

Identification of a seed dormancy gene in soybean sheds light on crop domestication

Xin Wei* & Xuehui Huang*

College of Life Sciences, Shanghai Normal University, Shanghai 200234, China

Received October 11, 2018; accepted October 17, 2018; published online November 1, 2018

Citation: Wei, X., and Huang, X. (2018). Identification of a seed dormancy gene in soybean sheds light on crop domestication. *Sci China Life Sci* 61, 1439–1441. <https://doi.org/10.1007/s11427-018-9410-y>

Crop domestication, usually referring to the process of genetic adaptation of wild plants to crops, had a profound influence on mankind. Moreover, it will have ramifications for future human development. Crop domestication began at least 10,000 years ago at several domestication centers (Doebley et al., 2006). More than 100 seed and fruit crops, such as maize, wheat, rice, soybean, and tomato, have been domesticated. Compared to the progenitors, modern crops have advantageous changes in many morphological and physiological traits. Some of these modifications, such as the decreasing of seed shattering or fruit abscission, increasing of fruit number and size, loss of seed dormancy and photoperiodic controls, and better taste, are common domestication traits in different crops. These convergent domesticated or parallel selected traits have been called the ‘domestication syndrome’. A fundamental question in the study of crop domestication is whether there are common orthologs underlying the domestication syndrome in different crops (Lenser and Theißen, 2013).

Large efforts have been made to find a genetic basis of the convergent domestication from about two decades ago (Patterson et al., 1995). From that time, numerous domesticated genes had been identified and these discoveries are still increasing quickly. However, only a few genes have been found to be domesticated in parallel within cereals, such as *Sh1*, *Hdl* and *Waxy*. Take *Sh1* for example, it was first identified in sorghum, and then parallel selection was de-

tected in the orthologue genes, *OsSh1* in rice and *ZmSh1* in maize (Lin et al., 2012). To date, few such domesticated genes across different plant families have been reported.

Recently, research using genome-wide association studies (GWAS), genomics analysis, and functional analysis together suggested that the green-stay G gene (*GmG*) in soybean also controlled seed dormancy and had been selected during domestication (Wang et al., 2018). In addition, the orthologs of *GmG* in rice (*OsG*) and tomato (*SolyG*) were found to have similar functions and be selected during domestications as well (Table 1). The identification of the G gene may help to domesticate new crops to feed the rapid increasing population in the world.

In this study, the groups from Prof. Zhixi Tian and Prof. Chengcai Chu of Institute of Genetics and Developmental Biology, Chinese Academy of Sciences identified using GWAS, a gene encoding a CAAX amino-terminal protease protein in soybean that controlled the green seed coat, using cultivated soybeans with green or yellow seed coats. The mutation in *GmG* resulted in the loss of a small part of the protein in the C termini, which in turn led to the yellow seed coat of soybean. The function of G was then confirmed through genetically transforming the G allele into the g allele soybean.

Soybean was domesticated from wild soybean (*Glycine soja*) in East Asia approximately 5,000 years ago (Zhang et al., 2017). Interestingly, the wild soybean has the G allele but displays a black seed coat. The proportion of the G allele in wild, landraces and varieties was 100%, 21% and 4%, re-

*Corresponding authors (Xin Wei, email: xwei@shnu.edu.cn; Xuehui Huang, email: xhhuang@shnu.edu.cn)

Table 1 Conserved functions of G gene in plants and its parallel selection in several crops

Plants	Gene Name	Gene ID in database	Phenotypes of transgenic lines	Population genetics analysis	Candidate causative variation
Soybean	<i>GmG</i>	<i>Glyma.01G198500</i>	Change of seed coat and dormancy	Under domestication selection	A/G mutation (changing the splicing site)
Rice	<i>OsG</i>	<i>LOC_Os03g01014</i>	Change of germination rate	Under domestication selection	G/T mutation in 9th exon (non-synonymous)
Tomato	<i>SolyG</i>	<i>Solyc08g005010</i>	Not performed	Under domestication selection	G/C mutation in 5th exon (non-synonymous)
<i>Arabidopsis</i>	<i>AtG</i>	<i>At2G35260</i>	Change of germination rate	Not involved	Not involved

spectively, indicating the possible selection on *GmG* during domestication. The domestication of *GmG* was further validated by selective sweep analysis. Additionally, the germination rate was compared between the g allele and its transgenic lines, which showed that *GmG* contributed to seed dormancy in the soybean.

This work also identified the orthologous genes of *GmG* from different plants (Table 1), and found that *OsG* had also undergone domestication. Three non-synonymous SNPs were found in the *OsG*, clustering the wild and cultivated rice accessions into five haplotypes. The haplotype analysis suggested that *OsG* was fixed at the early stage of rice domestication, which was in line with the routes of rice domestication (Huang et al., 2012).

