Available online at www.sciencedirect.com
SciVerse ScienceDirect
Journal of Genetics and Genomics 40 (2013) 93−96
JGG

LETTER TO THE EDITOR

Cloning of *Ln* Gene Through Combined Approach of Map-based Cloning and Association Study in Soybean

Increasing yield is one of the most important goals in crop breeding. Soybean (*Glycine max* L. Merr.), one of the most economically important leguminous seed crops, provides the majority of plant proteins, and more than a quarter of the world's food and animal feed (Graham and Vance, 2003). The yield of soybean is finally determined by the number of seeds per unit area, which affected by many characters, such as height, branching number, photosynthesis, seed size, seed number. The number of seeds per pod is taken for one of the critical components that related to yield (You et al., 1995), and it has long been considered in soybean production (Takahashi, 1934). Genetics studies indicated one major gene (termed as *Ln*) contributed to the variations of leaflet and seed number per pod (Domingo, 1945; Sawada, 1988). Ovate leaflet usually associated with non-4-seeded pod, and narrow leaflet linked with 4-seeded pod. It was suggested that the ovate leaflet and non-4-seeded pod were dominant over narrow leaflet and 4-seeded pod (Domingo, 1945; Sawada, 1988; Jeong et al., 2011). The inconsistent results from different studies on the effects of narrow leaflet and 4-seeded pod on final yield indicated that it may related to the circumstance (Mandl and Buss, 1981; You et al., 1995; Dinkins et al., 2002). Additionally, it was suggested that the trait is associated with geographical origin (Chen and Nelson, 2004), which indicated that this trait might undergo domestication process and has been used for breeding to adapt particular environment.

The seed was developed from ovule, thus the number of seeds per pod was mainly determined by the number of ovules per placenta. To check if the difference of seed numbers between *Ln* and *ln* was caused by the inherent ovules per ovary, we selected two typical cultivars to investigate the ovules under microscope. *Han2296*, a cultivar from Korea, has the characters of ovate leaflet (Fig. S1A), and a high ratio of 2-seed pod and 1-seed pod (account for ∼60% and ∼35% of the total pods respectively) (Fig. S1B and C). Whereas, *Lvbaoshi*, a cultivar from Shanxi Province, has the properties of narrow leaflet (Fig. S1A), and a high ratio of 4-seeded pod and 3-seed pod (account for ∼40% and ∼45% of the total pods respectively) (Fig. S1B and C). The investigation showed that most numbers of ovules per ovary in *Han2296* were 3 and 2, but for *Lvbaoshi*, the numbers of 4 and 3 take the most (Fig. 1A), which indicated that the *Ln* gene may initially affect the formation of ovule.

A fundamental goal of evolutionary biology is to understand the genetic basis of adaptation (Orr and Coyne, 1992). Map-based cloning, one of the major technique served for forward genetics, has been proved to be an invaluable approach for the identification of genes that determine traits of interest (Peters et al., 2003). Although the ever-increasing release of genome sequences has greatly facilitated its manipulation, map-based cloning is still a rather tedious, labour intensive and often unsuccessful approach (Fitzgerald et al., 2012). In complicated polyploid plant genomes, the manipulation becomes more difficult (Fitzgerald et al., 2012). For instance, it took about 10 years to clone the *E1* genes in soybean (Xia et al., 2012). Association mapping, through association detecting between phenotypes and genetic base variants, provides a powerful tool to dissect complex traits, especially for domesticated traits (Nordborg and Weigel, 2008; Tian et al., 2010; Jimenez-Gomez, 2011). However, accompanied with the positive associated loci, false associations always cannot be avoided (Ingvarsson and Street, 2011). In addition, due to the linkage disequilibrium (LD) of artificial selected region, even higher in soybean (Chan et al., 2012), it is hard to narrow an association to a particular candidate gene. Thus, an efficient approach to quick isolate gene controlling trait of interest will greatly benefit functional genomics.

