

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Wild Soybeans: An Opportunistic Resource for Soybean Improvement

Muhammad Amjad Nawaz, Seung Hwan Yang and
Gyuhwa Chung

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.74973>

Abstract

Reduced genetic diversity in cultivated soybean coupled with changing dietary expectations, climate change, and increase in population demands expansion of current gene pool. Wild soybeans are an opportunistic resource and a rational choice to discover novel genes and gene families for alternative crop production systems and to improve soybean. Multiple agronomic traits, lineage-specific genes, and domestication-related traits have been studied in wild soybeans in contrast to cultivated soybeans, and it has been proved that wild soybeans are an essential genomic resource containing unique and useful genetic resources that have been lost during domestication to expand the gene pool in order to improve soybean. Wild soybean is very often a plant of disturbed habitats of Southeast Asia. The vulnerability of these habitats to agriculture systems and urban expansion causes a reduction in the area of distribution and hence the diversity. To capture the wild soybean genetic diversity in its main distribution areas, a unique and comprehensive germplasm collection, characterization, and conservation platform is direly needed. Chung's Wild Legume Germplasm Collection is preserving and maintaining a representative wild soybean germplasm collection guided by the principles of conservation genetics. These wild legumes and particularly wild soybean is a promising genetic resource for soybean breeders.

Keywords: CWLGC, genetic diversity, genetic resource, *Glycine soja*, rediscovery of landraces, wild legumes, germplasm conservation

1. Introduction: rediscovery of crop wild relatives

Challenged by limited land and water resources and a concomitant increase in population, changing dietary expectations and climatic change are demanding escalated food supplies

[1]. The nutritional value of grain legumes is far better than cereals even if their production is low, making them a unique and essential component of balanced diet [2]. Grain legumes have suffered a reduction in genetic diversity largely due to plant breeding activities aimed at artificial selection of desirable traits. The new varieties, as well as land races in farmer's field, have desirable characteristics which have become genetically diverged from their ancestors or wild progenitors [3].

In order to cope with the global warming led climatic variations and limited water supplies, there is a constant need of crop improvement; the crop potential has been reduced due to domestication, genetic bottlenecks, and artificial selection [2]. To explore more genes and gene families for alternative production systems, crop wild relatives are a rational choice mainly due to limited or no breeding barriers [4]. The wild progenitors of crops are sometimes easily available, but this is not the case for all species as some of the wild species have gone extinct, or in other cases, multiple progenitors contributed to the genome of the domesticated plants, e.g., wheat. In some cases, some species are indirectly expanding the genomes of the domesticated crops as they may be related species of wild progenitors or wild cousins [5]. Wild crop relatives are mostly adapted to larger climatic variations and are evolved to withstand biotic and abiotic stresses [6]. Therefore, for crop improvement, we have two possibilities, namely, genetic modification or introduction of genetic materials through breeding with crop wild relatives. Of course the use of genetic engineering to create genetically modified plants is relatively quick and efficient but the acceptance of genetically modified plants among the consumers is still controversial. On the other hand, the desirable traits including resistance to biotic and abiotic stresses, nutritional values can be incorporated into the current agricultural crop by using conventional and new breeding technologies. This practice is sometimes quite challenging mainly due to linkage drag; however, recent advances in genetics and genomic approaches have expanded our understanding of evolution, linkage, and heredity of complex traits [4–6] (**Table 1**).

2. Wild soybean

The *Glycine* genus comprises of two subgenera, namely, *Glycine* Willd and *Soja* (Moench) F. J. Hermann. Among 28 species classified under two subgenera, only two annual species *G. soja* Sieb & Zucc (wild) and *G. max* (cultivated) are consumed as food or feed either directly or indirectly [13]. On the basis of cytological, proteomic, and genomic evidence, the wild species *G. soja* is considered as the progenitor of cultivated species *G. max*. Both herbaceous annual species are Asian inborn; mainly distributed in Southeast and Far East Asia including China, Korea, Japan, Taiwan, and Russia [14]. Wild soybean harbors treasured genetic resource and extraordinarily important gene pool; genes and gene families responsible for higher oil and protein contents, resistance to drought and high temperatures, disease resistance, and insect pest resistance [15].

