



Genetic adaptation of phenological stages in Chinese and European elite soybeans (*Glycine max* [L.] Merr.) across latitudes in Central Europe

Xindong Yao¹  | Martin Pachner¹ | Leopold Rittler² | Volker Hahn³ | Willmar Leiser³ | Christine Riedel⁴ | Raluca Rezi⁵ | Claude-Alain Bétrix⁶ | Jerzy Nawracała⁷ | Inna Temchenko⁸ | Vuk Đorđević⁹ | Li-Juan Qiu¹⁰ | Johann Vollmann¹ 

¹Department of Crop Sciences, University of Natural Resources and Life Sciences Vienna (BOKU), Tulln an der Donau, Austria

²Donau Soja, Vienna, Austria

³State Plant Breeding Institute, University of Hohenheim, Stuttgart, Germany

⁴Bavarian State Research Center for Agriculture (LfL), Institute for Crop Science and Plant Breeding (IPZ), Ruhstorf an der Rott, Germany

⁵Research and Development Station for Agriculture Turda (RDSA Turda), Turda, Cluj County, Romania

⁶Agroscope, Nyon, Switzerland

⁷Department of Genetics and Plant Breeding, Faculty of Agronomy, Horticulture and Bioengineering, Poznań University of Life Sciences, Poznań, Poland

⁸Institute of Feed Research and Agriculture of Podillia, Vinnytsia, Ukraine

⁹Department of Legumes, Institute of Field and Vegetable Crops, Novi Sad, Serbia

¹⁰National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, People's Republic of China

Correspondence

Johann Vollmann, Department of Crop Sciences, University of Natural Resources and Life Sciences Vienna (BOKU), Konrad Lorenz Strasse 24, 3430 Tulln an der Donau, Austria.
Email: johann.vollmann@boku.ac.at
Li-Juan Qiu, National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, No. 12 Zhongguancun South Street, Haidian District, Beijing 100081, People's Republic of China.
Email: qiulijuan@caas.cn

Funding information

National Key Research and Development Program of China, Grant/Award Number: 2019YFE0105900; German State of Bavaria; German Federal State of Baden-Württemberg; Saatgut Austria; Federal Authorities of the Swiss Confederation; University of Natural Resources and Life Sciences, Vienna

Abstract

Genetic adaptation of soybean phenological stages to high-latitude long-day environments is the major pre-requisite for enhancing regional plant-based protein production. Both *E*-genes controlling flowering and growing environment determine timing of flowering and maturity, and *E*-gene composition might differ between Chinese and European soybeans bearing the potential for improving adaptability. Therefore, 140 early maturity elite soybean cultivars of either Chinese or European origin were genotyped for the *E1* to *E4* flowering loci, and genotypes were tested across 17 European environments spanning a latitude range from 45 to 52°N in order to determine effects of various *E*-allele combinations. Differences in *E*-allele composition between Chinese and European cultivars were largest for the loci *E1* and *E3*. Wild-type alleles significantly delayed flowering, and effects of particular *E*-alleles were depending on geographic latitude. Consequently, photoperiod-insensitive *E*-haplotypes carrying several non-functional alleles proved to be suitable for cultivation in higher latitudes, whereas photoperiod-sensitive late-maturity *E*-haplotypes are adapted to lower latitudes only. Thus, breeding for new *E*-haplotypes through combining Chinese and

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Plant Breeding* published by Wiley-VCH GmbH.

European alleles could enhance the potential for further soybean adaptation to northern growing regions.

KEYWORDS

genetic adaptation, growing environment, latitude, photoperiod sensitivity, soybean, time to flowering

1 | INTRODUCTION

Genetic adaptation of crop plants to a given agro-ecological environment is focusing on optimization of life cycle for achieving the best agronomic performance under specific conditions. Soybean (*Glycine max* [L.] Merr.) is a short-day plant in terms of photoperiodic induction of flowering, in which long day-length of northern latitudes delays flowering and maturity. In order to describe the latitudinal region of adaptation of individual soybean cultivars, the system of maturity groups (MG) is commonly used in North America and Europe (Kurasch et al., 2017; Langewisch et al., 2017). Recently, the complex ecotype and cropping system categorization established for Chinese soybeans has also been translated into the more precise relative maturity group classification system (Song et al., 2023), which is helpful for optimum cultivar selection and germplasm exchange. In soybean breeding, genetic adaptation of phenology to higher latitude environments is an important activity at present, as soybean production is assumed to expand to northern growing areas in many soybean production regions of the world. This is due to the requirement of mitigating climate change effects as well as for covering the growing need of a plant-based protein supply from domestic production (Nendel et al., 2023; Niedbała et al., 2022; Ort et al., 2022; Vollmann & Škrabišová, 2023). Besides stem architecture and symbiotic nitrogen fixation, flowering time is considered as most relevant for determining soybean yield (Fang et al., 2024). Thus, in-depth knowledge on the genetic architecture and diversity of early maturity germplasm as well as on photoperiodic reactions of particular genotypes is of great interest for developing suitable soybean cultivars meeting these new requirements.

Time to flowering (TTF) and time to maturity (TTM) are quantitative characters which are controlled by a number of genes with major and minor effects. This includes at least 11 classical flowering genes (*E*-genes, *J*-locus) and several additional factors associated with photoperiodic or thermal regulation of flowering (Wu et al., 2023). Based on *E*-gene composition, a molecular model of maturity group classification has been proposed for North American soybean cultivars to increase the efficiency of breeding (Langewisch et al., 2017). Furthermore, numerous QTL affecting flowering have been identified in genome-wide association studies using different diversity panels or other collections of selected genotypes (Fu et al., 2020; Liu et al., 2021; Zhu et al., 2023). These findings illustrate the high complexity of the regulatory networks associated with soybean photoperiodic reactions; simultaneously, the large number of factors and their interactions with day-length and other environmental characteristics bear a considerable potential for

further adaptation through molecular breeding approaches (Wu et al., 2023; Zhang et al., 2017).

In early maturity soybeans, *E*-genes *E1* to *E4* are considered as most relevant in controlling the TTF (Dong et al., 2023; Kurasch et al., 2017; Liu et al., 2020; Perfil'ev et al., 2023), and their dominant wild-type alleles have inhibitory effects on flowering (Wu et al., 2023). These genes have been reviewed in full detail in several recent publications (see Dietz et al., 2021; Dong et al., 2023; Hou et al., 2023; Liu et al., 2020; Perfil'ev et al., 2023); *E1* is a legume specific transcription factor of the B3-domain suppressing the expression of florigen genes, which consequently inhibits flowering; *E2* is an orthologue of the *Arabidopsis GIGANTEA* gene regulating diurnal expression of florigenes in a circadian clock manner; the loci *E3* and *E4* are phytochrome A photoreceptor genes coding for sensitivity to red and far-red light under long-day conditions, thereby modulating *E1* expression. Example proposals of the regulatory network of photoperiodically controlled flowering in soybean based on established pathways between the four *E*-genes have been presented by Hou et al. (2023) and Fang et al. (2024). For all *E*-genes, a dominant wild-type allele and different mutant alleles including fully or partly functional alleles or non-functional null alleles have been identified in different soybean genotypes. Subsequently, specific allele combinations (*E*-haplotypes) have been described with respect to cultivars used in different soybean growing regions such as China (Liu et al., 2020), North America (Goettel & An, 2017; Langewisch et al., 2017), Europe (Kurasch et al., 2017; Miladinović et al., 2018) or Russia (Perfil'ev et al., 2023). These allele combinations represent the results of numerous cycles of both indirect and direct selection for adaptation to the conditions of specific agro-ecological environments.

