild.

Flowering-time genes and flowering time variation

Flowering-time genes have been extensively characterized at the molecular level in *B. distachyon* (Ream et al. 2012; Ruelens et al. 2013; Ream et al. 2014; Woods et al. 2014; Woods et al. 2017[a,b]; Lomax et al. 2018; Cao et al. 2020; Kennedy and Geuten 2020; Woods et al. 2020; Bouché et al. 2022; Raissig and Woods 2022). Here we selected 22 flowering-time genes known to play a major role in vernalization- (e.g., *VRN1*, *RVR1*) or photoperiod-sensing (e.g. *PPD1*, *PHYC*, *ELF3, VRN2*) to characterize the magnitude of their allelic effects on flowering time. We found *VRN1*, *FTL10*, *POLD3*, *FTL1*, *CO1*, *PHYC*, *VRN2*, *CO2*, *RVR1*, and *ELF3* (AFT-genes) to be significantly associated with flowering time measured in the greenhouse, many of which underlie significant QTL peaks in biparental populations or significant SNPs in genes from previous flowering GWAS (Bettgenhaeuser et al. 2017; Woods et al. 2017[a]; Wilson et al. 2019). SNPs in *VRN1*, *FTL10*, *POLD3*, and *FTL1* explained individually more than 40% of the variance in flowering time while *RVR1*, *PPD1* and *ELF3* were only marginally associated with phenotypic variation. The additive model explained altogether 58% of the variance, demonstrating that genetic factors play a larger role in shaping flowering time under greenhouse conditions than the climate at the locality of origin.

The current molecular model of flowering in *B. distachyon* involves all of the AFT's found in this study. Briefly, prior to vernalization expression of a long day photoperiod

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

repressor, *VRN2* prevents flowering by repressing the expression of the long day mobile floral signal *FTL1* from being expressed (Woods et al. 2017[a]). After vernalization, the basal repressive signal from *VRN2* is overcome by the strong flowering inductive signal provided by *VRN1,* which is induced during vernalization (Raissig and Woods 2022). *VRN1* remains repressed prior to vernalization due in part by *RVR1*, which is important in establishing a vernalization requirement in *B. distachyon* by repressing *VRN1* prior to cold (Woods et al. 2017[b]). *EZL1* and *POLD3* are also involved in the repression of *VRN1* but also impact the expression of many other floral homeotic genes and genes with diverse functions (Lomax et al. 2018; Woods et al. 2020). In addition to a vernalization requirement that must be met by cold exposure in *B. distachyon* the exposure of plants to a prolonged period in short days can also provide competence to flower referred to as short day vernalization, which requires the expression of *FTL9* (a close paralog to *FTL10*) first in SD to allow flowering in long days (Woods et al. 2019). Ultimately, long days are essential for the expression of *FTL1* and thus flowering in *B. distachyon* (Wu et al. 2013; Ream et al. 2014). *PHYC* is essential for long day photoperiodic flowering and is required for the long day activation of a number of genes in the photoperiod pathway including *PPD1*, *CO1*, and *CO2* that in turn are important in the activation of *FTL1* (Woods et al. 2014; Qin et al. 2019). The translation of the light signal perceived by PHYC into a flowering response is mediated by *ELF3*, which itself directly regulates the expression of *PPD1* in LD (Bouché et al. 2022; Alvarez et al. 2023; Woods et al. 2023). Taken together, several of the genes shown to play an important role in flowering via molecular, biochemical, and genetic methods also clearly contribute to flowering variation in diverse *B. distachyon* accessions.

However, it is interesting that genes like *ELF3* and *PPD1* play a minor (3.7%) or are not associated at all in shaping the flowering time diversity of *B. distachyon* whereas variation in these two genes is important in wheat and barley (Turner et al., 2005; Wilhelm et al., 2009; Bendix et al., 2015). This may reflect the different evolutionary histories of a wild grass, such as *B. distachyon* versus domesticated crops.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Moreover, the polyploid nature of crops like wheat have frequently found that variation in flowering often involves either copy number variation of specific genes or dominant mutations in cis-regulatory regions that increase gene expression and were selected unknowingly by breeders (e.g. (Wilhelm et al. 2009; Würschum et al. 2015). Interestingly, while *pold3* and *elf3* loss of function mutants flower faster than the wildtype (Woods et al. 2020; Bouché et al. 2022), their contribution to flowering time variation (46% and 3.7% respectively) differ drastically. This discrepancy might be due to the fact that mutants used to characterize gene functions usually harbor mutations with deleterious or large effect size (loss of function). In contrast, our natural flowering-time gene variants are mostly predicted as having low to moderate impacts.