Wild soybeans grow on roadsides, riversides villages, lakeshores, wastelands, and fertile valleys. Apart from numerous phenotypic distinctions among both species, their annual growing habit with similar ploidy level and ability to produce fertile offspring without genetic isolation

| Crop | Wild relatives | Reference |
|--------------------------------------|---------------------------------------------------------------------------|-----------|
| Rice | <i>O. glaberrima</i> Steud. | [7] |
| <i>Oryza sativa</i> L. | <i>O. nivara</i> Sharma et Shastry | |
| | <i>O. rufipogon</i> Griff. | |
| | <i>O. breviligulata</i> | |
| | <i>O. barthii</i> | |
| | <i>O. meridionalis</i> | |
| Wheat <i>Triticum aestivum</i> L. | <i>O. glumaepatula</i> | [8] |
| | <i>T. boeoticum</i> | |
| | <i>T. monococcum</i> | |
| | <i>T. dicoccoides</i> | |
| | <i>T. araraticum</i> | |
| | <i>T. dicoccum</i> | |
| | <i>T. palaeocolchicum</i> | |
| | <i>T. timopheevii</i> | |
| | <i>T. durum</i> | |
| | <i>T. turgium</i> | |
| | <i>T. persicum</i> (= <i>T. carthlicum</i> Nevski) | |
| | <i>T. polonicum</i> | |
| | <i>T. turanicum</i> (= <i>T. orientale</i> Perc.) | |
| | <i>T. parvicoccum</i> | |
| | <i>T. spelta</i> | |
| | <i>T. macha</i> | |
| | <i>T. vavilovii</i> | |
| | <i>T. aestivum</i> | |
| | <i>T. compactum</i> (= <i>T. aestivum</i> grex aestivo-compactum Schiem.) | |
| | <i>T. sphaerococcum</i> | |
| Corn | <i>Z. luxurians</i> | [9] |
| <i>Zea mays</i> L. | <i>Z. diploperennis</i> | |
| | <i>Z. Mexicana</i> | |
| | <i>Zea nicaraguensis</i> | |
| | <i>Zea perennis</i> | |
| Soybean | <i>G. soja</i> Sieb. & Zucc. | [10] |
| <i>Glycine max</i> L. | See Table 2 for wild cousins of soybean | |
| Common bean | <i>P. coccineus</i> | [11] |
| <i>Phaseolus vulgaris</i> L. | <i>P. dumosus</i> | |
| | <i>P. maculatifolius</i> | |
| | <i>P. costaricensis</i> | |
| | <i>P. acutifolius</i> | |
| | <i>P. lunatus</i> | |
| | <i>P. parvifolius</i> | |
| | <i>P. costaricensis</i> | |
| | <i>P. albescens</i> | |
| | <i>P. filiformis</i> | |

| Crop | Wild relatives | Reference |
|---------------------------|-------------------------|-----------|
| Chickpea | <i>C. reticulatum</i> | [12] |
| <i>Cicer arietinum</i> L. | <i>C. echinospermum</i> | |
| | <i>C. pinnatifidum</i> | |
| | <i>C. judaicum</i> | |
| | <i>C. bijugum</i> | |
| | <i>C. cuneatum</i> | |
| | <i>C. chorassanicum</i> | |
| | <i>C. yamashitae</i> | |

Table 1. List of wild relatives of three major cereals and three major legumes.

results in a flow of certain characteristics from wild to cultivated populations [16]. Wild soybeans exhibit distinct geographical patterns as well as interspecific horizontal mechanisms of flow of genetic information to cultivated species mainly because of sharing the same gene pool and close proximity [10] (**Table 2**).

Many parts of China and South Korea which were previously regarded as habitat for wild soybean are now being used for agricultural, commercial purposes (roads, buildings or dams) or are now part of the sea. Destruction of natural soybean habitats due to land clearance for agricultural or industrial purposes has led to decreased wild germplasm resources [19]. Furthermore, reduction in genetic diversity has been witnessed due to the domestication of soybean during past three decades. Progenitor wild species exhibit discrete geographic patterns with greater genetic diversity but the selection and allele frequency changes during domestication has curtailed genetic variability. Various studies have reported a reduction of genetic diversity up to 50% in domesticated/improved cultivars as compared to wild progenitors [3, 15]. Artificial selection and domestication mainly focused dominant selection of desirable traits such as oil content, seed size, and seed coat luster, imposing selection pressure on particular traits and ignoring other important traits. Another factor involved in diminishing genetic diversity is habitat fragmentation [20] (**Figure 1**).

