



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect



RESEARCH ARTICLE

## The identification of presence/absence variants associated with the apparent differences of growth period structures between cultivated and wild soybeans



LI Yan-fei<sup>1\*</sup>, HONG Hui-long<sup>1,2\*</sup>, LI Ying-hui<sup>1</sup>, MA Yan-song<sup>1,3</sup>, CHANG Ru-zhen<sup>1</sup>, QIU Li-juan<sup>1</sup>

<sup>1</sup> National Key Facility for Crop Gene Resources and Genetic Improvement (NFCRI)/Key Lab of Germplasm Utilization, Ministry of Agriculture/Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, P.R.China

<sup>2</sup> College of Agriculture, Northeast Agricultural University, Harbin 150030, P.R.China

<sup>3</sup> Soybean Research Institute, Heilongjiang Academy of Agricultural Sciences, Harbin 150086, P.R.China

### Abstract

The cultivated soybean (*Glycine max* (L.) Merr.) was distinguished from its wild progenitor *Glycine soja* Sieb. & Zucc. in growth period structure, by a shorter vegetative phase (V), a prolonged reproductive phase (R) and hence a larger R/V ratio. However, the genetic basis of the domestication of soybean from wild materials is unclear. Here, a panel of 123 cultivated and 97 wild accessions were genotyped using a set of 24 presence/absence variants (PAVs) while at the same time the materials were phenotyped with respect to flowering and maturity times at two trial sites located at very different latitudes. The major result of this study showed that variation at PAVs is informative for assessing patterns of genetic diversity in *Glycine* spp. The genotyping was largely consistent with the taxonomic status, although a few accessions were intermediate between the two major clades identified. Allelic diversity was much higher in the wild germplasm than in the cultivated materials. A significant domestication signal was detected at 11 of the PAVs at 0.01 level. In particular, this study has provided information for revealing the genetic basis of photoperiodism which was a prominent feature for the domestication of soybean. A significant marker-trait association with R/V ratio was detected at 14 of the PAVs, but stripping out population structure reduced this to three. These results will provide markers information for further finding of R/V related genes that can help to understand the domestication process and introgress novel genes in wild soybean to broaden the genetic base of modern soybean cultivars.

**Keywords:** wild soybean, cultivated soybean, genomic structure variation, interspecific differentiation, photoperiod

## 1. Introduction

The process of domestication followed by selection has sharply narrowed the genetic base of crop plants implying a reduction in their capacity to adapt to a changing environment (Fernie *et al.* 2006). In contrast, their wild relatives have generally retained a much larger reservoir of genetic variation. Soybean (*Glycine max* (L.) Merr.) was domesti-

Received 26 January, 2015 Accepted 15 April, 2015  
Correspondence LI Ying-hui, Tel/Fax: +86-10-82105843,  
E-mail: liyinghui@caas.cn; QIU Li-juan, Tel/Fax: +86-10-82105843,  
E-mail: qiulijuan@caas.cn

\* These authors contributed equally to this study.

© 2016, CAAS. All rights reserved. Published by Elsevier Ltd.  
doi: 10.1016/S2095-3119(15)61048-6

cated in East Asia from its wild progenitor *Glycine soja* Sieb. & Zucc (Hymowitz 2004). *G. soja* is fully intercrossable with *G. max*, so is regarded as a straightforward source of genes for soybean improvement (Cooper et al. 2001; Hyten et al. 2006; Lam et al. 2010; Li et al. 2010, 2013) although very few utilization is reported to date.

Presence/absence variants (PAVs) are substantial length segments of DNA which are not uniformly present in the genome of a given species. The phenomenon has been widely explored in maize (Lai et al. 2010), and has also been documented in rice (Xu et al. 2012), soybean (Lam et al. 2010; Li et al. 2014; Wang et al. 2014) and sorghum (Zhang et al. 2014). Along with other DNA-based markers, PAVs, which can be established by comparing a crop's genome to that of their wild relatives, have been used to locate the geographical origin of a crop species and chart the process of its domestication (Huang et al. 2008; Kim et al. 2010; Lam et al. 2010; Swanson-Wagner et al. 2010; He et al. 2011; Chia et al. 2012). PAVs have been associated with the determination of plant traits. A 3.8-kb segment of the gene *Gmsalt3* has been shown to have a demonstrable effect on the salinity tolerance of soybean (Guan et al. 2014). The recently completed sequencing of seven phylogenetically and geographically representative *G. soja* accessions has been used to establish a “pan-genome” of the wild species (Li et al. 2014). By comparing the wild soybean genome with the cultivated soybean reference genome, we have revealed the presence of 338 PAVs in genic regions (Li et al. 2014).