To isolate the *Ln* gene, we firstly took the map-based cloning approach. We generated a F_2 mapping population with 1868 individual lines derived from a cross between '*Han2296*' and '*Lvbaoshi*'. The segregation of the leaflet shapes and seed number per pod of F_2 plants fitted a ratio of 3:1 (ovate leaflet and non-4-seeded pod: narrow leaflet and 4-seeded pod = 1360:508). Through linkage mapping, *Ln* was primarily delimited in an interval ∼8.9 Mb between the two molecular markers BARCSOYSSR_20_0663 and sat_324 on the long arm of chromosome 20 (Fig. S2A). Fine-mapping further narrowed *Ln* locus to a 0.97 Mb region between the

1673-8527/$ - see front matter
http://dx.doi.org/10.1016/j.jgg.2013.01.002

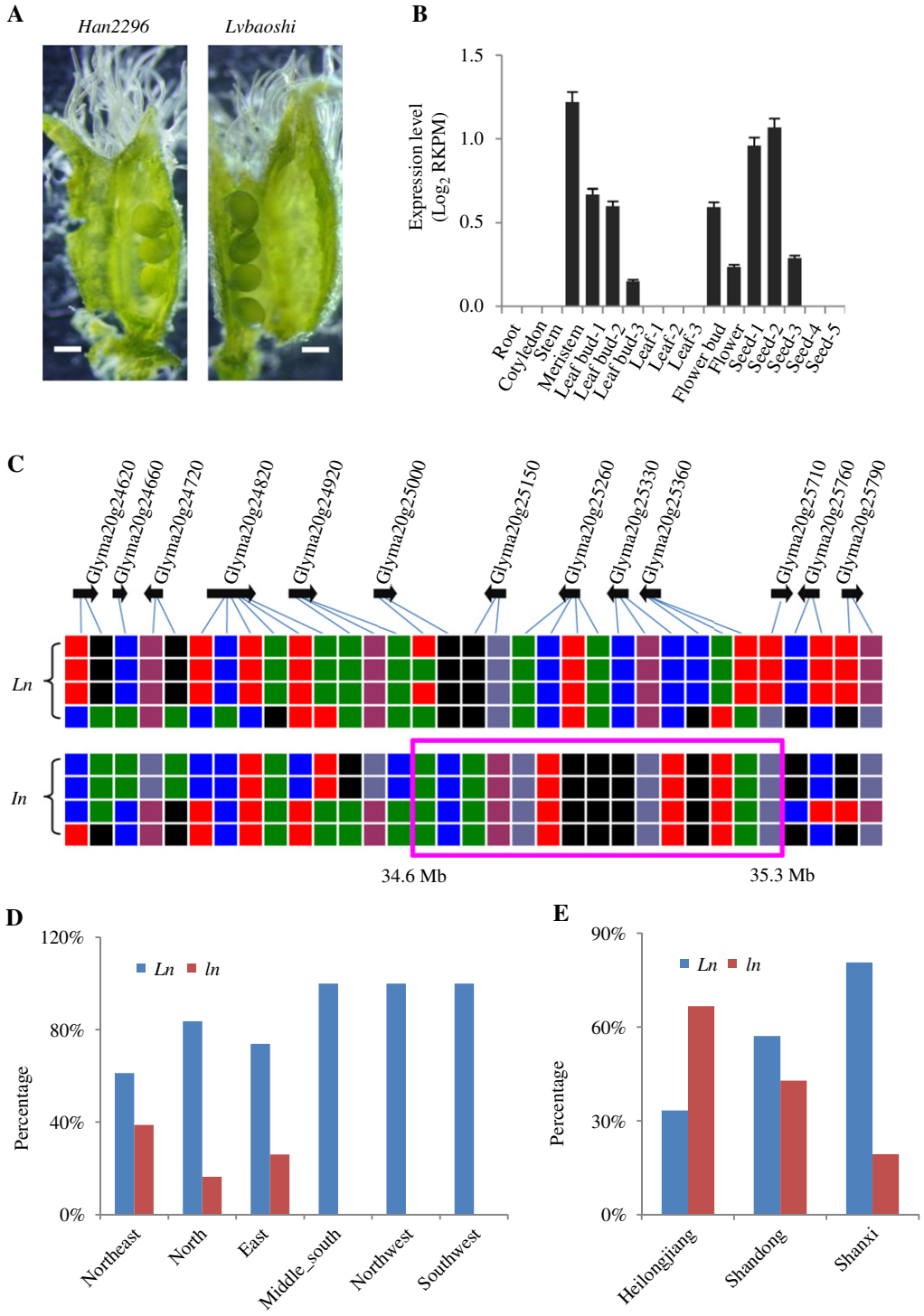

Fig. 1. Expression pattern of *Ln* and selection of *ln*.