2.1. Wild soybean genome: rediscovering the lost diversity

Whole genome sequencing of wild soybean genome started a new era for soybean functional and comparative genomics and has substantially increased our understanding about soybean domestication history, bottlenecks, lost diversity and has created a way forward towards its potential use in expanding the gene pool of soybean. The wild and cultivated soybeans have significant and useful genomic differences which highlight the phenotypic differences and as well as the domestication-related traits. Kim et al. [21] aligned 915.4 Mb genomic sequence of wild soybean with soybean reference genome excluding the gaps and found that wild soybean genome covered 97.65% of soybean genome with a difference of 35.2 Mb (3.76% of 937.5 Mb). The difference region consisted of 0.267% substitution bases, 0.043% insertion/deletions (indels), and 3.45% of large deleted sequences. Single nucleotide polymorphisms (SNPs) and

| Genus | Subgenus | Species | Authority | Reference |
|----------------|----------------|--------------------------|-------------------------------|-----------|
| <i>Glycine</i> | <i>Soja</i> | <i>G. max</i> | Merr. | [17, 18] |
| | | <i>G. soja</i> | Sieb. & Zucc. | |
| | <i>Glycine</i> | <i>G. canescens</i> | F. J. Hermann | |
| | | <i>G. clandestine</i> | Wendl. | |
| | | <i>G. latrobeana</i> | Benth. | |
| | | <i>G. tabacina</i> | Benth. | |
| | | <i>G. tomentella</i> | Hayata | |
| | | <i>G. falcate</i> | Benth. | |
| | | <i>G. latifolia</i> | (Benth) Newell & Hymowitz | |
| | | <i>G. argyrea</i> | Tindale | |
| | | <i>G. cyrtoloba</i> | Tindale | |
| | | <i>G. curvata</i> | Tindale | |
| | | <i>G. arenaria</i> | Tindale | |
| | | <i>G. microphylla</i> | (Benth.) Tindale | |
| | | <i>G. albicans</i> | Tindale & Craven | |
| | | <i>G. hirticaulis</i> | Tindale & Craven | |
| | | <i>G. lactovirens</i> | Tindale & Craven | |
| | | <i>G. dolichocarpa</i> | Tateishi & Ohashi | |
| | | <i>G. pindanica</i> | Tindale & Craven | |
| | | <i>G. stenophita</i> | B. Pfeil & Tindale | |
| | | <i>G. peratosa</i> | B. E. Pfeil & Tindale | |
| | | <i>G. rubiginosa</i> | Tindale & B. E. Pfeil | |
| | | <i>G. aphyonota</i> | B. Pfeil | |
| | | <i>G. pullenii</i> | B. Pfeil, Tindale & Craven | |
| | | <i>G. gracei</i> | B. E. Pfeil & Craven | |
| | | <i>G. montis-douglas</i> | B. E. Pfeil, Tindale & Craven | |
| | | <i>G. pescadrensis</i> | Hayata | |
| | | <i>G. syndetika</i> | B. E. Pfeil & Craven | |

Table 2. Members of genus *Glycine*.

insertions/deletions (indels) in precisely aligned areas differed by 0.31% between cultivated and wild soybean. The complex genome rearrangement is mainly caused by indels, inversions, and translocations (up to thousands of base pairs); along with SNPs and indels [22]. The wild soybean genome has greater allelic diversity than that of soybean. Resequencing of 17 wild and 14 cultivated soybean genomes to an average of $\times 5$ depth and $>90\%$ coverage identified higher allelic diversity [15].

Based on whole-genome SNP analysis using the parameter $\theta\pi$, a higher level of genetic diversity was found in wild soybeans (2.97×10^{-3}) as compared to cultivated soybeans (1.89×10^{-3}).



Figure 1. Geographic distribution of wild soybeans.

Similar findings were also reported when 302 wild and cultivated soybeans were whole-genome sequenced to an average depth of $> \times 11$; the genetic diversity (π) decreased from 2.94×10^{-3} in *G. soja* to 1.40×10^{-3} in landraces and to 1.05×10^{-3} in improved cultivars suggesting that nearly half of the annotated resistance-related sequences were lost during the domestication [3]. Noticeably, total number of SNPs in wild (5,924,662) and cultivated soybeans (4,127,942) was comparable (35 and 5%, respectively) and the ratio of nonsynonymous SNP to synonymous SNPs was higher in cultivated (1.38) than wild soybeans (1.36). It was observed that the number of fixed loci were lower in the wild (463,409) as compared to cultivated soybeans (2,148,585). These findings suggest that there should be low-frequency alleles in domesticated soybean owing to domestication bottleneck. However, contrary to this, low-frequency alleles were abundant in wild soybean suggesting that the wild soybean habitat has reduced and the cultivated soybean population has expanded [15].