Differences in the patterns of genetic diversity in elite cultivars between soybean production regions such as China and Europe (Yao et al., 2023) or China and North America (Rossi et al., 2013) are bearing a considerable potential for reciprocally enhancing breeding progress. This also applies to adaptation, where new allele combinations of *E*-genes would support a more precise tuning of phenological traits such as TTF and TTM or the optimum partition between vegetative and reproductive development phases. Therefore, the objectives of the present study were (i) to compare *E*-gene and *E*-haplotype distribution in elite soybean cultivars of early maturity groups from Chinese and European breeding programmes, (ii) to analyse effects of different *E*-alleles on TTF and (iii) to determine the sensitivity of *E*-alleles and *E*-haplotypes to geographic latitude or day-length as an indicator of latitude adaptation potential. In order to meet these objectives, 140 soybean cultivars of maturity groups 0000 to III were tested in a large collaborative experiment in 17 Central European environments spanning a latitude range from 45°N to 52°N. Results are anticipated to

support soybean breeding and cultivation in various regions for enhancing domestic plant-based protein production towards a mitigation of climate change effects.

2 | MATERIALS AND METHODS

2.1 | Elite soybean cultivars

An initial set of 156 early maturity soybean varieties was collected for the present study (see descriptive information in Table S1 and Yao et al., 2023). The total set consisted of 79 Chinese (CN) and 77 European (EU) elite cultivars that were each assigned to one of four different experimental groups of comparable group size based on their maturity classification (Table S2). European cultivars were provided from soybean breeding companies in Austria, France, Germany, Hungary, Italy, Poland, Serbia, Romania, Switzerland and Ukraine spanning a wide range of geographic adaptation and maturity groups from 0000 to II. Chinese soybean cultivars were provided from breeding institutions located in the north-east of China comprising soybean maturity groups 000 to III, that is, Northeastern Spring (NEsp) or Northern Spring (Nsp) cultivars, according to the Chinese soybean ecotype classification.

2.2 | Genotyping of flowering/maturity *E* loci

The allelic composition at soybean flowering and maturity loci *E1*, *E2*, *E3* and *E4* was genotyped for all 156 varieties at the State Plant Breeding Institute, University of Hohenheim (Stuttgart, Germany) as described earlier (Kurasch et al., 2017; Tsubokura et al., 2014; Yao et al., 2023). For genotyping, DNA from pre-germinated soybean root tips was isolated using the Wizard[®] DNA extraction kit from Promega (Promega Corp., Madison, WI, USA). For PCR amplification with allele-specific primers for loci *E1* to *E4*, individual PCR conditions were applied depending on the respective allele (see Kurasch et al., 2017 for primer sequences utilized and PCR details). Based on the allelic composition at the four different *E* loci, genotypes were also classified into 16 distinct *E*-haplotypes (Table S3) for further analyses.

2.3 | Environments, experimental design and data collection

Out of all cultivars, 140 entries (Table S1a, entries 1–140) were utilized in field experiments, while 16 cultivars (Table S1b, entries 141–156) were excluded due to limited seed availability. Field experiments were carried out in nine different locations between latitudes 45.34°N and 51.69°N (Figure S1) covering a wide range in agro-ecological and photoperiod (i.e. day-length) conditions. Data were collected from nine locations in 2019 and from eight locations in 2020 comprising 17 environments in total (Table S4). Individual experiments were arranged as randomized complete block design (RCB) with two replications. With respect to seed sources for each location, seed materials were planted as received from the different breeding

companies in the first experimental year, whereas in the second year, seed was used from the previous harvest at the respective location. Plots were sown as single-row plots of approx. 2 m length at a sowing rate of 50 seeds m⁻². In terms of agronomic management, field trials were carried out according to local soybean cultivation needs and practical experience of individual experimenters. TTF (expressed in days after May 31) and TTM (expressed in days after July 31) were recorded from field plots, and the duration of reproductive phase (R-phase, days from TTF to TTM) was calculated thereof.

2.4 | Statistical analyses and data visualization

From environments 1 to 10 (Table S4), phenological data were collected from both replications of the 140 cultivars grown, whereas from environments 11–17 phenological data were available only from one replication. Accordingly, analysis of variance (ANOVA) was carried out for environments 1–10 using full ANOVA models. In addition, phenological data were analysed across all 17 environments using mean data from environments 1–10 and single replication data from environments 11–17. For testing the significance of various model effects, either fixed or mixed linear model ANOVAs using Proc GLM or Proc MIXED of SAS version 9.4 package (SAS Institute, Cary, NC, USA) were carried out. A GLM-ANOVA was utilized to test *E*-haplotypes and regions of cultivar origin (China vs. Europe) within *E*-haplotypes as fixed effects over all 17 environments (see Table S6). In addition, a mixed-model ANOVA was applied for 10 environments with replicated field data (RCB design) available for testing random effects such as replications within environments and environments, and their interactions with the fixed effects regions of origin and genotypes within regions (see Table S7). For additional statistical analyses including graphical data visualizations and PCA bi-plot calculations, the OriginPro version 2023b SR1 software package (OriginLab Corp., Northampton, MA, USA) was utilized.

3 | RESULTS

3.1 | Distribution of *E*-genotypes

At the *E1* locus, 72% of genotypes are carrying the missense allele *e1-as*, whereas the dominant wild-type allele *E1* is present in 16% and the non-functional mutant (null) allele *e1-nl* in 12% of all genotypes. At *E2*, the nonsense mutant allele *e2-ns* is predominant occurring in 92% of genotypes, whereas the dominant wild-type allele *E2* was found in 8% of genotypes only. Similarly, at *E3*, the dominant allele *E3* was identified in 8%, *E3-Ha* in 39% and *e3-tr* in 50% of cultivars; in addition, three rare alleles—*e3-fs*, *e3-Mo* and *e3-ns*—were found in 1.6%, 0.3% and 0.3% of genotypes, respectively. In contrast, at the *E4* locus, the dominant wild-type allele *E4* was predominant and found at a frequency of about 91%, whereas alleles *e4-kes* and *e4-SORE-1* were present in only 1% and 8%, respectively.

Striking differences in the frequency of *E*-alleles are existing not only between experimental groups (Figure 1a) but also between