A: ovule of wild type and *ln* mutant, Bar = 0.2 mm. **B:** expression of *Ln* in different tissues. **C:** selection sweep of *Ln* gene between wild type and *ln* mutant groups. The accessions from top to end are *Williams 82*, *Harosoy*, *Clark*, *Han2296*, *Qingpidou*, PI548160, *Lvbaoshi* and PI547811. The red box indicates nucleotide 'T', the green box indicates nucleotide 'A', the blue box indicates nucleotide 'C', the black box indicates nucleotide 'G', the crimson box indicates multiple nucleotides insertion, the dark blue box indicates deletion. **D:** geographical distribution of the *Ln* and *ln* genotypes in China. **E:** distribution of the *Ln* and *ln* genotypes in Heilongjiang, Shandong and Shanxi Provinces.

two SSR markers SSR-820 and SSR-790 (Fig. S2B). Screening with six newly developed SNP markers SNP1−SNP6 (primers sequences in Table S1), *Ln* was located within a 57.6-kb region between the SNP3 and SNP4 (Fig. S2B). Within this region, there are four predicated open reading frames (ORFs), *Glyma20g24970*, *Glyma20g24980*, *Glyma20g24990*, and *Glyma20g25000* (Schmutz et al., 2010). No SNPs were detected between the parental cultivars within the first three gene models, whereas, a G → C polymorphism which changed the amino acid Asp (*GAT*) to His (*CAT*) was found at the first exon of *Glyma20g25000* (Fig. S2C). Blast search revealed that this gene was the homologous gene of *AT1G68480* (*JAG*) in *Arabidopsis thaliana*, which belong to C_2H_2 and C_2HC zinc fingers superfamily. Similar to *ln* in soybean, the development of leaflet (Ohno et al., 2004) and lateral organs (Dinneny et al., 2004) in *jag* mutant were also affected. Thus, *Glyma20g25000* was considered as the candidate gene of *Ln* controlling leaflet and seed number per pod in soybean.

To further confirm the candidate gene, we sequenced the SNP locus in 19 isogenic lines of *ln* from Clark or Harosoy genetic background (Table S2). All the lines were found to have *ln* mutation, a result consistent with our presumption. We also sequenced the *Ln* locus in other 54 isogenic lines from Clark or Harosoy genetic background but not *ln* mutants, and did not detect any mutations in the *Ln* locus (Table S2). In addition, we investigated this SNP in six cultivars with typical phenotype, and found that all the cultivars with narrow leaflet and high ratio of 4-seeded pod number contained this point mutation 'C', but the cultivars with ovate leaflet and non-4-seeded pod did not (Table S2). The association strengthened the inference that *Glyma20g25000* is *Ln*.

Transcriptional profiling showed that the *Ln* expression level is high in meristem, young leaf buds, young flower bud and young developed seeds (Fig. 1B). In addition, its expression level decreased along with the development of the organ. The transcription of *Ln* was not detected in old developed seeds. These results indicate that *Ln* might affect the cell division but not the cell expansion. Combining the expression pattern of *Ln* with the investigation that the number of ovule in *Ln* was less than that in the *ln* mutant, we deduced that *Ln* might affect the formation of ovule through regulating the cell division.

It was suggested that *Ln* had undergone artificial selection and been used for breeding to adapt particular environment (Chen and Nelson, 2004). We sequenced the flanking gene models of *Ln* locus in different cultivars to check their selection sweep between two groups. Sequences analysis indicated that *ln* group had high LD with about 700 kb (Fig. 1C), which is much stronger than that of the *Ln* group (Fig. 1C) and also much higher than common LD of nature soybean populations (~150 kb) (Lam et al., 2010). Another bigger population of 21 cultivars from *ln* group showed the high LD was commonly existed (Fig. S3). The strong LD of *ln* mutants suggested that the gene had undergone artificial selection during soybean breeding. More cultivars were sequenced to address the selection of *Ln* to adapt different circumstance. One hundred and ninety-five cultivars from the different geography loci in China were selected, sequencing result indicated that the *ln* mutation type mainly located in the Northeast of China (especially Heilongjiang Province and Shandong Province), while the *Ln* type dispersed the whole country (Fig. 1D and E). These data indicated that *ln* might be recently selected in the breeding or seed distribution for the soybean production.

In this study, through the combination of map-based cloning and association study, we quickly cloned the gene controlling leaflet and seed number per pod in soybean. The approach described herein can also be applied in other forward genetics investigations of domestication agronomic traits in soybean and other important crops.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos. 31271297 and 31222042), "One-hundred talents" Startup Funds from Chinese Academy of Sciences, and National Key Basic Research Program (No. 2009CB118402).

SUPPLEMENTARY DATA

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jgg.2013.01.002.