In fact, none of the AFT-genes co-localize with the GWAs candidate identified with the outdoor experiment (Figure 5D), questioning the role of flowering time genes on the fine-tuning of flowering time in the wild. While we should extent our common garden experiment to a larger number of sites, our results are in line with a previous studies in *A. thaliana* which, by using natural accessions and RILs, found that flowering time variation scored in the field experiment poorly correlated with the flowering time variation obtained under greenhouse conditions (Weinig et al. 2002; Malmberg et al. 2005; Brachi et al. 2010; Wilczek et al. 2010). As a consequence, a limited overlap is observed between the genomic regions detected in field experiments and those detected under greenhouse conditions (Brachi et al. 2010). Flowering time has a complex polygenic architecture (Buckler 2009; Brachi et al. 2010; Navarro et al. 2017; Zan and Carlborg 2019; Gaudinier and Blackman 2020) and the use of EMS-induced mutants suggested that many additional genes might play a role in shaping this trait in *B. distachyon* (Raissig and Woods 2022), some of which are not described or only play a minor role in flowering time in other plant models (Woods et al. 2020). In addition, alleles may affect phenotypes only in specific populations (Zan and Carlborg 2019; Yan et al. 2021; Gloss et al. 2022) or seasons (Weinig et al. 2002; Gould and Stinchcombe 2017). Taking into account the polygenic architecture, gene-by-environment association and phenotypic plasticity (Gaudinier and Blackman 2020; Yan

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

et al. 2021) will therefore be essential to better capture the adaptive potential of flowering time and flowering-time genes in natural populations of *B. distachyon.* Although flowering-time genes are undoubtedly essential in the perception of environmental cues and overwintering (for review Raissig and Woods 2022), our results show that the effect of their variants in the wild are more difficult to predict and remain to be carefully quantified.

Flowering-time genes and polygenic selection

The greenhouse experiment further highlighted a striking pattern of long-range linkage disequilibrium (LD), both intra- and inter-chromosomal, among eight AFT-genes associated with flowering time variation under greenhouse conditions (Figure 3). Polygenic selection is expected to result in LD between regions under selection (Yeaman et al. 2016; Yeaman et al. 2018; Gupta et al. 2023) but only few cases have been reported so far in plants and animals (Hohenlohe et al. 2012; Yeaman et al. 2016; Park 2019; Gupta et al. 2023). In *A. thaliana,* Zan and Carlborg (2019) also identified long-range LD among four clusters of flowering-time genes.

Disentangling the effect of population structure and demographic effects from polygenic selection can be challenging and there needs to be caution interpreting these patterns (Zan and Carlborg 2019; Gupta et al. 2023). The long-range LD we observed among our eight AFT-genes is associated with high genetic differentiation between the A and B lineages (Figure 2A, B, C) and could result from the recent bottlenecks experienced by *B. distachyon* in the recent past (Stritt et al. 2018; Minadakis et al. 2023). Three lines of evidence rule out this demographic scenario. First, FST analyses performed with our real- as well as forward-simulated data show that the population size reduction experienced by *B. distachyon* did not lead to genome-wide highly differentiated alleles and AFT-gene variants constitute clear FST outliers. Second, we also demonstrated that the levels of long-range LD (intra- and inter-chromosomal) we observed among these eight AFT-genes (Figure 3) are not observed among random genic loci in the genome. Eventually, AFT-genes do not display signs

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

of duplication or insertion polymorphisms which imply that structural rearrangements did not bring AFT-genes physically together in non-reference accessions. We therefore believe that we present clear evidence for polygenic selection on key genes involved in the flowering pathway.

Expression analyses further support the functional connection among these loci in *B. distachyon* and grasses in general. For example, *VRN1* and *FTL1* are expressed in a positive feedback loop, which overcomes the flowering repression of *VRN2* (Ream et al. 2014; Woods et al. 2016; Woods et al. 2017[b]). Additionally, there is an intricate connection between many of these flowering genes both transcriptionally and at the protein level. For example, among the pairwise interactions tested between *PHYC*, *PHYB*, *ELF3*, *PPD1*, *VRN2*, *CO1*, and *CO2* more than 80% showed positive interactions in yeast two hybrid assays some of which have been verified *in planta* (Shaw et al. 2020; Alvarez et al. 2023). Thus, many of these genes can interact at multiple levels.

The X^TX analysis we performed with the five genetic clades as focal populations shows that AFT-genes harbor SNPs above the threshold of significance, *POLD3* and *CO1* presenting more extreme outliers. This footprint of positive selection (Gautier 2015) is yet not accompanied by extended haplotypes around our candidate genes (except for *CO2*) which implies that the initial selective sweeps may have eroded with time and that selection did not occur in a recent past. AFT-genes and flowering-time genes are yet not colocalizing with regions we previously identified with GEA analyses (Figure 4B). Thus, although we find evidence of positive selection at single AFT-genes in additional of signs of polygenic selection, the selective constraint at play remains yet to be identified. We eventually found evidence for positive selection but no association with GEA regions for the newly identified GWAs candidate (Figure 5E).

It is yet essential to keep in mind that GEA performed in species with strong population structure can lead to high rate of false negatives as they typically include structure correction (Booker et al. 2023; Lottheros 2023). We did find significant associations between bioclimatic variables and SNPs at AFT-genes and GWAs peak

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

when not correcting for population structure (Figure S5). The fact AFT-genes display signs of positive selection argue in favor of an effect of the environment in shaping their diversity but also prove that disentangling the confounding effects of population structure and local adaptation remain challenging. Common garden experiments will thus be key to test the effect of genotype-by-environment interactions in our system.