2.2. Expanding the gene pool for soybean improvement

Wild soybeans are a potential genetic resource for the improvement of cultivated soybean and aid greatly in exploring alternative production systems. Wild soybeans, as in case of another wild relative of cultivated crop species, contain higher genetic diversity as they had a long time opportunity to evolve and withstand under varied environmental conditions without inference by humans [4, 15]. Wild soybeans are interfertile with cultivated soybeans and represent an easily accessible or primary gene pool for soybean improvement [10]. However, the global climate change and increase in human population have developed a scenario of securing, conserving, characterizing, and using wild soybeans as a resource for soybean improvement. Loss of genetic diversity during the process of soybean domestication and presence of a domestication bottleneck, i.e., domestication syndrome has led towards changes in growth habits, loss of germination inhibition and mechanisms of seed dispersal [4]. This domestication has also enabled the crop plants to withstand and adapt to modern agriculture and farming system, which is very encouraging. However, loss of diversity in cultivated soybeans calls for revisiting natural diversity reservoirs, i.e., wild soybeans in search of potential genes/alleles

for higher yield. Multiples sequences that are unique to wild soybeans have been discovered but a report from Korea by Chung et al. [23] also witnessed gene loss events in wild soybeans. However, this discrepancy might be due to diversity of wild collections [4]. Multiple agronomic traits, lineage-specific genes, and domestication-related traits have been studied in wild soybeans in contrast to cultivated soybeans, and it has been proved that wild soybeans are an essential genomic resource containing unique and useful genetic resources that have been lost during domestication to expand the gene pool in order to improve soybean [3, 15, 24, 25]. One recent example is the salt-resistant gene GmCH1X identified in wild soybean. The salt-resistant gene originally did not have a Ty1/copia retrotransposon insertion into its exon 3 in wild soybean and controls 80% salt tolerance in wild soybean (W05) as compared to its counterpart C08 (cultivated soybean) which had retrotransposon insertion possibly due to recent round of whole genome duplication [24], strongly implying that wild soybeans' genetic diversity must be explored.

2.3. Traits from the wild

Recent development in high-throughput sequencing technologies is clearly promoting a revolution in the comparative genomic sequencing of major crops. A rapid growth in the number of sequenced genomes of crops and their wild relatives has established that wild species tend to have higher genetic diversities, making the wild relatives promising natural resources of novel genes/alleles for crop improvement. Many studies have provided the details on wild soybean specific genes/alleles controlling major abiotic and biotic stress tolerance-related traits. Contrastingly, cultivated soybeans also have unique genes/alleles which have been possibly lost during the evolution of wild soybeans. However, the results of each comparative study must be based on the genetic diversity present within the subject population [4, 21, 26].

2.3.1. Domestication-related traits

Identification of genes for domestication-related traits is an important task to maintain diversity in crops for improvement. Such a knowledge provides essential understanding of how and what genetic signatures have brought necessary changes in plant phenotype and physiology during the process of domestication. In soybean, the domestication-related traits are the increased size of inflorescence, grain yield, seed size, seed color, hilum color, pubescence form, apical dominance, stem determinacy, and plant height. Many of the domestication QTLs have been identified, such as twining habit (Ch. 02 and 18), hard seededness (Ch. 02 and 06), determinate habit (Ch. 17), maximum internode length (Ch. 06, 18 and 19), flowering time (Ch. 06 and 16), pod dehiscence (Ch. 16), seed weight (Ch. 17), stem determinacy (Ch. 17), oil content (Ch. 03, 11, 12, 13, 15, 17), flower color (Ch. 13), seed coat color (Ch. 08), pubescence form (Ch. 01, 12, 18, 19, 20), and plant height (Ch. 18) [3, 15].