The cultivated soybean is a short day plant, which has a shorter vegetative growth period (V) and a longer reproductive growth period (R) than its wild ancestor (Shu et al. 1986). The R/V ratio, a parameter for evaluating growth period structure, is important not only in expanding the geographical range of cultivation, but also in increasing the crop yield. Although several major genes controlling flowering time and maturity time in soybean are known (Liu et al. 2008; Watanabe et al. 2009, 2011; Xia et al. 2012), the genetic basis of R/V was limited (Wang et al. 2015). Here, a germplasm panel including both *G. max* and *G. soja* accessions were genotyped at a set of 24 PAVs in an attempt to document the distribution of PAVs in both species, detect signals of selection during domestication and establish correlations between individual PAVs and phenotype, with a particular focus on R/V.

## 2. Results

### 2.1. Allelic variation at PAVs

All 24 PAVs were variable across the full set of 220 acces-

sions (Table 1). The frequency of the “presence” allele varied from 0.118 (C-23) to 0.468 (A-107), and the Nei index of gene diversity from 0.223 (C-23) to 0.499 (A-107), with mean value of 0.433. Although each of the PAVs was informative within each species, there was a large inter-specific difference in the frequency of the “presence” allele (0.592 and 0.140, respectively, in *G. soja* and *G. max*). The Nei index of gene diversity calculated from 23 of the 24 loci was greater among the *G. soja* than among the *G. max* accessions, except for A-107. For A-107, *G. max* was slightly more polymorphic than *G. soja*. Based on the  $F_{st}$  outlier test, 11 of the loci, namely A-3, A-58, A-94, B-120, B-122, B-13, D-13, D-142, D-98, F-72, and F-78 were identified as outliers at 0.01 level. The  $F_{st}$  values associated with these loci ranged from 0.408 (A-3) to 0.705 (B-120), with a mean of 0.563, representing a ~2.5-fold higher value than the average  $F_{st}$  value (0.218) associated with the other 13 PAVs.

### 2.2. Genetic differentiation between *G. soja* and *G. max*

The phylogenetic analysis separated the germplasm into two recognizable major clades which corresponded with taxonomic status (Fig. 1). The shared-allele distances separating most of the cultivated accessions were smaller than those separating the wild accessions, as expected following a domestication and selection bottleneck. However, 18 wild and 21 cultivated accessions fell between the two major clades implying a degree of gene flow between the two species. Of the 21 intermediate cultivated accessions, 14 expressed one or more trait generally associated with *G. soja*, such as small seed, black testa or a vinyl habit. Similarly, 13 of the 18 intermediate wild accessions expressed one or more cultivated traits, such as large seed, non-black testa or erect habit.

### 2.3. Genotype/phenotype associations for R/V and the detection of signals of selection

All 220 accessions reached maturity by the end of the experiment at the low latitude site, but only 153 did so at the high latitude site. Most of the accessions which failed to reach maturity originated from southern China. The distribution of R/V at each site is shown in Fig. 2. At the high latitude site, R/V varied from 0.38 to 2.37 (coefficient of variation of 40.1%), while at the low latitude site, it was much less variable (1.09 to 2.04, coefficient of variation of 15.1%) (Table 2). The mean R/V for the *G. soja* germplasm was 0.76 at the high latitude site, a value which was only ~56% that was recorded at the low latitude site (1.36). For

Download English Version:

<https://daneshyari.com/en/article/10180042>

Download Persian Version:

<https://daneshyari.com/article/10180042>

[Daneshyari.com](https://daneshyari.com)