Conclusion and perspectives

Our results suggest that i) flowering time is a locally adapted trait in *B. distachyon* but ii) only part of the variation in flowering time can be explained by the environment and iii) eight key flowering-time genes are co-evolving but their effect in the wild remain to be clarified. In the face of global warming, plant phenology has recently advanced significantly (e.g. Anderson et al. 2012) therefore, investigating the polygenetic architecture of flowering time remains therefore a timely question. Polygenic selection, epistatic and pleiotropic effects might limit the evolution of traits (Yan et al. 2021; Yeaman 2022) and future experiments in *B. distachyon* should focus on disentangling these effects. Mimicking the combination of the various clues that trigger flowering is yet impractical under greenhouse conditions and common garden experiments will thus be essential to place flowering time in a natural context in this system. We, however, know relatively little about the basic ecology of *B. distachyon*. When are plants emerging in the wild? How plastic is flowering time and phenotypes in general? What is the contribution of seed banks to the effective population size and hence selection strength? constitute fundamental questions to address in the future to more broadly characterize the process of local adaptation in this species.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Materials and Methods

Biological materials and genomic resources

The *Brachypodium distachyon* diversity panel is composed of 332 natural accessions for which whole-genome sequencing data are publicly available (Gordon, Contreras-moreira, et al. 2017; Gordon et al. 2020; Skalska et al. 2020; Stritt et al. 2022). For the flowering time experiment, we selected a subset of 61 accessions representing all five genetic clades as described by Stritt et al. 2022 and occurring along latitudinal gradiants. Maps were drawn QGIS (version 3.16).

We also made use of the raw vcf produced by Minadakis et al. (2023) for the entire diversity panel. We used vcftools (Danecek et al. 2011) to apply the following filtering criteria: --max-alleles 2 --max-missing-count 200 --minQ 20. We further filtered heterozygous SNPs as those have been shown to result from duplicated sequenced and be mostly artifactual in selfing species (Stritt et al. 2022). All the analyses were performed on version3 of the *B. distachyon* genome (https://phytozome-next.jgi.doe.gov).

Flowering time measurement

We performed an experiment in controled conditions from October 2021 until May 2022 in order to test the flowering phenology of the five genetic clades of *Brachypodium distachyon.* Twelve accessions per genetic clade were selected, in addition to the reference accession Bd21. All accessions were treated with five vernalization periods spanning from 2 to 10 weeks, with three replicates per treatment and per accession. Seeds were stratified for at least two weeks before the experiment, and then sowed in pots that were placed in greenhouse conditions (16 h day at 20°C and 8 h dark at 18°C with a light intensity of 200 µMol/m2/s). We distributed the replicates randomly across trays to minimize bias due to position effects. Three weeks after germination the plants were transfered to a cooling chamber (constant temperature at 4°C, 8h light 80 µMol/m2/s, and 16 h dark) for 2, 4, 6, 8, and 10 weeks. At the end of the vernalization treatment, plants were moved back to the greenhouse.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Flowering time was measured as the number of days after return to greenhouse to the first day of spike emergence which is consistent with stage 50 of the Zadoks scale that was used in (Ream et al. 2014). Measurements were taken every two days until the end of the experiment in May. During the experiments, trays were permuted on the table to further limit position effects. We calculated the mean of flowering time per vernalization treatment and accession to estimate flowering from i) the day plants were move to the vernalization chamber or ii) the day after the return to warm conditions. This allowed us to estimate the fastest time for an accession to flower. Results were plotted in R version 4.0.2 (R Core Team, 2018) with the package ggplot2 (Wickham 2016).

Extraction of current and past bioclimatic variables

Raster maps for current monthly solar radiation and altitude were retrieved from worldclim (https://www.worldclim.org). In addition, raster maps for monthly Global Aridity Index (GAI) were obtained from https://cgiarcsi.community/data/global-aridity-and-pet-database/. Bioclimatic variables were then extracted using the R packages raster (v.3.5-2, Hijmans & van Etten, 2012) and rgdal (v.1.5-27, Keitt et al., 2010) for each of the 332 accessions. For solar radiation and GAI, data were also average over spring months (April to June).

For paleo-bioclimatic variables, we used the niche suitability projections for the last glacial maximum (LGM) computed by Minadakis et al (2022). For each genetic clade, we extracted the coordinates of a set of 200 random points per clade in highly suitable habitats (>0.85) with the raster package function rasterToPoints. We retrieved raster maps for LGM from https://www.worldclim.org/data/v1.4/paleo1.4.html and extracted the 19 paleo-bioclimatic variables for the corresponding sites as described above.

Linear Mixed model analyses

Association between flowering time and bioclimatic variables were tested with linear mixed model analyses (LMM) using the R package lme4 (Bates et al. 2015). We subset

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

the 19 classical bioclimatic variables as well as altitude, aridity and solar radiation in spring for the 56 accessions for which flowering time was measured. We performed a principal component analysis (PCA) with the R base function prcomp using the resulting 22 variables. We then ran LMM where flowering time was entered as the response variable, the 22 environmental variables mentioned above as well as the first 2 PCA axes entered separately as fixed factors and the clade of origin as a random factor. Accessions were nested within genetic clades.

We also tested the association between flowering time and SNPs in flowering genes (FT genes) using LMM. To do so, we first extracted and converted the SNPs located in flowering-time genes as a dataframe with vcftools -- bed --extract-FORMAT-info GT, where the bed file contained the position of FT-genes. We then ran LMM where flowering time (fastest time per accession) was entered as the response variable, the SNPs at FT genes entered separately as fixed factors and the clade of origin as random factors. Accessions were nested with genetic clades. FT-gene IDs and position in the genome can be found in tables S4. Epistatic effects were tested with Plink 1.9 (Purcell et al. 2007). Note that because many accessions from the A_East, A_Italia and C clades did not flower without at least six weeks of vernalization (Figure S1), we did not compute a full model including the vernalization treatment as a random factor due to missing data.