2.3.2. Other traits

Many useful QTLs/genes have been obtained by characterizing wild soybeans or using genetic populations resulting from crosses between wild and cultivated soybeans. These

genes/QTLs have been characterized to understand the stress resistance mechanisms and various biochemical pathways related to plant development, yield, and local breeding traits. (1) Multiple genes/alleles responsible for flower color, i.e., pinkish-white and white flowers have been identified from wild soybeans. Flavonoid 3'5'-hydroxylase (F3'5'H) and dihydroflavonol-4-reductase (DFR) are responsible for anthocyanin production. Different loci control the anthocyanin content and decide the fate of flower color. Different loci, i.e., W1, W3, W4, w1-s1, w1-s2, w1-lp, and w1-p2 have been reported to control white color and pinkish-white [27, 28]. (2) Other studies on seed antioxidant, phenolics, and flavonoid contents have identified GmMATE1, 2, 4 genes [29]. Astringent taste in soy products is caused by group A saponins. *Glyma15g39090* has been successfully characterized as sg-5 gene in natural wild soybean mutant line CWS5095. The gene oxygenates the C-21 position of soyasapogenol B or other intermediate which results in the production of saponin A [30, 31]. Another gene Sg-1 of this pathway has also been characterized from Korean wild soybean natural mutant which controls Ab series of saponins [32]. (3) Soybean cyst nematode is a global threat to soybean and host plant resistance is an ideal way of managing the damages. Wild soybean was used to discover SNPs and candidate genes significantly associated with soybean cyst nematode resistance, i.e., 10 SNPs and genes related to disease resistance-related proteins with leucine-rich region. Two genes, namely, a mitogen-activated protein kinase gene (Ch. 18) and a MYB transcription factor (Ch. 19) were found to be strong candidates [33]. (4) Many genes and transcription factors have also been identified from wild soybeans which play important role in drought stress tolerance [34]. (5) A salt-tolerant gene, i.e., GmCHX1 has been identified from wild soybean through whole-genome de novo sequencing approach. (6) A phosphatase 2C-1 (PP2C-1) allele has been reported from wild soybean to be involved in seed weight and seed size. Apart from these genes many QTLs related to (7) linolenic acid production, (8) yield, height and maturity, (9) soybean cyst nematode, (10) seed yield, seed weight, seed filling period, maturity, plant height, and lodging, (11) salt tolerance, (12) sclerotinia stem rot, (13) root traits, (14) oil and local breeding, (15) shoot fresh weight, (16) seed antioxidant, phenolics and flavonoids, and many other traits have been mapped and identified from wild soybeans or populations developed by crossing wild and cultivated soybeans [24, 29, 34–41]. Taking the advantage of higher genetic diversity and identified genes and QTLs for important traits from wild soybean will gear up the soybean yield improvement in changing climatic conditions and modern dietary demands. Joint ventures guided by the principles of plant breeding, genetics, genomics, and modern biotechnology are underway in many parts of the world to improve soybean in terms of resistance to biotic and abiotic stresses, adaptation to low water and higher temperature conditions, as well as intensive agricultural systems.

3. Wild soybean germplasm conservation

Wild soybean (the presumed progenitor of soybean) is very often a plant of disturbed habitats of Southeast Asia. Such habitats are mostly on the roadsides, intensive agricultural lands, and areas with higher human disturbances in terms of land use. The adaptation to these disturbed areas actually predisposes the wild soybeans to agricultural systems; this is one of the reasons for its domestication in East Asia [2, 19]. The vulnerability of these habitats to agriculture

systems and urban expansion causes reduction in area of distribution and hence the diversity. As discussed earlier, wild soybean is an efficient resource for identification and characterization of furnished and important genes for soybean improvement [10]. Economically, genebanks and wild genetic resources have been unambiguously reported to have led towards higher economic return by increasing soybean productivity [5]. Wild soybean germplasm preservation is underway in many countries mainly China, Korea, Japan, and the United States of America [10, 19]. Surely, the collections are growing by following the principles of conservation genetics; however, the complete representative collections are yet to be achieved as there remain many unexplored uninhabited natural habitats of wild soybeans which might carry many useful genes, alleles, or mutations. The undiscovered variations are greatly in demand by plant breeders to increase soybean production for an ever-increasing population [18]. Wild soybean germplasm should be collected mainly to (a) to understand the taxonomy and phylogenetic relationships, (b) to understand the biosystematics of certain yield-related pathways, (c) characterize and conserve germplasm, and (d) make it available for soybean breeders across the globe [10, 18–19]. Currently, there are many gene banks which are working on wild soybean germplasm conservation. In Southeast Asia, China holds the largest wild soybean germplasm collection of 6172 accessions under Chinese Crop Germplasm Information system [42], followed by Chung's Wild Legume Germplasm Collection (CWLGC) holding 6012 accessions and National Institute of Agrobiological Sciences Genebank in Japan which holds 1131 accessions [43]. Outside East Asia, the largest collection of wild soybean germplasm is USDA Soybean Germplasm Collection holding nearly 21,810 accessions belonging to 21 species of genus *Glycine* [44]. Almost 1179 accessions belong to 20 wild relatives of cultivated soybean including the wild soybean. N. I. Vavilov Institute of Plant Genetic Resources (VIR) holds ~350 accessions [45]. All of these germplasm collections are either focused on cultivated soybean or limited to a particular country/region. There is a dire need of germplasm center particularly focused on the collection, characterization, and dissemination of wild soybean accessions from the main distribution area of the species i.e., Southeast Asia. Out of the above-mentioned germplasm collections, CWLGC is primarily focused on wild soybean germplasm collections from China, Korea, Japan, and Far East Russia near Chinese border.

