Biochemical Systematics and Ecology 62 (2015) 229–235

Contents lists available at ScienceDirect

Biochemical Systematics and Ecology

journal homepage: www.elsevier.com/locate/biochemsyseco

Evaluation of the geographical pattern of genetic diversity of *Glycine soja* and *Glycine max* based on four single copy nuclear gene loci: For conservation of soybean germplasm

Yunsheng Wang [a, b], Muhammad Qasim Shahid [c], Fozia Ghouri [c], Faheem Shehzad Baloch [d], Ying Wang [b, e, *], Hongwen Huang [b, e, **]

[a] College of Environment and Life Science, Kaili University, Guizhou Province, 556011, China
[b] Wuhan Botanical Garden, Chinese Academy of Science, Wuhan City, Hubei Province, 430074, China
[c] College of Agriculture, South China Agricultural University, Guangzhou, Guangdong Province, 510642, China
[d] Department of Field Crops, Faculty of Agricultural and Natural Science, Abant İzzet Baysal University, Bolu, Turkey
[e] South China Botanical Garden, Chinese Academy of Science, Guangzhou City, Guangdong Province, 510642, China

ARTICLE INFO

Article history:
Received 23 May 2015
Received in revised form 4 September 2015
Accepted 6 September 2015
Available online 19 September 2015

Keywords:
Demographical history
Gene locus
Genetic diversity
Nucleotide polymorphism
Population genetics

ABSTRACT

Genetic diversity and its geographical patterns play a very important role in species conservation and exploitation. Here, nucleotide polymorphism patterns of four single copy nuclear gene loci in wild (*Glycine soja*) and cultivated soybean (*Glycine max*) populations from different geographical regions as well as their demographic history were analyzed. The results showed that: (1) Southern subpopulation has the highest, while central subpopulation revealed the lowest genetic diversity among three Chinese *G. soja* subpopulations. (2) Northern Chinese *G. max* subpopulation depicted higher genetic diversity than other two Chinese, Korean, Japanese and American *G. max* subpopulations. (3) Significant genetic differentiation ($P < 0.001$) was observed among Chinese *G. soja* subpopulations from three ecological zones. There was also a significant genetic differentiation ($P < 0.01$) between three Chinese and Japanese subpopulations of *G. max*. (4) The demographic dynamics revealed that effective population size of *G. soja* is expanding, while it was constant in *G. max*. *G. soja* is a useful germplasm resource to widen the genetic diversity of *G. max*. This study suggests that native populations of *G. soja* from different geo-ecological regions should be protected to conserve the genetic diversity.

1. Introduction

Genetic diversity is an important aspect of bio-diversity reflecting the differences in phenotype, DNA and protein between different individuals of a population or a species. The mutant, recombination and exchange happened within or between chromosomes is inherent reason for genetic diversity, while genetic drift, migration and selection are exterior impetus of genetic diversity. Genetic diversity and geographic patterns of a species always played a very crucial role in making policies for proper species conservation and exploitation (Frankham and Briscoe, 2002).

* Corresponding author. Wuhan Botanical Garden, Chinese Academy of Science, Wuhan City, Hubei Province, 430074, China.
** Corresponding author. Wuhan Botanical Garden, Chinese Academy of Science, Wuhan City, Hubei Province, 430074, China.
E-mail addresses: yingwang@wbgcas.cn (Y. Wang), huanghw@mail.scbg.ac.cn (H. Huang).

http://dx.doi.org/10.1016/j.bse.2015.09.006
0305-1978/

Cultivated soybean (*Glycine max*) is an economically important crop and is the world's primary source of vegetable protein and oil (Zhuang, 1999; Shahid et al., 2009). The genetic diversity and demographic patterns of cultivated soybean are key issues for plant biologists. Annual wild soybean (*Glycine soja*) is believed to be the wild ancestor of *G. max*, and natively distributed in the northern-east Asia, including China, far-east of Russia, Korea and Japan (Zhuang, 1999). However, natural habitats of *G. soja* have considerably decreased in recent few decades and now its existence is in severe danger because of inappropriate human activities (Li et al., 2005). *G. soja* is known to possess many useful genes for important traits such as high yield, resistance to various biotic and abiotic factors (Wang and Li, 2000).

