Please cite this article in press as: Yue et al., A Single Nucleotide Deletion in *J* Encoding GmELF3 Confers Long Juvenility and Is Associated with Adaption of Tropic Soybean, Molecular Plant (2017), http://dx.doi.org/10.1016/j.molp.2016.12.004

Molecular Plant
Letter to the Editor

CellPress
PARTNER JOURNAL

A Single Nucleotide Deletion in *J* Encoding GmELF3 Confers Long Juvenility and Is Associated with Adaption of Tropic Soybean

Dear Editor,

Wild soybean is a typical short-day plant that begins flowering when the days are shorter than its critical photoperiod. Soybean was domesticated in the temperate region of East Asia at the relatively high latitude, and the breeding and release of soybean varieties have historically centered on mid- and high-latitude temperate regions. Low-latitude areas with tropical and sub-tropical climates were previously considered unsuitable for soybean production because most temperate soybean varieties exhibited precocious flowering and early maturity and suffered from low yields.

The discovery and introduction of the long juvenile trait into soybean varieties in the 1970s (Hartwig and Kiihl, 1979) fundamentally changed global soybean production in a way that has had an enormous influence on commodity markets. This trait delays flowering and thereby ensures sufficient vegetative growth prior to the developmental transition to reproductive growth. The long juvenile trait thus solved the early maturation and low yield problems that had hitherto prevented economically viable soybean production in low-latitude regions (Destro et al., 2001). The United States and Brazil pioneered the introduction of the long juvenile trait in low-latitude soybean breeding programs. Brazil has expanded its soybean production enormously, from 1 million hectares in 1970 (Brown, 2004) to over 33 million hectares in 2016 (http://gain.fas.usda.gov/Recent%20GAIN%20Publications/Oilseeds%20and%20Products%20Update_Brasilia_Brazil_12-1-2016.pdf).

The existence of a *J*/*j* locus controlling the long juvenile trait in tropic soybean was proposed through genetic studies nearly 40 years ago, but its genetic basis has not been identified. Long juvenility is known to be a recessive trait (Ray et al., 1995; Carpentieri-Pípolo et al., 2000), and some geneticists have asserted that it is likely controlled by a single gene (Ray et al., 1995). However, others including Hartwig and Kiihl (1979) have suggested that several genes are likely involved in controlling this trait. To date, however, the gene or quantitative trait locus (QTL) responsible for *J* has not been reported, even though several other maturity loci have been mapped and identified (Xia et al., 2012).

The long juvenile trait is only expressed in tropical soybeans. Guangzhou, China (23.17° N, 113.36° E) is uniquely located at the boundary of the sub-tropical and tropical climates, making it especially suitable for phenotyping the long juvenile trait: day lengths in the soybean growing season range from 10.8 to 13.5 h, and average temperatures are low in spring but very high in summer (Supplemental Figure 1A). Growth in Guangzhou therefore allows the screening of long juvenile varieties without the phenotype masking that occurs in the long-day conditions at high-latitude sites (e.g., Beijing, 39.97° N, 116.33° E). We thus investigated the flowering time of six soybean varieties in Guangzhou (Supplemental Table 1). We planted seeds of each variety once each month from March to September, and found, for all of the date-of-planting experiments, that the flowering times of the conventional juvenile varieties Zhonghuang 24, Dongnong 42, and AGS292 were less than 31 days, while the flowering times of the long juvenile varieties Huaxia 3, Conquista, and Vencedora were more than 35 days (Supplemental Figure 1B).

We constructed an RIL population by crossing the conventional juvenile variety Zhonghuang 24 (Huang-Huai-Hai Rivers Valley [HHH]/North China) and the long juvenile variety Huaxia 3 (South China), which consistently had the longest flowering times in our experiments in Guangzhou (Supplemental Figure 1B). The F2 population derived from the Zhonghuang 24 × Huaxia 3 cross consisted of 266 lines, and the flowering time phenotypes segregated at the expected Mendelian ratio; almost three quarters of the F2 lines flowered earlier than the conventional juvenile parent (Supplemental Figure 2 and Supplemental Table 2). The F2 population exhibited a bimodal distribution for flowering time. High-generation RIL populations (F8 and F10) planted over two seasons (spring 2012 and summer 2014) also had solid bimodal distributions for flowering time (Figure 1A, Supplemental Figure 3B and Supplemental Table 3). These results indicate that the long juvenile trait is controlled by a major recessive allele in this population.