For all LMM analyses, the part of the variance explained by the fixed- (marginal R^2) and random effects were computed following Nakagawa and Schielzeth (2013). All the plots were produced with the R package ggplot2 (Wickham 2016). Classical linear models and correlation (Kendall) were ran in R basics.

F_{ST} calculation

Single SNP F_{ST} between accessions of the A and B lineages were calculated with vcftools (Danecek et al. 2011). To account for shifts in the observed F_{ST} values caused by the population structure of *B. distachyon*, the expected F_{ST} distribution under

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

neutral evolution was also estimated using forward simulations run in SLiM version 3.4 (Haller and Messer 2019). The population structure, effective population sizes and time of divergences between lineages of *B. distachyon* during its evolution were modeled in SLiM based on the results of Minadakis et al. (2023). No migration between the different populations in the simulations was allowed, as a lack of interbreeding between the distinct *B. distachyon* clades was reported (Stritt et al. 2022). The simulation, was run 100 times and single SNP F_{ST} were calculated for each simulation to generate the expected F_{ST} distribution under neutrality.

Linkage disequilibrium analyses

To plot linkage disequilibrium (LD) decay, we first thinned the vcf with vcftools --thin 20000 to keep one SNP every 20 kilobases (kb). Intrachromosomal LD (r^2) was calculated with vcftools --geno-r2. We repeated this step by further filtering the vcf for a minimum allele frequency of 0.05 with vcftools -- maf 0.05. For both outputs, we visualized LD decay by plotting r^2 and its 95 CI as a function of the physical distance between SNPs with the R package ggplot2. LD between focal SNPs in flowering-time genes was calculated separately with vcftools –geno-r2 and added to the LD decay plot for comparison. LD plots for focal SNPs were produced with the R package gaston (Pedry and Dandine-Roulland 2020).

Inter-chromosomal LD (r^2) was calculated with vcftools --interchrom-geno-r2 for a subset of 100,000 loci selected randomly in the genome. We then re-sampled 50,000 times five and six loci and calculated the mean LD each time. This allowed us to further calculate the CI around the mean and compare our real data to this distribution.

Scans of selection

The X^TX analysis was performed with BayPass v2.3 using the five genetic clades or three genetic lineages as populations (Gautier 2015). We generated the input file by using vcftools --count to calculate the allele frequency of each SNP present in our vcf (no filtering on minor minimum allele frequency). We then ran Baypass on our actual

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

dataset with the following parameters: -pilotlength 500 -npilot 15 -burnin 2500 -nthreads 6. To calibrate the X^TX and define a threshold of significance for differentiated SNPs, we then used the simulate.baypass function from baypass_utils (Gautier et al. 2015) to generate a pseudo-observed dataset (POD) of 100,000 loci based on the co-variance matrix computed with our real dataset. We then ran Baypass on the POD with the parameters described above. We used the 0.99 quantile of the X^TX calculated for the POD as a threshold of significance for the real dataset. Integrated Haplotype Scores (IHS) were also computed for accessions of the A and B lineages separately with the R package Rehh 3.1 (Gautier et al. 2017)

Functional effect and age estimates of variants

The functional effect of variants was annotated using SnpEff version 5.0e (Cingolani et al. 2012) using default parameters and the provided database for *Brachypodium distachyon*.

The age of each single SNP was computed with GEVA (Albers and McVean 2020) to estimate the average SNP age for each annotated gene (https://phytozome-next.jgi.doe.gov) in the derived A and B genetic lineages as well as in the five genetic clades. All private SNPs to the combined A and B lineages were polarized using the ancestral C lineage using custom R scripts. GEVA was run on the five main scaffolds (corresponding to the five chromosomes) using the genetic map produced by (Huo et al. 2011) and the polarized SNP dataset.

Genome-wide association analysis

We further used the flowering time measurements performed outdoor in Zurich in 2017 (Stritt et al. 2022). We extracted the 19 classical bioclimatic variables as well as altitude, aridity and solar radiation for the 332 accessions as well as a site in Zurich (lat = 43.73693085, lon = 3.69295907) and performed a principal component analysis (PCA) as described above. Linear-mixed model analyses were performed as described above for the greenhouse experiment.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

To identify loci associated with flowering time variation, we performed a genome-wide association analysis with GEMMA (Zhou and Stephens 2012). Due to population structure, we corrected the analysis by a centered relatedness matrix, first produced with the option -gk 1. Association tests were performed using the option -maf 0.05 to exclude SNPs with minor allele frequency with values less than 0.05. Regions were considered significantly associated if displaying at least four markers above FDR threshold in 8Kb windows (overlap of 4Kb). Upset plots were drawn with the R package UpSetR (Gehlenborg 2019).

Overlaps with Environmental association analyses

We made use of the environmental association analyses performed by Minadakis et al. (2023) to assess whether flowering-time genes and the candidate gene identified by the GWAs were associated with current environmental variables. Upset plots were drawn in R with the R package UpSetR (Gehlenborg 2019).