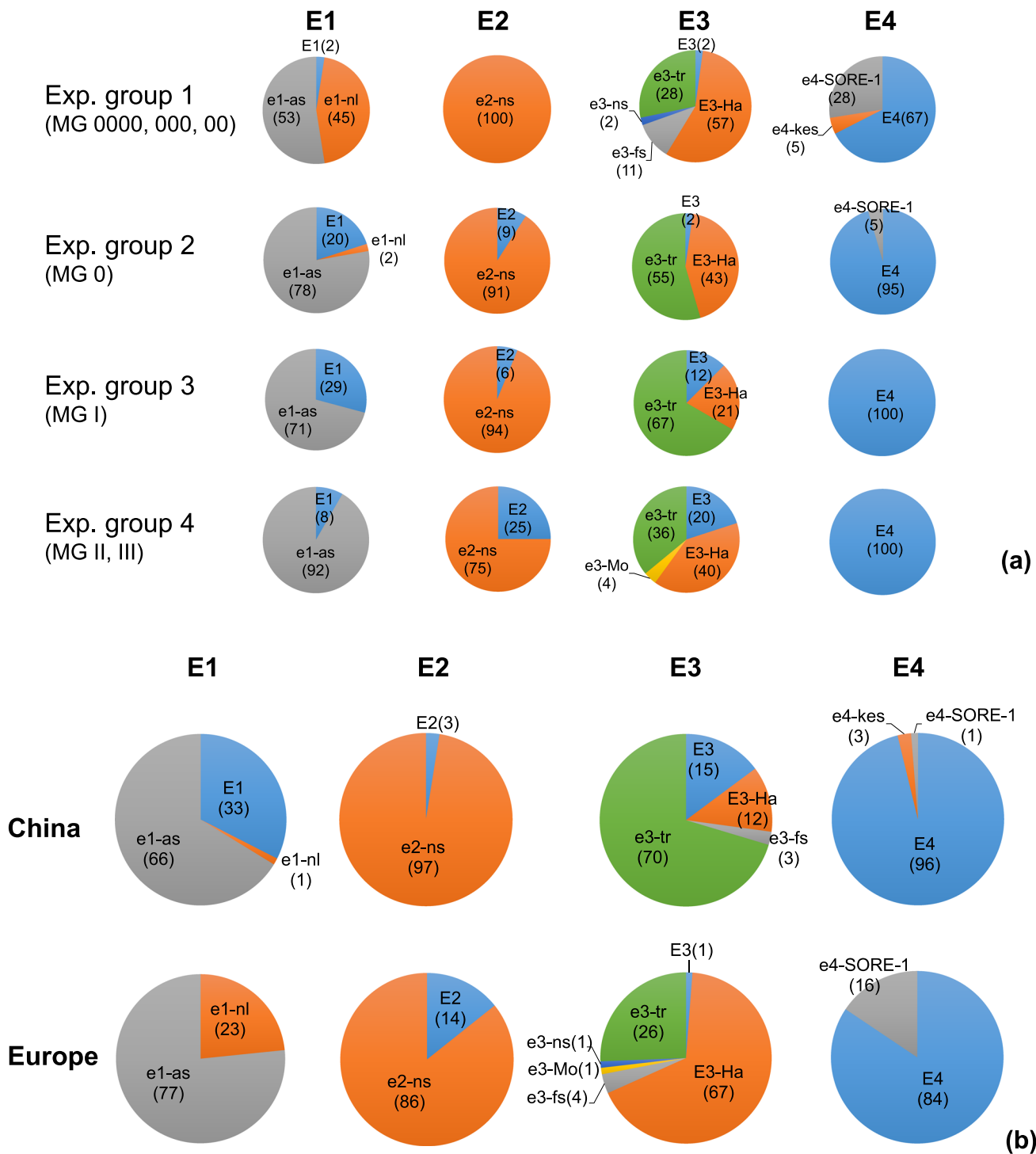


FIGURE 1 Distribution of alleles at the loci E1, E2, E3 and E4 in 156 Chinese and European elite cultivars across four experimental groups (a) and the two regions of origin (b); percentage of cultivars per pie chart section in parentheses.

Chinese and European cultivars as well (Figure 1b). Wild-type alleles E1-E4 causing delayed flowering were found at highest frequencies in the late maturity experimental groups 3 and 4 (i.e. soybean MGs I-III). Consequently, alleles causing early flowering such as *e1-nl* or *e4-SORE-1* were present in early maturity experimental groups 1 and 2 only (Figure 1a). In European cultivars, the dominant wild-type allele E1 is absent, and dominant alleles E3 and E4 are present in lower

frequency than in the Chinese cultivars (Figure 1b). Remarkably, the non-functional mutant allele *e1-nl* is present in many early maturity cultivars from Europe, whereas it is rarely found in Chinese ones. Similarly, alleles *E3-Ha* and *e4-SORE-1* are more often found in European than in Chinese cultivars (Figure 1b). In contrast, allele *e3-tr* is predominant in Chinese cultivars but present at a lower rate in European ones.

Specific combinations of alleles at the four *E*-loci are considered as *E*-haplotypes (Table S3), and 16 different *E*-haplotypes have been identified across the whole set of genotypes (Table S5). While seven *E*-haplotypes were found both in China and Europe, five were unique to the Chinese sub-set; remarkably, four *E*-haplotypes were present in the European sub-set only (Figure S2).

3.2 | Overall phenology across environments

Growing environments, the *E*-haplotype and the region of origin of *E*-haplotypes played a significant role in soybean phenology (ANOVA results; Table S6). Both the beginning of flowering and TTM were affected by environments (Figure S3) with earliest flowering and maturity occurring in the Serbian and Romanian environments and latest flowering occurring in the northernmost Polish environments. Generally lower temperatures during the spring period of 2020 as compared to 2019 caused a delay of flowering in 2020 at all locations, which also affected TTM in most of the locations later on (Figure S3). Apart from the environmental influence, *E*-gene composition (resulting in 16 different *E*-haplotypes) played the major role in phenology, as shown for TTF (Figure S4). *E*-haplotypes 8 and 15 carrying several mutant alleles (8: *e1-as e2-ns E3-Ha e4-kes*; 15: *e1-nl e2-ns E3-Ha e4-SORE-1*) had the shortest TTF, whereas *E*-haplotypes with dominant and late flowering alleles such as haplotypes 4 and 5 (4: *e1-as E2 E3-Ha E4*; 5: *e1-as E2 e3-tr E4*) had an average TTF up to 17 days later than the earliest one. Variation in TTF within particular *E*-haplotypes (e.g. *E*-haplotypes 2, 7, 4) is rather wide, which indicates the presence of additional genetic background effects controlling the onset of flowering (Figure S4). The significance of environment, region of cultivar origin and genotype effects were also confirmed by results of mixed model ANOVAs (Table S7, based on full data from 10 environments).

3.3 | Effects of *E* alleles on phenology traits

Mutant alleles at the *E1*, *E2* and *E4* loci caused an earlier flowering of about 7–8 days as compared to the respective dominant wild type alleles (Table 1), whereas for *E3*, a significant effect of mutant alleles was found for TTM and the duration of the reproductive phase only. For *E1*, the non-functional null-allele *e1-nl* causes an earlier flowering than the dominant *E1* allele by about 8.2 days on average (Table 1 and Figure 2), whereas the allele *e1-as* appears as fully functional similar to allele *E1* with respect to delaying flowering. As the interaction between the *E1*-locus and environment was particularly high for TTF (ANOVA, details not shown), that locus was investigated in more detail (see Table 2) separately for each environment. The difference between *E1* and *e1-nl* alleles in TTF was between 1.3 and 19.8 days (Table 2) depending on the environment. As the *E1* locus is sensitive to day-length, and as day-length is a function of latitude, both the functional alleles *E1* and *e1-as* were significantly correlated to latitude with respect to TTF, whereas the null-allele *e1-nl* was not correlated to day-length (Table 2). As the TTF is depending both on day-length

TABLE 1 Average effects of mutant alleles at *E1* to *E4* loci on time to flowering (TTF), time to maturity (TTM) and duration of reproductive phase (RPH) compared to their respective dominant wild type allele (difference to wild type in days, data across 17 environments).

| <i>E</i> locus allele | TTF | TTM | RPH |
|-----------------------|-------|--------|--------|
| <i>E1</i> | 0.0a | 0.0a | 0.0a |
| <i>e1-nl</i> | −8.2b | −17.3c | −9.3b |
| <i>e1-as</i> | 0.3a | −2.5b | −2.9a |
| <i>E2</i> | 0.0a | 0.0a | 0.0a |
| <i>e2-ns</i> | −7.3b | −7.5b | −0.3a |
| <i>E3</i> | 0.0a | 0.0a | 0.0a |
| <i>E3-Ha</i> | −0.4a | −7.6c | −7.4c |
| <i>e3-tr</i> | −0.7a | −4.8b | −4.2b |
| <i>E4</i> | 0.0a | 0a | 0a |
| <i>e4-kes</i> | −8.3b | −17.2b | −8.9b |
| <i>e4-SORE-1</i> | −7.0b | −18.5b | −11.5b |

Note: Different letters within each *E* locus indicate significant differences between respective alleles (Tukey–Kramer multiple comparison at $P = .05$ level).