Restriction site-associated DNA sequencing (RAD-seq) (median depth, 0.24×) of the F8 population (Supplemental Table 4) identified 52 478 high-quality single nucleotide polymorphisms (SNPs) (Supplemental Figure 4), 50% of which covered a majority of the RIL lines (Supplemental Figure 5). These SNPs were integrated into 1586 recombination bin units using the MPR package in R to construct a high-density bin linkage map (2513.8 cM) with an average distance of 1.6 cM between adjacent markers (Supplemental Table 5 and Supplemental Figure 6).

Linkage analysis with this map and the phenotype data for flowering time in spring 2012, summer 2012, and summer 2014 led to the identification of two major QTLs on chromosome 4 that had phenotypic effects as high as 51.0% (Figure 1B and

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and IPPE, SIBS, CAS.

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Please cite this article in press as: Yue et al., A Single Nucleotide Deletion in *J* Encoding GmELF3 Confers Long Juvenility and Is Associated with Adaption of Tropic Soybean, Molecular Plant (2017), http://dx.doi.org/10.1016/j.molp.2016.12.004

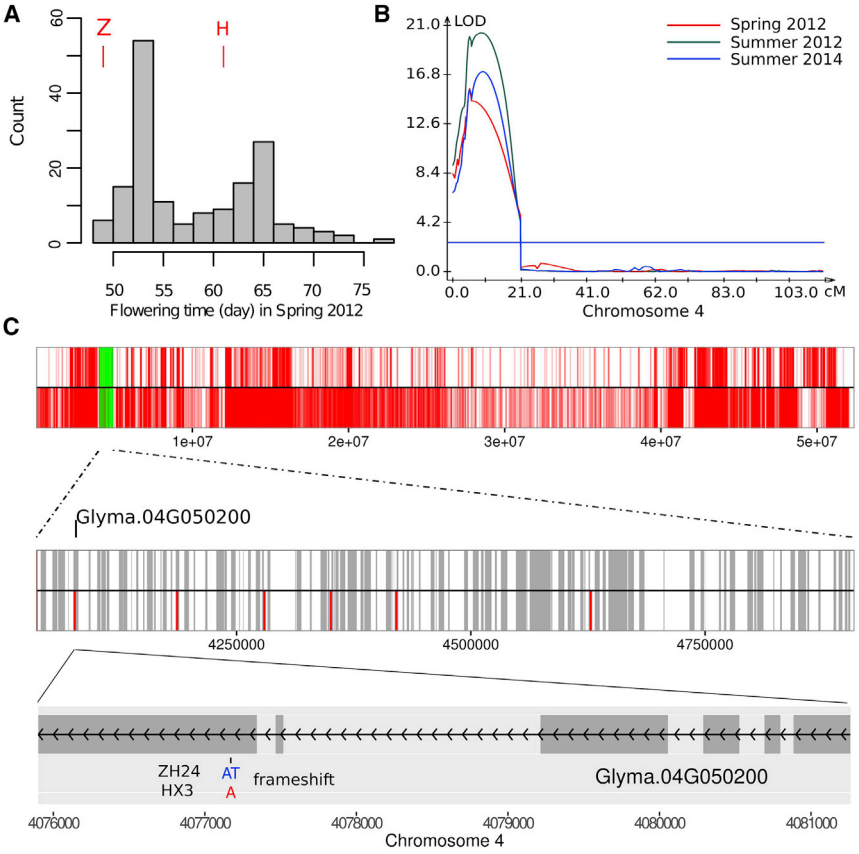

Figure 1. A Polymorphism in Glyma.04G050200 Confers the Long Juvenile Trait.

(A) Flowering time (days) for high-generation Zhonghuang 24 × Huaxia 3 populations exhibit a bimodal distribution (spring 2012). Z, Zhonghuang 24. H, Huaxia 3.

(B) Chromosome 4 harbors a QTL with phenotypic effects that was identified via QTL mapping of the Zhonghuang 24 × Huaxia 3 RIL population with the flowering time data for three seasons (spring 2012, summer 2012, and summer 2014).

(C) Whole-genome sequencing of the parental genotypes and SnpEff analysis indicated that only one polymorphism strongly affected a gene in the QTL region (an indel at Chr04:4077171 AT>A at the Glyma.04G050200). Top, distribution of polymorphisms on chromosome 4 identified via RAD-seq of the Zhonghuang 24 × Huaxia 3 RIL population (upper layer) and that of polymorphisms identified via high-depth whole-genome re-sequencing of Zhonghuang 24 and Huaxia 3 (lower layer). The green bar indicates the position of the identified QTL. Middle, distribution of polymorphisms in the vicinity of the identified QTL. Gray bars indicate annotated genes. Bottom, an AT>A indel that results in a frameshift in Glyma.04G050200. Red lines, polymorphisms. ZH24, Zhonghuang 24. HX3, Huaxia 3.