Gene duplication

We checked for potential gene duplication with detettore (https://github.com/cstritt/detettore), a program developed to study structural variation based on short-read sequences (Stritt 2021). We also calculated the proportion of heterozygous sites over the 22 flowering-time genes and 332 accessions, using the raw vcf produced by Minadakis et al. (2023) not filtered for heterozygous SNPs but with the following criteria: --max-alleles 2 --max-missing-count 200 --minQ 20.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Author contribution

NM conceived the study, performed the greenhouse experiment, contributed to the analysis and the writing of the manuscript. MT collected samples, contributed to the analysis. WX, RH and LK contributed to the analysis. DPW contributed with the writing of the manuscript. ACR conceived the study, contributed to the analysis and with the writing of the manuscript.

Data availability

Seeds will be distributed through GRIN (https://www.ars-grin.gov). Accession numbers for sequencing data can be found in (Gordon et al. 2017; Gordon et al. 2020; Skalska et al. 2020; Stritt et al. 2022; Minadakis et al. 2023).

Acknowledgments

We would like to thank the Swiss National Science foundation (project 31003A_182785) and the Research Priority Program Evolution in action from the University of Zurich for their generous funding. The authors are grateful to Beat Keller and Ueli Grossniklaus for providing space for the flowering time experiments as well as the Amasino lab, Sam Yeaman and Dieter Ebert for stimulating discussion on flowering time and polygenic selection.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Supplementary Materials

The supplementary materials consist in six figures and seven tables.

Table S1: Raw results of the flowering time experiment.

Table S2: Differences in flowering time among the five genetic clades.

Table S3: Output of the linear model analyses between flowering time under greenhouse conditions and environmental variables.

Table S4: Flowering time gene position and ID.

Table S5: output of the LMM between flowering time and SNP at FT genes for significantly associated SNPs.

Table S6: Flowering-time gene age estimates.

Table S7: Output of the linear model analyses between flowering time under oputdoor conditions and environmental variables.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

References

Ågren J, Oakley CG, Lundemo S, Schemske DW. 2017. Adaptive divergence in flowering time among natural populations of Arabidopsis thaliana: Estimates of selection and QTL mapping. Evolution. **71**:550–564.

Ågren J, Schemske DW. 2012. Reciprocal transplants demonstrate strong adaptive differentiation of the model organism Arabidopsis thaliana in its native range. New Phytol. **194**:1112–1122.

Albers PK, McVean G. 2020. Dating genomic variants and shared ancestry in population-scale sequencing data. PLoS Biol. **18**:1–26.

Alvarez MA, Li C, Lin H, Joe A, Padilla M, Woods DP, Dubcovsky J. 2023. EARLY FLOWERING 3 interactions with *PHYTOCHROME B* and *PHOTOPERIOD1* are critical for the photoperiodic regulation of wheat heading time. PLOS Genet. **19**(5): e1010655.

Anderson JT, Inouye DW, McKinney AM, Colautti RI, Mitchell-Olds T. 2012. Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. Proc. R. Soc. B Biol. Sci. **279**:3843–3852.

Anderson JT, Lee CR, Rushworth C, Colautti R and Mitchell-Olds T. 2013. Genetic tradeoffs and conditional neutrality contribute to local adaptation. Mol Ecol **22**:699–708.

Andrés F, Coupland G. 2012. The genetic basis of flowering responses to seasonal cues. Nature Reviews Genetics. **13**, 627–639.

Bates D, Machler M, Bolker BM, Walker SC. 2015. Fitting linear mixed-effects models using LME4. Journal of Statistical Software **67**: 1–48.

Bettgenhaeuser J, Corke FMK, Opanowicz M, Green P, Hernández-Pinzón I, Doonan JH, Moscou MJ. 2017. Natural variation in brachypodium links vernalization and flowering time loci as major flowering determinants. Plant Physiol. **173**:256–268.

Blackman BK. 2017. Changing responses to changing seasons: Natural variation in the plasticity of flowering time. Plant Physiol. **173**:16–26.

Blümel M, Dally N, Jung C. 2015. Flowering time regulation in crops-what did we learn from Arabidopsis? Curr. Opin. Biotechnol. **32**:121–129.

Booker TR, Yeaman S, Whiting JR, Whitlock MC. 2023. The WZA : A window-based method for characterizing genotype–environment associations. Mol. Ecol. Resour.**00** : 1–16.

Bouché F, Woods DP, Linden J, Li W, Mayer KS, Amasino RM, Périlleux C. 2022. EARLY FLOWERING 3 and Photoperiod Sensing in Brachypodium distachyon. Front. Plant Sci. **12**:1–14.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Bourgeois Y, Stritt C, Walser JC, Gordon SP, Vogel JP, Roulin AC. 2018. Genome-wide scans of selection highlight the impact of biotic and abiotic constraints in natural populations of the model grass Brachypodium distachyon. Plant J. **96**:438–451.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F. 2010. Linkage and association mapping of Arabidopsis thaliana flowering time in nature. PLoS Genet. **6**:40.

Buckler ES et al. 2009. The Genetic Architecture of Maize. **325**:714.

Cao S, Luo X, Xie L, Gao C, Wang D, Holt BF, Lin H, Chu C, Xia X. 2020. The florigen interactor BdES43 represses flowering in the model temperate grass Brachypodium distachyon. Plant J. **102**:262–275.

Cingolani P , Platts A, Wang L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain. Fly **6**:80-92.

Chouard P. 1960. Vernalization and its Relations to Dormancy. Annu. Rev. Plant Physiol. **11**:191–238.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, Depristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. Bioinformatics **27**:2156–2158.