Chao Fang[a,b], *Weiyu Li*[a], *Guiquan Li*[c], *Zheng Wang*[a], *Zhengkui Zhou*[a], *Yanming Ma*[a], *Yanting Shen*[a,b], *Congcong Li*[a,b], *Yunshuai Wu*[a], *Baoge Zhu*[a], *Weicai Yang*[a], *Zhixi Tian*[a,*]

[a]*Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China*

[b]*University of Chinese Academy of Science, Beijing 100039, China*

[c]*Shanxi Agriculture University, Taigu 030801, China*

*Corresponding author. Tel: +86 10 8261 3935, fax: +86 10 8261 6905.
E-mail address: zxtian@genetics.ac.cn (Z. Tian)

Received 11 January 2013
Revised 15 January 2013
Accepted 15 January 2013
Available online 19 January 2013

REFERENCES

Chan, C., Qi, X., Li, M.W., Wong, F.L., Lam, H.M., 2012. Recent developments of genomic research in soybean. J. Genet. Genomics 39, 317−324.

Chen, Y.W., Nelson, R.L., 2004. Evaluation and classification of leaflet shape and size in wild soybean. Crop Sci. 44, 671−677.

Dinkins, R.D., Keim, K.R., Farno, L., Edwards, L.H., 2002. Expression of the narrow leaflet gene for yield and agronomic traits in soybean. J. Hered. 93, 346−351.

Dinneny, J.R., Yadegari, R., Fischer, R.L., Yanofsky, M.F., Weigel, D., 2004. The role of JAGGED in shaping lateral organs. Development 131, 1101−1110.

Domingo, W., 1945. Inheritance of number of seeds per pod and leaflet shape in the soybean. J. Agric. Res. 70, 251−268.

Fitzgerald, T.L., Kazan, K., Manners, J.M., 2012. The application of reverse genetics to polyploid plant species. Crit. Rev. Plant Sci. 31, 181−200.

Graham, P.H., Vance, C.P., 2003. Legumes: importance and constraints to greater use. Plant Physiol. 131, 872−877.

Ingvarsson, P.K., Street, N.R., 2011. Association genetics of complex traits in plants. New Phytol. 189, 909−922.

Jeong, N., Moon, J.K., Kim, H.S., Kim, C.G., Jeong, S.C., 2011. Fine genetic mapping of the genomic region controlling leaflet shape and number of seeds per pod in the soybean. Theor. Appl. Genet. 122, 865−874.

Jimenez-Gomez, J.M., 2011. Next generation quantitative genetics in plants. Front. Plant Sci. 2, 77.

Lam, H.M., Xu, X., Liu, X., Chen, W., Yang, G., Wong, F.L., Li, M.W., He, W., Qin, N., Wang, B., Li, J., Jian, M., Wang, J., Shao, G., Wang, J., Sun, S.S., Zhang, G., 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nat. Genet. 42, 1053−1059.

Mandl, F.A., Buss, G.R., 1981. Comparison of narrow and broad leaflet isolines of soybean. Crop Sci. 21, 25−27.

Nordborg, M., Weigel, D., 2008. Next-generation genetics in plants. Nature 456, 720−723.

Ohno, C.K., Reddy, G.V., Heisler, M.G.B., Meyerowitz, E.M., 2004. The *Arabidopsis JAGGED* gene encodes a zinc finger protein that promotes leaf tissue development. Development 131, 1111−1122.

Orr, H.A., Coyne, J.A., 1992. The genetics of adaptation: a reassessment. Am. Nat. 140, 725−742.

Peters, J.L., Cnudde, F., Gerats, T., 2003. Forward genetics and map-based cloning approaches. Trends Plant Sci. 8, 484−491.

Sawada, S., 1988. Inheritance of leaflet shape in soybeans. Soybean Genet. Newsl. 15, 61−65.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., Jackson, S.A., 2010. Genome sequence of the palaeopolyploid soybean. Nature 463, 178−183.

Takahashi, N., 1934. Linkage relation between the genes for the form of leaves and the number of seeds per pod of soybeans. Jpn. J. Genet. 9, 208−225.

Tian, Z., Wang, X., Lee, R., Li, Y., Specht, J.E., Nelson, R.L., McClean, P.E., Qiu, L., Ma, J., 2010. Artificial selection for determinate growth habit in soybean. Proc. Natl. Acad. Sci. USA 107, 8563−8568.

Xia, Z., Watanabe, S., Yamada, T., Tsubokura, Y., Nakashima, H., Zhai, H., Anai, T., Sato, S., Yamazaki, T., Lu, S., Wu, H., Tabata, S., Harada, K., 2012. Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. Proc. Natl. Acad. Sci. USA 109, E2155−E2164.

You, M., Liu, Y., Zhao, T., Gai, J., 1995. Effects of leaf shape on seed yield and its components in soybeans. Soybean Genet. Newsl. 22, 66−70.