Supplemental Table 6): Chr04_bin11-12-13 (Chr04:3513908-4909316, data for spring 2012) and Chr04_bin13-13-13 (Chr04:4036172-4909316, data for the summers of 2012 and 2014). These QTLs overlapped, and were therefore merged as a single, major QTL (approximately 900 kb) for the long juvenile trait. Note that this QTL is unrelated to any of the previously reported maturity loci (e.g., E1–E4) (Xia et al., 2012).

To increase SNP density, we performed high-depth, whole-genome re-sequencing of the parents of the RILs, Zhonghuang 24 and Huaxia 3 (Supplemental Table 7). A total of 1 396 541 SNPs, 224 820 indels, and 788 polymorphisms of other types were identified between the two parents (a 25.6-fold increase in marker density over the RAD-seq data) (Figure 1C and Supplemental Figure 4). The re-sequencing of the parents covered virtually all (>99.9%) of the QTL region (Supplemental Figure 7) and led to the identification of seven SNPs and one indel in the QTL region for the long juvenile trait that differed between the two parents (Figure 1C). SnpEff analysis indicated that only one polymorphism strongly affected a gene (an indel at Chr04:4077171 AT>A at the Glyma.04G050200 locus).

The gene at this locus, which is 5360 bp in length and contains six exons and five introns, is a homolog of *EARLY FLOWERING 3* (*ELF3*) of *Arabidopsis* (Zagotta et al., 1996). The AT>A indel results in a truncated *GmELF3* gene product coding for a peptide of 443 amino acids, which is much shorter than the 714 amino acids of the original, ostensibly functional GmELF3 protein (Supplemental Figure 8). Given that the rice homolog *OsELF3* is known to promote flowering (Yang et al., 2013), it is reasonable to assume this gene likely promotes flowering in soybean. Hence, we refer to the functional allele of *GmELF3* as *J*, and refer to the allele encoding the truncated form of GmELF3 as *j*. Consistently, when we genotyped the RIL population via sequencing a 630 bp region of *J* that spans the indel, we found that the RIL lines with the *j* allele flowered later than those with the *J* allele in most instances; this trend was evident across three separate growing seasons (Supplemental Figure 9 and Supplemental Table 8), indicating that this indel co-segregates with the long juvenile trait. Phenotypic analysis of transgenic Huaxia 3 plants constitutively expressing a functional *J* allele from Zhonghuang 24 revealed that these transgenic plants flowered significantly earlier than untransformed Huaxia 3 plants (Supplemental Figure 10).

We cloned and Sanger sequenced the full-length *GmELF3* in 170 varieties from Brazil, China, and the United States (Supplemental Table 9), and found eight polymorphisms in the open reading frame (Supplemental Figure 11A and Supplemental Table 10). Based on these polymorphisms, we divided *GmELF3* loci into four haplotypes: HT1, HT1m, HT2, and HT3 (Supplemental Figure 11B). HT1, HT2, and HT3 were found in the US varieties while all four haplotypes were found in Brazilian and Chinese varieties. In China, HT1-type varieties were mostly from the northeast, where soybeans require a functional GmELF3 to flower and mature normally under long-day conditions; HT2 and HT3 types were mainly from the HHH region and in South China while HT1m type was only from South China (Supplemental Figure 11C) where soybean varieties require either a function-attenuation or

Please cite this article in press as: Yue et al., A Single Nucleotide Deletion in *J* Encoding GmELF3 Confers Long Juvenility and Is Associated with Adaption of Tropic Soybean, Molecular Plant (2017), http://dx.doi.org/10.1016/j.molp.2016.12.004

loss-of-function allele of *GmELF3* to ensure enough vegetative growth. Interestingly, one long juvenile variety Conquista is HT1 type, indicating that the long juvenility trait of Conquista may be controlled by other gene(s). In the previous studies, two or even more independent unknown genes were proposed for the genetic basis of the long juvenality in soybean (Ray et al., 1995; Carpentieri-Pípolo et al., 2000).