(photoperiodic response) as well as on other environmental factors (e.g. temperature, sowing date, soil and nutritional conditions), the day-length insensitive null-allele *e1-nl* could be used as an indicator of those other environmental factors. Consequently, the difference calculated between alleles *E1* and *e1-nl* in TTF might reveal an unbiased day-length effect on flowering; actually, the correlation between latitude and that difference (allelic difference) *E1-e1nl* was $r = .897$ (Table 2) showing a high correspondence between day-length/latitude and TTF with respect to *E1*; in the plots given in Figure 3, the different reactions of *E1* (Figure 3a), *e1-nl* (Figure 3b) and their allelic difference (Figure 3c) to latitude are clearly visible. This also applies to the functional allele *e1-as*, in which the allelic difference to *e1-nl* is again highly correlated to the latitude ($r = .936$; Table 2 and Figure S5a,b). In addition, the difference between *E1*-alleles with respect to timing of flowering and its dependence on latitude is also evident from comparing distributions of individual high and low latitude environments (Figure S6a,b). For alleles at the flowering loci *E2*, *E3* and *E4*, similar correlations between latitude and TTF were found except for alleles *e4-SORE-1* and *e4-kes*, which were not correlated to latitude similar to *e1-nl* (detail results not shown).

3.4 | Adaptation to latitude

Sensitivity to day-length is a major feature for adaptation of soybean to different geographic locations. The reaction of the three different *E1* alleles to latitude as outlined above is visualized in a PCA bi-plot analysis of flowering data in Figure 4: The 17 environments are represented by PCA scores, whereas latitude and *E1* alleles and allelic differences are shown as vectors. Vector loadings for allelic differences of both *E1* and *e1-as* to *e1-nl* are pointing to the same direction as the

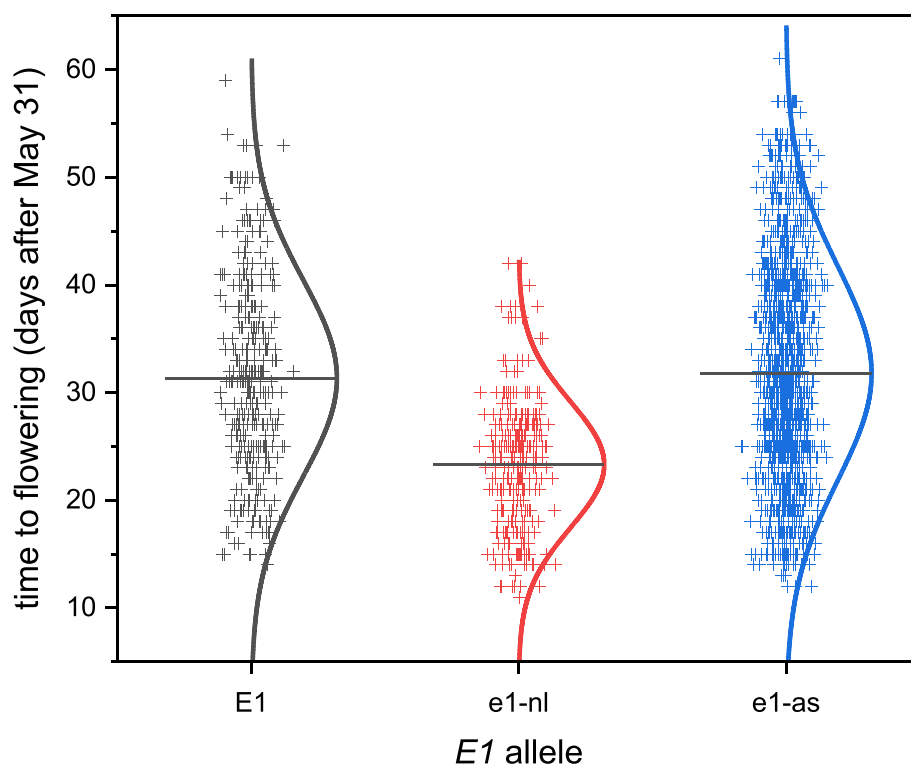


FIGURE 2 Time to flowering across all cultivars and 17 environments as affected by the *E1* allele.

TABLE 2 Time to flowering (days after May 31) for different environments with allelic effects at the *E1* locus, allelic differences between *E1* alleles, and correlations to latitude (degree N).

| Environment | | | | <i>E1</i> allele | | | Allelic difference (in days) | | |
|--------------------------------------|------------------------|------|----------|------------------|--------------|--------------|------------------------------|-------------------------|---------------------------|
| No. | Location | Year | Latitude | <i>E1</i> | <i>e1-nl</i> | <i>e1-as</i> | <i>E1</i> - <i>e1nl</i> | <i>E1</i> - <i>e1as</i> | <i>e1as</i> - <i>e1nl</i> |
| 1 | Tulln, AT | 2019 | 48.32 | 33.5 | 22.4 | 31.3 | 11.0 | 2.1 | 8.9 |
| 2 | Kühnham, DE | 2019 | 48.41 | 29.0 | 19.3 | 28.1 | 9.7 | 0.9 | 8.8 |
| 3 | Willstätt, DE | 2019 | 48.52 | 22.7 | 15.5 | 24.6 | 7.2 | -1.9 | 9.1 |
| 4 | Táplánszentkereszt, HU | 2019 | 47.25 | 36.1 | 26.9 | 34.2 | 9.1 | 1.8 | 7.3 |
| 5 | Nyon, CH | 2019 | 46.40 | 28.0 | 23.4 | 30.3 | 4.6 | -2.3 | 6.9 |
| 6 | Tulln, AT | 2020 | 48.32 | 34.1 | 26.3 | 35.1 | 7.9 | -1.0 | 8.9 |
| 7 | Kühnham, DE | 2020 | 48.41 | 38.8 | 26.7 | 39.4 | 12.1 | -0.6 | 12.7 |
| 8 | Willstätt, DE | 2020 | 48.52 | 33.7 | 27.4 | 37.1 | 6.3 | -3.4 | 9.7 |
| 9 | Táplánszentkereszt, HU | 2020 | 47.25 | 43.3 | 36.4 | 44.5 | 6.9 | -1.2 | 8.1 |
| 10 | Nyon, CH | 2020 | 46.40 | 29.8 | 26.1 | 30.2 | 3.7 | -0.4 | 4.1 |
| 11 | Dłóń, PL | 2019 | 51.69 | 36.5 | 16.7 | 36.7 | 19.8 | -0.2 | 20.0 |
| 12 | Turda, RO | 2019 | 46.59 | 20.4 | 15.6 | 19.3 | 4.8 | 1.1 | 3.8 |
| 13 | Novi Sad, SRB | 2019 | 45.34 | 16.8 | 15.4 | 16.6 | 1.3 | 0.2 | 1.2 |
| 14 | Vinnitsia, UA | 2019 | 49.23 | 30.6 | 21.4 | 32.4 | 9.1 | -1.9 | 11.0 |
| 15 | Turda, RO | 2020 | 46.59 | 33.6 | 25.4 | 33.5 | 8.2 | 0.2 | 8.1 |
| 16 | Novi Sad, SRB | 2020 | 45.34 | 25.9 | 22.8 | 24.3 | 3.1 | 1.6 | 1.5 |
| 17 | Dłóń, PL | 2020 | 51.69 | 43.1 | 28.3 | 43.2 | 14.8 | -0.2 | 14.9 |
| Average allelic difference | | | | | | | 8.2 | -0.3 | 8.5 |
| Correlation to latitude (<i>r</i>) | | | | 0.568 | 0.019 | 0.586 | 0.897 | -0.199 | 0.936 |
| Significance (<i>P</i> -value) | | | | 0.0173 | <i>n.s.</i> | 0.0134 | <0.0001 | <i>n.s.</i> | <0.0001 |