The genomic sequence of *OsG* from wild rice was introduced into two *japonica* varieties (termed *Osg*). Seed dormancy was observed in the transgenic lines. CRISPR of *Osg* in a *japonica* variety showed faster seed germination rate. Overexpression of *OsG* and *Osg* in the *japonica* variety background showed that transgenic lines with *OsG* had a stronger dormancy than that with *Osg*.

By analyzing nucleotide polymorphism related indices in the re-sequenced genomes of 360 tomato accessions (including wild relatives of tomato, cherry tomatoes and cultivated tomatoes), *SolyG* was detected in a selective sweep region. Further analysis of the variations revealed that the frequency of a non-synonymous SNP in *SolyG* of wild and cultivated tomatoes was significantly different. These results suggested that *SolyG* was also a domesticated gene.

In addition, the coding regions of *GmG* and *Gmg* were introduced to two *Arabidopsis* mutants (*Atg*). *GmG* transgenic fully complemented the weak dormancy of *Atg*, whereas the *Gmg* transgenic lines did not, indicating G orthologs had a conserved function to control dormancy in different plants. Correspondently, the haplotypes in the *Arabidopsis* population showed various levels of dormancy.

Yeast two-hybrid screening was used to characterize the molecular mechanisms of G underlying seed dormancy. Two key enzymes involved in ABA biosynthesis were found to interact with *AtG*. The contents of ABA and the biosynthesis precursors in *AtG* were higher than that in *Atg*. Consistently, the soybean with *GmG* allele and the transgenic lines with *GmG* allele both had a significantly higher ABA content than

soybean with the *Gmg* allele. These results revealed that G affected seed dormancy by interacting with abscisic acid synthesis genes.

Seed dormancy is one of the most important targets in crop domestication. A large number of seed dormancy QTLs in plants have been identified (Chen and Penfield, 2018; Sugimoto et al., 2010; Torada et al., 2016). However, few genes were found to be parallel selected. The G gene in soybean, which is responsible for seed dormancy, has been parallel selected in multiple crop families. This gene is the first gene identified to affect seed dormancy and undergo domestication across several crop families. Loss-of-function mutations are the most common mutational lesion in domesticated genes. However, in rice and tomato, the SNPs in G resulted in weak mutations rather than loss-of-function mutations, suggesting that weak selection might be one clue to the detection of parallel selection in different crops. The findings in this research will not only help us to understand the process of domestication but also to facilitate gene functional studies on seed dormancy.

Compliance and ethics The author(s) declare that they have no conflict of interest.

- Chen, M., and Penfield, S. (2018). Feedback regulation of COOLAIR expression controls seed dormancy and flowering time. *Science* 360, 1014–1017.
- Doebley, J.F., Gaut, B.S., and Smith, B.D. (2006). The molecular genetics of crop domestication. *Cell* 127, 1309–1321.
- Huang, X., Kurata, N., Wei, X., Wang, Z.X., Wang, A., Zhao, Q., Zhao, Y., Liu, K., Lu, H., Li, W., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490, 497–501.
- Lenser, T., and Theißen, G. (2013). Molecular mechanisms involved in convergent crop domestication. *Trends Plant Sci* 18, 704–714.
- Lin, Z., Li, X., Shannon, L.M., Yeh, C.T., Wang, M.L., Bai, G., Peng, Z., Li, J., Trick, H.N., Clemente, T.E., et al. (2012). Parallel domestication of the shattering1 genes in cereals. *Nat Genet* 44, 720–724.
- Paterson, A.H., Lin, Y.R., Li, Z., Schertz, K.F., Doebley, J.F., Pinson, S.R.M., Liu, S.C., Stansel, J.W., and Irvine, J.E. (1995). Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science* 269, 1714–1718.
- Sugimoto, K., Takeuchi, Y., Ebana, K., Miyao, A., Hirochika, H., Hara, N., Ishiyama, K., Kobayashi, M., Ban, Y., Hattori, T., et al. (2010). Molecular cloning of *sdr4*, a regulator involved in seed dormancy and domestication of rice. *Proc Natl Acad Sci USA* 107, 5792–5797.
- Torada, A., Koike, M., Ogawa, T., Takenouchi, Y., Tadamura, K., Wu, J., Matsumoto, T., Kawaura, K., and Ogihara, Y. (2016). A causal gene for seed dormancy on wheat chromosome 4a encodes a map kinase kinase.

[Curr Biol](#) 26, 782–787.

Wang, M., Li, W., Fang, C., Xu, F., Liu, Y., Wang, Z., Yang, R., Zhang, M., Liu, S., Lu, S., et al. (2018). Parallel selection on a dormancy gene during domestication of crops from multiple families. [Nat Genet](#) 50,

1435–1441.

Zhang, S.R., Wang, H., Wang, Z., Ren, Y., Niu, L., Liu, J., and Liu, B. (2017). Photoperiodism dynamics during the domestication and improvement of soybean. [Sci China Life Sci](#) 60, 1416–1427.