Molecular markers such as SSR (Microsatellite or simple sequence repeat), VNTR (Variable number of tandem repeat), RAPD (Random amplified polymorphism DNA), AFLP (Amplified fragment length polymorphism), RFLP (Restriction fragment length polymorphism), SRAP (Sequence related amplified polymorphism), SNP (Single nucleotide polymorphism), SSCP (Single-strand conformation polymorphism) had been commonly used to study the genetic diversity in different crops (Alsaleh et al., 2014). During the last two decades, with the development of sequencing technology, the single nucleotide polymorphism (SNPs) markers are increasingly being used for the evaluation of genetic diversity (Johnson, 2009; Baloch et al., 2015). Single copy nuclear gene sequence received high attention to study the genetic diversity of plant species. However, the nucleotide polymorphism is not evenly distributed along the whole gene region. Generally, the intron region of a gene had the highest polymorphism, while the coding region of a gene had the lowest (Arabidopsis Genome Initiative, 2000). Therefore, intron region of single copy nuclear gene was commonly used in the genetic diversity evaluation of intra-species in recent years. However, there are few reports about the genetic diversity evaluation of soybean by single copy nucleotide genes, especially based on intron sequence.

In this study, the geographical diversity patterns of four single copy nuclear gene loci, mostly from intron region, of 152 *G. soja* and 77 *G. max* accessions sampled from different eco-geographical regions were analyzed and compared with each other. The main objectives of this study were to (1) assess the genetic diversity of *G. soja* and *G. max* populations from different geo-ecological niches; (2) make effective strategies for conservation of *G. soja* germplasm. Taken together, these populations could be used for conservation programs and genomic studies.

2. Materials and method

2.1. Sampling

Plant materials: A total of 229 accessions including 152 *G. soja* (consisting of 141 accessions from China, 3 from Korea, 5 from Japan and 3 from Russia) and 77 *G. max* populations (including 52 accessions from China, 5 from Korea, 14 from Japan and 6 from America) were sampled from wide geographical regions and used as plant material in this study (Supplementary Table S1). *G. max* populations were divided into six geo-ecological subpopulations according to the theory of Lu et al. (1981) and Bu and Pan (1982): Japanese, Korean, American, Northern Chinese, Central Chinese and Southern Chinese sub-populations. *G. soja* populations were also divided into six geo-ecological subpopulations: Japanese, Korean, Russian, Northern Chinese, Central Chinese and Southern Chinese subpopulation.

2.2. DNA extraction

Seeds of each soybean accession were placed on the absorbent papers in a petri dish and sufficient quantity of water was added for seed germination. Seedlings were grown in pots at room temperature under natural conditions. Young leafs about 100 mg were collected to extract the genomic DNA according to the protocol of CTAB method (Doyle and Doyle, 1987).

2.3. Loci selection and sequencing

Four high-diversity gene loci were observed to compare the genetic variation of soybean populations in this study. We detected an intron splice site from soybean EST sequence (TC229661) published in dataset bank (The Institute for Genomic Research, http://www.tigr.org/) by intron detected software online (http://www.sgn.cornell.edu/cgi-bin/tools/intron_detection/find_introns.pl) and three gene loci, including locus B (AF105221), locus C (AJ003246) and locus D (J02746) were selected according to Van et al. (2005) and primer sequences at these loci were: Primer1-F 5′ -GCGTTGGAGATTGGAGATAA-3′, primer1-R 5′-TGGGACAGTAAGCAGTTGACC-3′; primer 2-F 5′-GCGACGCATTCAGTACACACTACA-3′C, primer 2-R 5′-GCGGCCAAAGAAAGACAAGTAGA-3′TA; primer 3-F 5′-GCGGGCAAAAAGGAAGAAAT-3′, primer 3-R 5′-GCGGGGAAAAGGT-GAAAATTA-3′; primer-4F 5′-GCGGGGTGTTCGAGGTTTCTAAT-3′ and primer 4-R 5′-GCGATGCGTTGGAATTTCAGGATA-3′, respectively (Wang et al., 2015).

PCR reaction was kept as follow: dNTP 1ul (100uM), $MgCl_2$ 4.0ul (25 mM), buffer 5.0ul (10×), Tagase 2.5u, primers 0.4ul (100uM), and ddH_2O was added to make the volume up to 50 μl. The PCR protocol was used as follow: 94 °C for 5 min, 35 cycles of 94 °C for 30 s, 56 °C for Locus B and Locus D; 54 °C for 45 s for Locus A and locus C and 72 °C for 90 s, and a final extension at 72 °C for 10 min (Wang et al., 2015). The sizes of the amplification products of locus A, B, C and D were 411 bp, 483 bp, 516 bp and 428 bp, respectively. Further, PCR products were purified by PCR purify kit from Clonetech Corporation. The 3730/ABI sequencer was used for sequencing. Homologous sequence alignment was done with clustalx program (Thompson et al., 1997) and modified by software bioedit v7.0.5 (Hall, 1999). A total of 1839 bp homologous genomic