Taken together, our results suggest that *J* is the dominant functional allele of *GmELF3* that promotes flowering, while *j* is the recessive, loss-of-function allele that causes delayed flowering and confers the long juvenile trait that is associated with the adaption of soybean for growth in tropic, low-altitude regions. These findings fulfill the criteria of the long-predicted recessive and single-locus genetic mechanism underlying the long juvenile trait. The identification of *GmELF3* as the causal gene for the *J* locus raises several possibilities in practical applications. It is conceivable that genome-editing tools such as CRISPR/Cas9 can be used to alter the functional status of *J* in high-yielding, temperate soybean varieties and thus enable the growth of these varieties with high agronomic performance in tropical agriculture. This *J*-based strategy may have dramatic impacts on both plant yields and soybean economics. Another potential application relating to *J* could be described as region-specific flowering time optimization. Since day lengths are entirely predictable from one year to the next, one can envision highly optimized varieties for particular regions and perhaps even for particular seasons within those regions. Increased control over flowering time in soybean promises to offer new, potentially yield-increasing management options to producers. Our identification of *GmELF3* as the causal gene for the economically and historically important long juvenile trait will likely enable significant, rapid progress in soybean breeding programs. Further, our insights about *J* promise to facilitate the investigation and identification of various interacting components in the signaling pathways associated with the long juvenile trait in soybean, and perhaps photoperiodism in other plants.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Molecular Plant Online*.

FUNDING

This work was supported by the China Agricultural Research System (CARS-04) and the Major Science and Technology Projects for New Varieties of Genetically Modified Organisms (2014ZX08004-002).

AUTHOR CONTRIBUTIONS

H.N. and T.H. managed and organized the project; H.N., Y.Y., T.H., and B.J. designed the experiments and led the data analysis; Y.Y., N.L., H.W., M.L., Z.J., and Q.X. carried out the phenotypic analyses; Y.Y., B.J., Z.J., H.P., and Q.M. carried out the genetic transformations and the gene functional analyses; Y.Y., B.J., T.H., and H.N. wrote the manuscript.

ACKNOWLEDGMENTS

We thank Dr. John H. Snyder for discussion, comments and language improvement, and Dr. Cunxiang Wu for providing soybean germplasm. No conflict of interest declared.

Received: September 2, 2016
Revised: November 10, 2016
Accepted: December 6, 2016
Published: December 12, 2016

Yanlei Yue[1,2,3,6], Nianxi Liu[1,3,6], Bingjun Jiang[2,6], Mu Li[1,3], Haijie Wang[1,4], Ze Jiang[1,3], Huanting Pan[1,3], Qiuju Xia[5], Qibin Ma[1,3], Tianfu Han[2,] and Hai Nian[1,3,*]*

[1]State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou 510642, China
[2]MOA Key Lab of Soybean Biology (Beijing), Institute of Crop Science, The Chinese Academy of Agricultural Sciences, Beijing 100081, China
[3]Guangdong Provincial Key Laboratory of Plant Molecular Breeding, South China Agricultural University, Guangzhou 510642, China
[4]Institute of Crop Sciences, Hainan Academy of Agricultural Sciences, Haikou 571100, China
[5]BGI-Shenzhen, Shenzhen 518083, China
[6]These authors contributed equally to this article.
***Correspondence: Tianfu Han (hantianfu@caas.cn), Hai Nian (hnian@scau.edu.cn)**
http://dx.doi.org/10.1016/j.molp.2016.12.004

REFERENCES

Brown, L.R. (2004). Outgrowing the Earth: The Food Security Challenge in an Age of Falling Water Tables and Rising Temperatures (New York: W.W. Norton).

Carpentieri-Pípolo, V., Almeida, L.A., Kiihl, R.A.S., and Rosolem, C.A. (2000). Inheritance of long juvenile period under short day conditions for the BR80-6778 soybean (*Glycine max* (L.) Merrill) line. Euphytica **112**:203–209.

Destro, D., Carpentieri-Pípolo, V., Kiihl, R.A.S., and Almeida, L.A. (2001). Photoperiodism and genetic control of the long juvenile period in soybean: a review. Crop Breed. Appl. Biotechnol. **1**:72–92.

Hartwig, E.E., and Kiihl, R.A.S. (1979). Identification and utilization of a delayed flowering character in soybeans for short-day conditions. Field Crops Res. **2**:145–151.

Ray, J.D., Hinson, K., Mankono, J.E.B., and Malo, M.F. (1995). Genetic control of a long-juvenile trait in soybean. Crop Sci. **35**:1001–1006.

Xia, Z., Zhai, H., Liu, B., Kong, F., Yuan, X., Wu, H., Cober, E.R., and Harada, K. (2012). Molecular identification of genes controlling flowering time, maturity, and photoperiod response in soybean. Plant Syst. Evol. **298**:1217–1227.

Yang, Y., Peng, Q., Chen, G.X., Li, X.H., and Wu, C.Y. (2013). OsELF3 is involved in circadian clock regulation for promoting flowering under long-day conditions in rice. Mol. Plant **6**:202–215.

Zagotta, M.T., Hicks, K.A., Jacobs, C.I., Young, J.C., Hangarter, R.P., and Meeks-Wagner, D.R. (1996). The *Arabidopsis* ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. Plant J. **10**:691–702.