Del'Acqua MD, Zuccolo A, Tuna M, Gianfranceschi L, Pè ME. 2014. Targeting environmental adaptation in the monocot model Brachypodium distachyon : a multi-faceted approach. BMC Genomics**15**:801.

Eichten S, Stuart T, Srivastava A, Borevitz J. 2016. DNA Methylation profiles of diverse Brachypodium distachyon aligns with underlying genetic diversity. Genome Res. **26**(11):1520-1531.

Gaudinier A, Blackman Benjamin K. 2020. Evolutionary processes from the perspective of flowering time diversity. New Phytol. **225**:1883–1898.

Gautier M. 2015. Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. Genetics **201**(4):1555-79.

Gautier M, Vitalis R (2012). rehh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics, **28**(8) 1176-1177.

Gehlenborg N. 2019. UpSetR: A More Scalable Alternative to Venn and Euler Diagrams for Visualizing Intersecting Sets. R package version 1.4.0. https://CRAN.R-project.org/package=UpSetR.

Gloss AD, Vergnol A, Morton TC, Laurin PJ, Roux F, Bergelson J. 2022. Genome-wide association mapping within a local Arabidopsis thaliana population more fully reveals the genetic architecture for defensive metabolite diversity. Philos. Trans. R. Soc. B Biol. Sci. **377**: 20200512.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Gordon SP, Contreras-Moreira B, Levy JJ, Djamei A, Czedik-Eysenberg A, Tartaglio VS, Session A, Martin J, Cartwright A, Katz A, et al. 2020. Gradual polyploid genome evolution revealed by pan-genomic analysis of Brachypodium hybridum and its diploid progenitors. Nat. Commun. **11**:1–16.

Gordon SP, Contreras-moreira B, Woods DP, Marais DL Des, Burgess D, Shu S, Stritt C, Roulin AC, Schackwitz W, Tyler L, et al. 2017. Extensive gene content variation in the Brachypodium distachyon pan-genome correlates with population structure. Nat. Commun. **8**:2184.

Gould BA, Stinchcombe JR. 2017. Population genomic scans suggest novel genes underlie convergent flowering time evolution in the introduced range of Arabidopsis thaliana. Mol. Ecol. **26**:92–106.

Gupta S, Harkess A, Soble A, Etten M Van, Leebens-Mack J, Baucom RS. 2023. Inter-chromosomal linkage disequilibrium and linked fitness cost loci associated with selection for herbicide resistance. New Phytol. **238**:1263-1277.

Hall, Willis, Hall MC, Willis JH. 2006. Divergent selection on flowering time contributes to local adaptation in Mimulus guttatuspopulations. Evolution. **60**:2466–2477.

Haller BC, Messer PW. 2019. SLIM 3: forward genetic simulations beyond the Wright–Fisher model. Molecular Biology and Evolution **36**: 632–637.

Hasterok R, Catalan P, Hazen SP, Roulin AC, Vogel JP, Wang K, Mur LAJ. 2022. Brachypodium: 20 years as a grass biology model system; the way forward? Trends Plant Sci. **27**:1002–1016.

Higgins JA, Bailey PC, Laurie DA. 2010. Comparative genomics of flowering time pathways using *Brachypodium distachyon* as a model for the temperate Grasses. PLoS One. **5**(4): e10065.

Hijmans RJ, Van Etten J. (2012) Geographic analysis and modeling with raster data. http://raster.r-forge.r-project.org/

Hohenlohe PA, Bassham S, Currey M, Cresko WA. 2012. Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. Philos. Trans. R. Soc. B Biol. Sci. **367**:395–408.

Huo N, Garvin DF, You FM, Luo SMM, Gu YQ, Lazo GR, Philip J. 2011. Comparison of a high-density genetic linkage map to genome features in the model grass Brachypodium distachyon. Theor. Appl. Genet. **123**:455–464.

International Brachypodium Initiative. 2010. Genome sequencing and analysis of the model grass {B}rachypodium distachyon. Nature **463**:763–768.

Izawa T. 2007. Adaptation of flowering-time by natural and artificial selection in Arabidopsis and rice. J. Exp. Bot. **58**:3091–3097.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Jaegle B, Pisupati R, Soto-Jiménez LM, Burns R, Rabanal FA, Nordborg M. 2023. Extensive sequence duplication in Arabidopsis revealed by pseudo-heterozygosity. Genome Biol. **24**:1–19.

Keitt TH, Bivand R, Pebesma E, Rowlingson B. 2010. RGDAL: bindings for the 'Geospatial' data abstraction library. R package v.1.5-27. https://cran.r-project.org/web/packages/rgdal/index.html

Kennedy A, Geuten K. 2020. The Role of FLOWERING LOCUS C Relatives in Cereals. Front. Plant Sci. **11:**617340.

Lomax A, Woods DP, Dong Y, Bouché F, Rong Y, Mayer KS, Zhong X, Amasino RM. 2018. An ortholog of CURLY LEAF/ENHANCER OF ZESTE like-1 is required for proper flowering in Brachypodium distachyon. Plant J. **93**:871–882.

Lottheros K. 2023. The paradox of adaptive trait clines with nonclinal patterns in the underlying genes. Proc. Natl. Acad. Sci. **120**: e2220313120

Malmberg RL, Held S, Waits A, Mauricio R. 2005. Epistasis for fitness-related quantitative traits in Arabidopsis thaliana grown in the field and in the greenhouse. Genetics **171**:2013–2027.