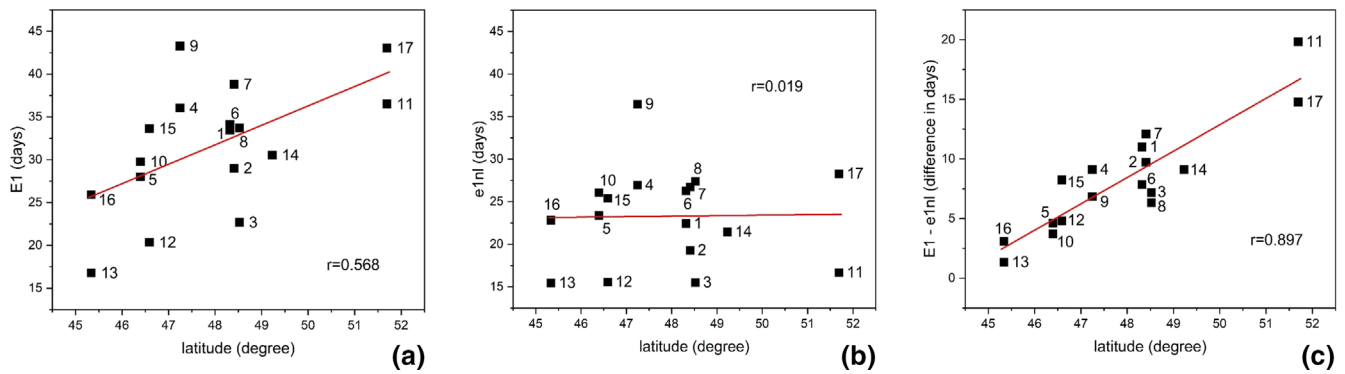
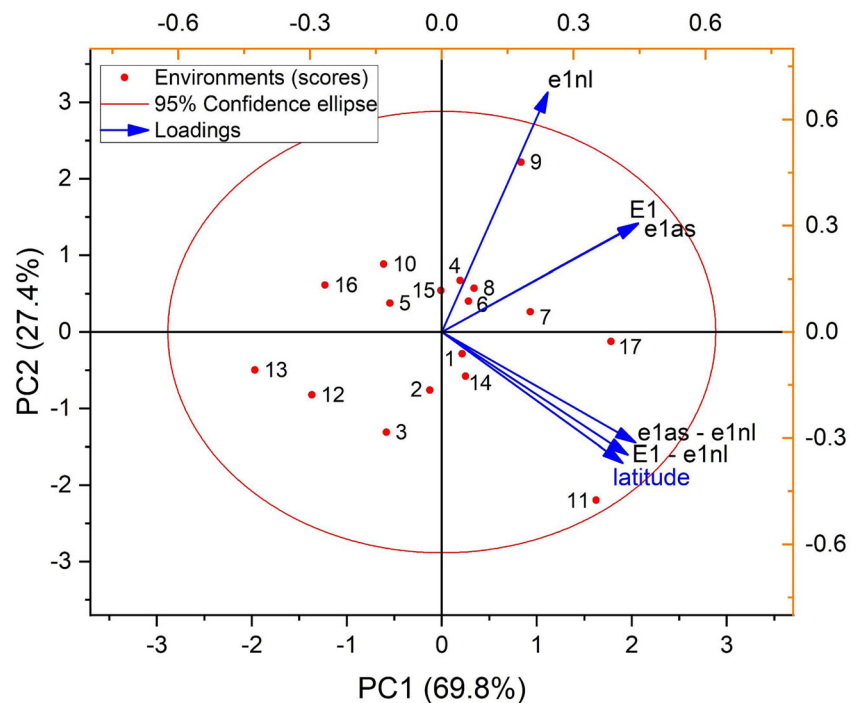


FIGURE 3 Correlation between latitude (degree N) and time to flowering for *E1* (a) and *e1-nl* (b) alleles and for the difference between *E1* and *e1-nl* alleles (c) over 17 environments (see Table 2 for environment codes).

FIGURE 4 PCA bi-plot of *E1* alleles and latitude for time to flowering in 17 environments (see Table 2 for environment codes).



latitude vector indicating their strong response to latitude, whereas the *e1-nl* vector forms an almost 90° angle to latitude, which indicates its full independence of latitude. On the level of the 16 different *E*-haplotypes, a similar PCA bi-plot (Figure 5) can be interpreted as an indicator of adaptation to latitude: Vectors of *E*-haplotypes 5, 4, 7 and 2 (late flowering *E*-haplotypes according to Figure S4) show closer correlations to latitude with respect to flowering time, whereas *E*-haplotypes 8, 15, 14 and 16 (early flowering *E*-haplotypes; Figure S4) are independent of latitude. Thus, *E*-haplotypes 5, 4, 7 and 2 are better adapted to lower latitudes, as their flowering and maturity would be much delayed by longer day-length in the north. In contrast, *E*-haplotypes 8, 15, 14 and 16 can be grown higher in the north, because their flowering is appropriately timed here. Thus, the angle between the latitude vector and the *E*-haplotype vector loading is indicating adaptation to day-length, and the larger the angle (e.g. *E*-

haplotype 8 or 15), the further north a haplotype could potentially be grown. This result is also clearly confirmed in another PCA bi-plot analysis, which includes the five most frequent *E*-haplotypes only (Figure S7).

4 | DISCUSSION

A clear differentiation between Chinese and European elite soybean populations of similar maturity has recently been disclosed on the level of different genetic marker systems (Yao et al., 2023). This differentiation is also evident with respect to allele distribution at the flowering/maturity loci *E1* to *E4* (Figure 1b). As these loci have been recognized as the most important genetic factors for soybean adaptation to regions of different latitude both in China (Liu et al., 2020) and