Des Marais DL, Lasky JR, Verslues PE, Chang TZ, Juenger TE. 2017. Interactive effects of water limitation and elevated temperature on the physiology, development and fitness of diverse accessions of Brachypodium distachyon. New Phytol. **214**:132–144.

Minadakis N, Williams H, Horvath R, Stritt C, Thieme M, Bourgeois Y, Roulin AC. 2023. New resources for environmental genomics in the wild Mediterranean grass *Brachypodium. distachyon*. bioRxiv https://doi.org/10.1101/2023.06.01.543285

Monnahan PJ, Kelly JK. 2017. The genomic architecture of flowering time varies across space and time in mimulus guttatus. Genetics **206**:1621–1635.

Nakagawa S, Schielzeth H. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol. Evol. **4:**133–142.

Navarro JAR, Wilcox M, Burgueño J, Romay C, Swarts K, Trachsel S, Preciado E, Terron A, Delgado HV, Vidal V, et al. 2017. A study of allelic diversity underlying flowering-time adaptation in maize landraces. Nature genetics **49**: 476–480.

Nunes TDG, Zhang D, Raissig MT. 2020. Form, development and function of grass stomata. Plant J. **101**:780–799.

Park L. 2019. Population-specific long-range linkage disequilibrium in the human genome and its influence on identifying common disease variants. Sci. Rep. **9**:1–13.

Perdry H, Dandine-Roulland C (2022). gaston: Genetic Data Handling (QC, GRM, LD, PCA) & Linear Mixed Models. R package version 1.5.9. https://CRAN.R-project.org/package=gaston.

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, et al. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. **81**:559–575.

Qian C, Yan X, Shi Y, Yin H, Chang Y, Chen J, Ingvarsson PK, Nevo E, Ma XF. 2020. Adaptive signals of flowering time pathways in wild barley from Israel over 28 generations. Heredity **124**:62–76.

Qin Z, Bai Y, Muhammad S, Wu X, Deng P, Wu J, An H, Wu L. 2019. Divergent roles of FT-like 9 in flowering transition under different day lengths in *Brachypodium distachyon*. Nat. Commun. **10**:812

Raissig MT, Woods DP. 2022. The wild grass Brachypodium distachyon as a developmental model system. Current Topics in Developmental Biology **147**:33-71

Ream TS, Woods DP, Amasino RM. 2012. The molecular basis of vernalization in different plant groups. Cold Spring Harb. Symp. Quant. Biol. **77**:105–115.

Ream TS, Woods DP, Schwartz CJ, Sanabria CP, Mahoy JA, Walters EM, Kaeppler HF, Amasino RM. 2014. Interaction of Photoperiod and Vernalization Determines Flowering Time of *Brachypodium distachyon*. Plant Physiol. **164**:694–709.

Ruelens P, Maagd RA, Proost S, Teissen G, Geuten K, Kaufmann K. 2013. FLOWERING LOCUS C in monocots and the tandem origin of angiosperm-specific. Nature Com. **4**:2280

Sharma N, Ruelens P, D'hauw M, Maggen T, Dochy N, Torfs S, Kaufmann K, Rohde A, Geuten K. 2017. A flowering locus C homolog is a vernalization-regulated repressor in *Brachypodium* and is cold regulated in wheat. Plant Physiol. **173**:1301–1315.

Shaw LM, Li C , Woods DP, Alvarez MA, Lin H, Lau MY, Chen A, Dubcovsky J. 2020. Epistatic interactions between modulate the photoperiodic response in wheat. PLoS Genet **16**(7): e1008812.

Skalska A, Stritt C, Wyler M, Williams HW, Vickers M, Han J, Tuna M, Tuna GS, Susek K, Swain M, et al. 2020. Genetic and methylome variation in Turkish brachypodium distachyon accessions differentiate two geographically distinct subpopulations. Int. J. Mol. Sci. **21**:1–17.

Nunes TDG, Berg LS, Slawinska MW, Zhang D, Redt L, Sibout R, Vogel J, Laudencia-Chingcuanco D, Jesenofsky B, Lindner H, Raissig MT. 2023. Regulation of hair cell and stomatal size by a hair cell-specific peroxidase in the grass *Brachypodium distachyon*. Curr. Biol. **33**, 1844–1854

Stritt C, Gimmi EL, Wyler M, Bakali AH, Skalska A, Hasterok R, Mur LAJ, Pecchioni N, Roulin AC. 2022. Migration without interbreeding: Evolutionary history of a highly selfing Mediterranean grass inferred from whole genomes. Mol. Ecol. **31**:70–85.

Stritt C, Gordon SP, Wicker T, Vogel JP, Roulin AC. 2018. Recent activity in expanding populations and purifying selection have shaped transposable element landscapes

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

across natural accessions of themediterranean grass brachypodium distachyon. Genome Biol. Evol. **10**:304–318.

Stritt C, Thieme O, Roulin A.C. 2021. Detecting signatures of TE polymorphisms in short-read sequencing data. Methods Mol Biol. **2250**:177-187.

Stritt C, Wyler M, Gimmi EL, Pippel M, Roulin AC. 2020. Diversity, dynamics and effects of long terminal repeat retrotransposons in the model grass {B}rachypodium distachyon. New Phytol. **227**:1736–1748.

Takou M, Wieters B, Kopriva S, Coupland G, Linstädter A, De Meaux J. 2019. Linking genes with ecological strategies in Arabidopsis thaliana. J. Exp. Bot. **70**:1141–1151.

Wadgymar SM, Lowry DB, Gould BA, Byron CN, Mactavish RM, Anderson JT. 2017. Identifying targets and agents of selection: innovative methods to evaluate the processes that contribute to local adaptation. Methods Ecol. Evol. **8**:738–749.

Weinig C, Ungerer MC, Dorn LA, Kane NC, Toyonaga Y, Halldorsdottir SS, Mackay TFC, Purugganan MD, Schmitt J. 2002. Novel loci control variation in reproductive timing in Arabidopsis thaliana in natural environments. Genetics **162**:1875–1884.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Available from: https://ggplot2.tidyverse.org

Wilczek AM, Burghardt LT, Cobb AR, Cooper MD, Welch SM, Schmitt J. 2010. Genetic and physiological bases for phenological responses to current and predicted climates. Philos. Trans. R. Soc. B Biol. Sci. **365**:3129–3147.

Wilhelm EP, Turner AS, Laurie DA. 2009. Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (Triticum durum Desf .). Theor. Appl Genet **118**:285–294.

Wilson PB, Streich JC, Murray KD, Eichten SR, Cheng R, Aitken NC, Spokas K, Warthmann N, Gordon SP, Vogel JP, et al. 2019. Global diversity of the Brachypodium species complex as a resource for genome-wide association studies demonstrated for agronomic traits in response to climate. Genetics **211**:317–331.

Woods DP, Li W, Sibout R, Shao M, Laudencia-Chingcuanco, Vogel J, Dubcovsky J, Amasino RM. 2023. PHYTOCHROME C regulation of photoperiodic flowering via PHOTOPERIOD1 is mediated by EARLY FLOWERING 3 in *Brachypodium distachyon*. PLoS Genet **19**(5): e1010706.

Woods D, Dong Y, Bouche F, Bednarek R, Rowe M, Ream T, Amasino R. 2019. A florigen paralog is required for short-day vernalization in a pooid grass. Elife **8**:1–16.

Woods DP, Bednarek R, Bouché F, Gordon SP, Vogel JP, Garvin DF, Amasino RM. 2017a. Genetic architecture of flowering-time variation in *Brachypodium distachyon*. Plant Physiol. **173**:269–279.

Woods DP, Dong Y, Bouché F, Mayer K, Varner L, Ream TS, Thrower N, Wilkerson C, Cartwright A, Sibout R, et al. 2020. Mutations in the predicted DNA polymerase

bioRxiv preprint doi: https://doi.org/10.1101/2023.07.11.548268; this version posted August 16, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

subunit POLD3 result in more rapid flowering of *Brachypodium distachyon* . New Phytol. **227**: 1725–1735

Woods DP, McKeown MA, Dong Y, Preston JC, Amasino RM. 2016. Evolution of VRN2/Ghd7-like genes in vernalization-mediated repression of grass flowering. Plant Physiol. **170**:2124–2135.

Woods DP, Ream TS, Bouché F, Lee J, Thrower N, Wilkerson C, Amasino RM. 2017b. Establishment of a vernalization requirement in *Brachypodium distachyon* requires *REPRESSOR OF VERNALIZATION1*. Proc. Natl. Acad. Sci. **114**:6623–6628.

Woods DP, Ream TS, Minevich G, Hobert O, Amasino RM. 2014. PHYTOCHROME C is an essential light receptor for photoperiodic flowering in the temperate grass, Brachypodium distachyon. Genetics **198**:397–408.

Wu L, Liu Dongfeng, Wu J, Zhang R, Qin Z, Liu Danmei, Li A, Fu D, Zhai W, Mao L. 2013. Regulation of FLOWERING LOCUS T by a MicroRNA in *Brachypodium distachyon*. The Plant Cell. **25** :4363–4377,

Würschum T, Boeven PHG, Langer SM, Longin CFH, Leiser WL. 2015. Multiply to conquer : Copy number variations at Ppd-B1 and Vrn-A1 facilitate global adaptation in wheat. BMC Genetics. **16**:96

Yan W, Wang B, Chan E, Mitchell-Olds T. 2021. Genetic architecture and adaptation of flowering time among environments. New Phytol. **230**:1214–1227.

Yeaman S. 2022. Evolution of polygenic traits under global vs local adaptation. Genetics **220**(1) : iyab134

Yeaman S, Gerstein AC, Hodgins KA, Whitlock MC. 2018. Quantifying how constraints limit the diversity of viable routes to adaptation. PLoS Genet. **14**:1–25.

Yeaman S, Hodgins KA, Lotterhos KE, Suren H, Nadeau S, Degner JC, Nurkowski KA, Smets P, Wang T, Gray LK, et al. 2016. Convergent local adaptation to climate in distantly related conifers. Science. **353**:1431–1433.

Zan Y, Carlborg Ö. 2019. A Polygenic Genetic Architecture of Flowering Time in the Worldwide Arabidopsis thaliana Population. Mol. Biol. Evol. **36**:141–154.

Zhang D, Abrash EB, Nunes TDG, Prados IH, Gil MXA, Jesenofsky B, Lindner H, Bergmann DC, Raissig MT. 2022. Opposite polarity programs regulate asymmetric subsidiary cell divisions in grasses. Elife **11**:e79913.

Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. **44**:821–824.