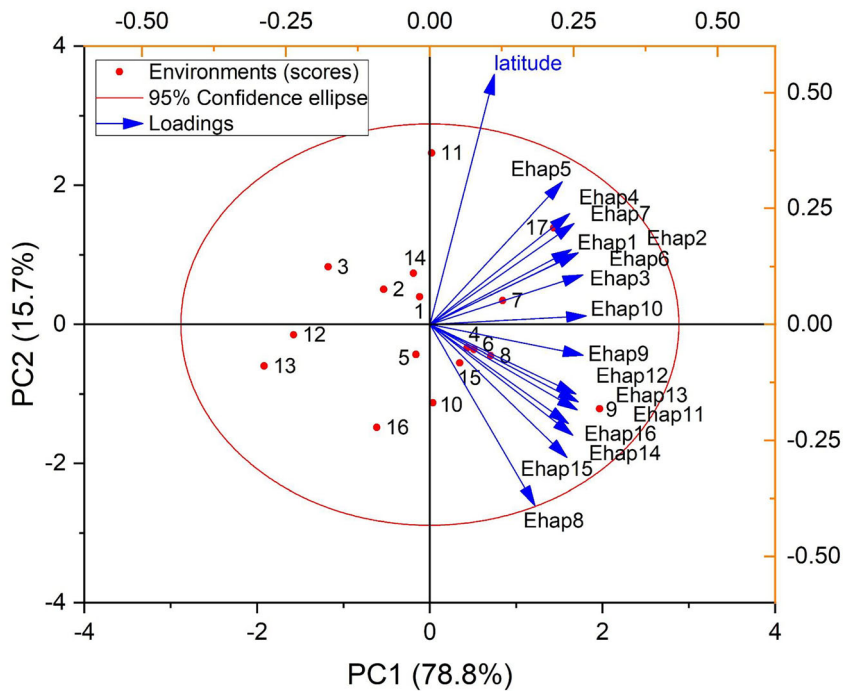


FIGURE 5 PCA bi-plot of 16 *E*-haplotypes and latitude for time to flowering in 17 environments (see Table 2 for environment codes).

in Europe (Kurasch et al., 2017), they reflect the need for adaptation to particular environments with specific photothermal conditions. Correspondingly, distinct patterns of *E*-gene differentiation due to adaptation processes of soybean have been reported for other regions (Dietz et al., 2021; Langewisch et al., 2017; Miladinović et al., 2018). Most remarkably, many European cultivars adapted to long day-length of higher latitudes are carrying the non-functional null allele *e1-nl* at the *E1* locus (Figure 1b), which is rarely present in China or other regions (Liu et al., 2020; Xia et al., 2012); due to its photoperiod-insensitivity, the *e1-nl* allele is found in early maturity groups only, while it does not play a role for late maturity soybeans (Figure 1a) or in breeding programmes for southern Europe (Miladinović et al., 2018). Similar to *e1-nl*, the dys-functional allele *e4-SORE-1* was as well more frequent in early maturity European rather than in Chinese cultivars (Figure 1a,b), whereas alleles *E1* and *e3-tr* were present in Chinese cultivars at higher frequencies (Figure 1b). As a consequence of allelic differences at the *E1* to *E4* loci, individual allelic combinations (i.e. *E*-haplotypes) were unevenly distributed between maturity groups (Table S5) and regions (Figure S2), which appears as a result of selection for specific adaptation and productivity; similarly, particular allele combinations at *E*-loci were also proposed for the different growing regions of China as well as for Europe (Liu et al., 2020; Miladinović et al., 2018). Thus, variation at loci *E1* to *E4* is considered as the main driver for soybean diversification in maturity and adaptation to different latitudes (Jiang et al., 2014).

The average effects of mutant alleles at *E1* to *E4* loci on the TTF (Table 1) are in a similar magnitude as reported by Kurasch et al. (2017), Liu et al. (2020) and Perfil'ev et al. (2023). While dominant wild-type alleles at *E1*, *E2* and *E4* delay flowering by 7–8 days on average, this is not the case for *E3*, where mutant alleles *E3-Ha* and the truncated *e3-tr* affect TTM only. This post-flowering effect of

mutant *E3* alleles reduces the duration of the reproductive stage and the number of mainstem nodes bearing pods (Xu et al., 2013); such an effect on accelerating maturity can be relevant in certain northern growing regions to avoid killing frost at the end of season but might also limit grain yield performance.

Different light conditions (i.e. red/far-red light ratios due to different day-length) are affecting the expression of phytochrome A genes, which regulates the flowering inhibition effect of the *E1* locus both on a transcriptional and posttranscriptional level (Lin et al., 2022). Therefore, differences in TTF between the photoperiod-sensitive dominant (*E1*) and the photoperiod-insensitive recessive (*e1-nl*) allele are larger at higher latitudes than at lower ones (Figure S6 and Table 2). This has earlier been reported for controlled day-length conditions (Cober et al., 1996) as well as for European (Kurasch et al., 2017) and Chinese locations (Liu et al., 2020) with respect to *E1* and other *E*-genes. As the diurnal expression of the flowering-inhibitory effect mediated by day-length appears to be of a quantitative nature, the present data set is suitable to study the effect of day-length on different *E* alleles as mediated by geographic latitude. For the locus *E1*, the recessive allele *e1-nl* is not correlated to latitude in terms of TTF (Figure 3b), which clearly demonstrates its photoperiod-insensitive nature as a loss-of-function (130 kb deletion) allele. In contrast, the allele *E1* (Figure 3a) and the difference between *E1* and *e1-nl* (Figure 3c) are clearly correlated to latitude (i.e. day-length). This applies to the missense allele *e1-as* as well, which appears to be fully functional and similar to the dominant allele *E1* (Figure S5) in terms of inhibiting early flowering with a minor reduction in TTM and reproductive phase only (Table 1) based on the present data set collected over a wide range of environments. In contrast, *e1-as* has previously been described as hypo-active (Tripathi et al., 2021), semi-functional (Perfil'ev et al., 2023), hypomorphic (Xu et al., 2013) or not fully functional and photoperiod insensitive

(Miladinović et al., 2018), which might be due to genetic background or specific environmental effects. In addition, the *e1-as* allele might be associated with high yield capacity under growing conditions of Central-South-Eastern Europe, as it has predominantly been selected for in respective breeding programmes targeting these regions (Miladinović et al., 2018). Allelic differences in TTF between functional and loss-of-function alleles, that is, between *E1* and *e1-nl* or between *e1-as* and *e1-nl* (Table 2 and Figures 3c and 55b) appear highly useful, as they can offer to observe a pure latitude/day-length effect on flowering independent of any other environmental influences.

For the *E1* locus, the adaptability of the different alleles is well illustrated in a PCA bi-plot graph of TTF data and latitude for all locations (see Figure 4). Alleles *E1* and *e1-as* have very similar vectors pointing in the same direction, whereas the vector for allele *e1-nl* forms an almost 90°-angle to the latitude-vector indicating no correlation to latitude and thus no sensitivity to day-length; contrastingly, the differences between both *E1* and *e1-as* and the null allele *e1-nl* are closely correlated to latitude. When this view is extended from one *E*-locus to all 16 *E*-haplotypes (Figure 5), the adaptation potential of genotypes can be visualized. Photoperiod-insensitive *E*-haplotypes such as the *E*-haplotype 8, 15, 14, 16 and 9 carry two to three non-functional *E*-alleles that enable a timely flowering even under long day-lengths of northern latitudes, whereas late maturity *E*-haplotypes 5, 4, 7 and 2 carry no or only one non-functional *E*-allele that prohibits cultivation of the respective genotypes in higher latitudes. This view on adaptation through *E*-haplotypes is supported by recent findings on adaptation to specific environments of China (Liu et al., 2020), Europe (Kurasch et al., 2017; Miladinović et al., 2018) or North America (Goettel & An, 2017). Through the consideration of additional QTL identified for thermal response, soybean genotypes could be selected with an optimum photothermal adaptation to specific environments or cropping systems (Wu et al., 2023). Moreover, the present results also suggest that developing new *E*-haplotypes by combining unique Chinese and European ones (see Figure S2) would probably enhance the adaptation potential due to a further differentiation in photosensitivity; as an example, early-flowering Chinese *E*-haplotype 8 might be combined with European *E*-haplotypes 14 and 15 for producing new photoperiod-insensitive genotypes suitable for planting in high latitudes.

The timing of flowering and maturity in soybean is a complex developmental process influenced by regulatory genetic networks controlling phenology and plant architecture (i.e. stem termination pattern) as well as by environmental conditions such as day-length and temperature. While the present results demonstrate that molecular breeding approaches are most helpful in the search for adaptation, there is considerable variation in TTF within some of the *E*-haplotypes identified here (Figure S4; Table S6: significant region within *E*-haplotype effect) indicating genetic background effects. Therefore, the consideration of additional *E*-gene alleles and further QTL for TTF and photothermal response as identified recently (see Fu et al., 2020; Liu et al., 2021; Zhang et al., 2017; Zhu et al., 2023) would promise a more precise prediction of adaptation of individual breeding lines to different environments as well as a better categorization into maturity groups. While the present study has mainly focused on the TTF, *E*-

gene and latitude effects were similarly present for TTM, which is of high relevance for farming at given locations. In addition, many agronomic characters such as number of nodes and pods, branching, plant height, duration of seed filling, seed size, maturity and finally grain yield formation are indirectly affected by photoperiod responses (Dong et al., 2023; Ort et al., 2022; Xu et al., 2013). Thus, precise adaptation to given cropping environments might assist in a more effective mitigation of climate change effects through the expansion of present soybean growing areas to northern latitudes (Vollmann & Škrabišová, 2023) as well as through changes in cropping patterns such as early or late (i.e. second crop) planting in soybean production.

AUTHOR CONTRIBUTIONS

Johann Vollmann, Volker Hahn, Leopold Rittler and Li-Juan Qiu developed the research concept of this study. Leopold Rittler and Xindong Yao cared for seed logistics and sample collection. Christine Riedel, Raluca Rezi, Claude-Alain Bétrix, Jerzy Nawracała, Inna Temchenko, Vuk Đorđević, Volker Hahn and Martin Pachner conducted field experiments. Willmar Leiser carried out the molecular genetic analysis. Xindong Yao, Martin Pachner and Johann Vollmann collected and analysed experimental data. Johann Vollmann and Xindong Yao wrote the manuscript and received input from all other authors.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support provided by the Friedrich Haberlandt Scholarship committee and administrative help by Donau Soja. The authors are also grateful to soybean breeding companies and institutions from both Europe and China for providing seed of their recent soybean cultivars included in the present study. We acknowledge the contributions by Mrs. Timea Jakab of Gabonakutató Nonprof. Közhasznú Kft., Hungary in conducting field experiments and data collection. Helpful advice provided by Mr. Matthias Krön (Vienna, Austria) and Dr Donal Murphy-Bokern (Lohne, Germany) is also gratefully acknowledged.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest related to the subject of the research presented.

DATA AVAILABILITY STATEMENT

E-gene data for all soybean cultivars are included among the supplementary materials of this publication. Phenotypic data from individual experiments are available from the corresponding authors on request.

ORCID

Xindong Yao  <https://orcid.org/0000-0002-2364-4226>

Johann Vollmann  <https://orcid.org/0000-0003-2057-8347>

REFERENCES

- Cober, E. R., Tanner, J. W., & Voldeng, H. D. (1996). Genetic control of photoperiod response in early-maturing, near-isogenic soybean lines. *Crop Science*, 36(3), 601–605. <https://doi.org/10.2135/cropsci1996.0011183X003600030013x>

- Dietz, N., Combs-Giroir, R., Cooper, G., Stacey, M., Miranda, C., & Bilyeu, K. (2021). Geographic distribution of the E1 family of genes and their effects on reproductive timing in soybean. *BMC Plant Biology*, 21(1), 441. <https://doi.org/10.1186/s12870-021-03197-x>
- Dong, L., Li, S., Wang, L., Su, T., Zhang, C., Bi, Y., Lai, Y., Kong, L., Wang, F., Pei, X., Li, H., Hou, Z., Du, H., Li, T., Cheng, Q., Fang, C., Kong, F., & Liu, B. (2023). The genetic basis of high-latitude adaptation in wild soybean. *Current Biology*, 33(2), 252–262.e4. <https://doi.org/10.1016/j.cub.2022.11.061>
- Fang, C., Du, H., Wang, L., Liu, B., & Kong, F. (2024). Mechanisms underlying key agronomic traits and implications for molecular breeding in soybean. *Journal of Genetics and Genomics*, 51(4), 379–393. <https://doi.org/10.1016/j.jgg.2023.09.004>
- Fu, M., Wang, Y., Ren, H., Du, W., Wang, D., Bao, R., Yang, X., Tian, Z., Fu, L., Cheng, Y., Su, J., Sun, B., Zhao, J., & Gai, J. (2020). Genetic dynamics of earlier maturity group emergence in south-to-north extension of northeast China soybeans. *Theoretical and Applied Genetics*, 133(6), 1839–1857. <https://doi.org/10.1007/s00122-020-03558-4>
- Goettel, W., & An, Y. C. (2017). Genetic separation of southern and northern soybean breeding programs in North America and their associated allelic variation at four maturity loci. *Molecular Breeding*, 37(1), 8. <https://doi.org/10.1007/s11032-016-0611-7>
- Hou, Z., Fang, C., Liu, B., Yang, H., & Kong, F. (2023). Origin, variation, and selection of natural alleles controlling flowering and adaptation in wild and cultivated soybean. *Molecular Breeding*, 43(5), 36. <https://doi.org/10.1007/s11032-023-01382-4>
- Jiang, B., Nan, H., Gao, Y., Tang, L., Yue, Y., Lu, S., Ma, L., Cao, D., Sun, S., Wang, J., Wu, C., Yuan, X., Hou, W., Kong, F., Han, T., & Liu, B. (2014). Allelic combinations of soybean maturity loci E1, E2, E3 and E4 result in diversity of maturity and adaptation to different latitudes. *PLoS ONE*, 9(8), e106042. <https://doi.org/10.1371/journal.pone.0106042>
- Kurasch, A. K., Hahn, V., Leiser, W. L., Vollmann, J., Schori, A., Béatrix, C., Mayr, B., Winkler, J., Mechtler, K., Aper, J., Sudaric, A., Pejic, I., Sarcevic, H., Jeanson, P., Balko, C., Signor, M., Miceli, F., Strijk, P., Rietman, H., ... Würschum, T. (2017). Identification of mega-environments in Europe and effect of allelic variation at maturity E loci on adaptation of European soybean. *Plant, Cell & Environment*, 40(5), 765–778. <https://doi.org/10.1111/pce.12896>
- Langewisch, T., Lenis, J., Jiang, G.-L., Wang, D., Pantalone, V., & Bilyeu, K. (2017). The development and use of a molecular model for soybean maturity groups. *BMC Plant Biology*, 17(1), 91. <https://doi.org/10.1186/s12870-017-1040-4>
- Lin, X., Dong, L., Tang, Y., Li, H., Cheng, Q., Li, H., Zhang, T., Ma, L., Xiang, H., Chen, L., Nan, H., Fang, C., Lu, S., Li, J., Liu, B., & Kong, F. (2022). Novel and multifaceted regulations of photoperiodic flowering by phytochrome A in soybean. *Proceedings of the National Academy of Sciences*, 119(41), e2208708119. <https://doi.org/10.1073/pnas.2208708119>
- Liu, F., He, J., Wang, W., Xing, G., Zhao, J., Li, Y., & Gai, J. (2021). Genetic dynamics of flowering date evolved from later to earlier in annual wild and cultivated soybean in China. *Crop Science*, 61(4), 2336–2354. <https://doi.org/10.1002/csc2.20462>
- Liu, L., Song, W., Wang, L., Sun, X., Qi, Y., Wu, T., Sun, S., Jiang, B., Wu, C., Hou, W., Ni, Z., & Han, T. (2020). Allele combinations of maturity genes E1–E4 affect adaptation of soybean to diverse geographic regions and farming systems in China. *PLoS ONE*, 15(7), e0235397. <https://doi.org/10.1371/journal.pone.0235397>
- Miladinović, J., Čeran, M., Đorđević, V., Balešević-Tubić, S., Petrović, K., Đukić, V., & Miladinović, D. (2018). Allelic variation and distribution of the major maturity genes in different soybean collections. *Frontiers in Plant Science*, 9, 1286. <https://doi.org/10.3389/fpls.2018.01286>
- Nendel, C., Reckling, M., Debaeke, P., Schulz, S., Berg-Mohnicke, M., Constantin, J., Fronzek, S., Hoffmann, M., Jakšić, S., Kersebaum, K.-C., Klimek-Kopyra, A., Raynal, H., Schoving, C., Stella, T., & Battisti, R. (2023). Future area expansion outweighs increasing drought risk for soybean in Europe. *Global Change Biology*, 29(5), 1340–1358. <https://doi.org/10.1111/gcb.16562>
- Niedbala, G., Kurasiak-Popowska, D., Piekutowska, M., Wojciechowski, T., Kwiatek, M., & Nawracała, J. (2022). Application of artificial neural network sensitivity analysis to identify key determinants of harvesting date and yield of soybean (*Glycine max* [L.] Merrill) cultivar Augusta. *Agriculture*, 12(6), 754. <https://doi.org/10.3390/agriculture12060754>
- Ort, N. W. W., Morrison, M. J., Cober, E. R., McAndrew, D., & Lawley, Y. E. (2022). A comparison of soybean maturity groups for phenology, seed yield, and seed quality components between eastern Ontario and southern Manitoba. *Canadian Journal of Plant Science*, 102(4), 812–822. <https://doi.org/10.1139/cjps-2021-0235>
- Perfil'ev, R., Shcherban, A., Potapov, D., Maksimenko, K., Kiryukhin, S., Gurinovich, S., Panarina, V., Polyudina, R., & Salina, E. (2023). Impact of allelic variation in maturity genes E1–E4 on soybean adaptation to central and west Siberian regions of Russia. *Agriculture*, 13(6), 1251. <https://doi.org/10.3390/agriculture13061251>
- Rossi, M. E., Orf, J. H., Liu, L.-J., Dong, Z., & Rajcan, I. (2013). Genetic basis of soybean adaptation to North American vs. Asian mega-environments in two independent populations from Canadian × Chinese crosses. *Theoretical and Applied Genetics*, 126(7), 1809–1823. <https://doi.org/10.1007/s00122-013-2094-9>
- Song, W., Liu, L., Sun, S., Wu, T., Zeng, H., Tian, S., Sun, B., Li, W., Liu, L., Wang, S., Xing, H., Zhou, X.-a., Nian, H., Lu, W., Han, X., Wang, S., Chen, W., Guo, T., Song, X., ... Han, T. (2023). Precise classification and regional delineation of maturity groups in soybean cultivars across China. *European Journal of Agronomy*, 151, 126982. <https://doi.org/10.1016/j.eja.2023.126982>
- Tripathi, R., Agrawal, N., Kumawat, G., Gupta, S., Kuchlan, M., Maranna, S., Nataraj, V., Kuchlan, P., Satpute, G. K., Ratnaparkhe, M. B., Rajesh, V., Chandra, S., Bhatia, V. S., & Chand, S. (2021). Novel role of photoin-sensitive alleles in adaptation of soybean [*Glycine max* (L.) Merr.] to rainfed short growing seasons of lower latitudes. *Genetic Resources and Crop Evolution*, 68(6), 2455–2467. <https://doi.org/10.1007/s10722-021-01142-3>
- Tsubokura, Y., Watanabe, S., Xia, Z., Kanamori, H., Yamagata, H., Kaga, A., Katayose, Y., Abe, J., Ishimoto, M., & Harada, K. (2014). Natural variation in the genes responsible for maturity loci E1, E2, E3 and E4 in soybean. *Annals of Botany*, 113(3), 429–441. <https://doi.org/10.1093/aob/mct269>
- Vollmann, J., & Škrabišová, M. (2023). Going north: Adaptation of soybean to long-day environments. *Journal of Experimental Botany*, 74(10), 2933–2936. <https://doi.org/10.1093/jxb/erad105>
- Wu, T., Lu, S., Cai, Y., Xu, X., Zhang, L., Chen, F., Jiang, B., Zhang, H., Sun, S., Zhai, H., Zhao, L., Xia, Z., Hou, W., Kong, F., & Han, T. (2023). Molecular breeding for improvement of photothermal adaptability in soybean. *Molecular Breeding*, 43(8), 60. <https://doi.org/10.1007/s11032-023-01406-z>
- Xia, Z., Zhai, H., Liu, B., Kong, F., Yuan, X., Wu, H., Cober, E. R., & Harada, K. (2012). Molecular identification of genes controlling flowering time, maturity, and photoperiod response in soybean. *Plant Systematics and Evolution*, 298(7), 1217–1227. <https://doi.org/10.1007/s00606-012-0628-2>
- Xu, M., Xu, Z., Liu, B., Kong, F., Tsubokura, Y., Watanabe, S., Xia, Z., Harada, K., Kanazawa, A., Yamada, T., & Abe, J. (2013). Genetic variation in four maturity genes affects photoperiod insensitivity and PHYA-regulated post-flowering responses of soybean. *BMC Plant Biology*, 13(1), 91. <https://doi.org/10.1186/1471-2229-13-91>
- Yao, X., Xu, J., Liu, Z., Pachner, M., Molin, E. M., Rittler, L., Hahn, V., Leiser, W., Gu, Y.-z., Lu, Y.-q., Qiu, L.-j., & Vollmann, J. (2023). Genetic diversity in early maturity Chinese and European elite soybeans: A comparative analysis. *Euphytica*, 219(1), 17. <https://doi.org/10.1007/s10681-022-03147-0>

- Zhang, S.-R., Wang, H., Wang, Z., Ren, Y., Niu, L., Liu, J., & Liu, B. (2017). Photoperiodism dynamics during the domestication and improvement of soybean. *Science China Life Sciences*, 60(12), 1416–1427. <https://doi.org/10.1007/s11427-016-9154-x>
- Zhu, X., Leiser, W. L., Hahn, V., & Würschum, T. (2023). The genetic architecture of soybean photothermal adaptation to high latitudes. *Journal of Experimental Botany*, 74(10), 2987–3002. <https://doi.org/10.1093/jxb/erad064>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Yao, X., Pachner, M., Rittler, L., Hahn, V., Leiser, W., Riedel, C., Rezi, R., Béatrix, C.-A., Nawracata, J., Temchenko, I., Đorđević, V., Qiu, L.-J., & Vollmann, J. (2024). Genetic adaptation of phenological stages in Chinese and European elite soybeans (*Glycine max* [L.] Merr.) across latitudes in Central Europe. *Plant Breeding*, 1–11. <https://doi.org/10.1111/pbr.13197>