3.1. Chung's wild legume germplasm collection

Chung's wild legume germplasm collection strives to develop a comprehensive conservation program resourcefully and efficiently to conserve and promote the genetic diversity within wild legumes with main focus on wild legumes. Guided with the principles of conservation genetics, CWLGC focuses on (a) direct collection, (b) acquirement, (c) conservation, (d) evaluation and characterization, and (e) documenting and distribution of wild soybean germplasm. CWLGC was established in 1983 by Professor Gyuhwa Chung at Department of Biotechnology, Chonnam National University, Yeosu campus, Republic of Korea. CWLGC holds 10,314 different legume genera. However, particular emphasis is on the germplasm collection, multiplication, evaluation, and utilization of *G. soja*, *Amphicarpaea edgeworthii*, *Vigna vexillata*, *Rhynchosia volubilis*, and *Phaseolus nipponensis*. The CWLGC efforts incredibly focused on wild soybeans, native to main centers of diversity, particularly East Asia. Wild soybean seeds are the main collection at CWLGC, and the seeds are considered as the currency of germplasm which is safeguarded for global food availability and security as well as



Figure 2. Wild soybean germplasm conservation at CWLGC.



Figure 3. A glimpse of diverse legume germplasm seed collection at CWLGC.

to preserve natural genetic diversity of wild soybeans and other legumes. This diverse collection of pinhead to large-sized seeds is important than ever as plant breeders, geneticists, conservationists, and biotechnologists need to cope with the changing climate, reduction in arable land area, occurrence of natural disasters, environmental degradation, and rising expectations in nutritional standards (**Figures 2 and 3**).

Conflict of interest

The authors declare that there exists no conflict of interest.

Author details

Muhammad Amjad Nawaz, Seung Hwan Yang and Gyuhwa Chung*

*Address all correspondence to: chung@chonnam.ac.kr

Department of Biotechnology, Chonnam National University, Yeosu, Republic of Korea

References

- [1] Considine MJ, Siddique KHM, Foyer CH. Nature's pulse power: legumes, food security and climate change. *Journal of Experimental Botany*. 2017;**68**:1815-1818. DOI: 10.1093/jxb/erx099
- [2] Foyer CH, Lam HM, Nguyen HT, Siddique KHM, Varshney RK, Colmer TD, Cowling W, Bramley H, Mori TA, Hodgson JM, Cooper JW, Miller AJ, Kunert K, Vorster J, Cullis C, Ozga JA, Qahlqvist ML, Liang Y, Shou H, Shi K, Yu J, Fodor N, Kaiser BN, Wong FL, Valliyodan B, Considine MJ. Neglecting legumes has compromised human health and sustainable food production. *Nature Plants*. 2016;**2**:16112. DOI: 10.1038/nplants.2016.112
- [3] Zhou Z, Jiang Y, Wang Z, Gou Z, Lyu J, Li W, Yu Y, Shu L, Zhao Y, Ma Y, Fang C, Shen Y, Liu T, Li C, Li Q, Wu M, Wang M, Wu Y, Dong Y, Wan W, Wang X, Ding Z, Gao Y, Xiang H, Zhu B, Lee SH, Wang W, Tian X. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nature Biotechnology*. 2015;**33**:408-415. DOI: 10.1038/nbt.3096
- [4] Munoz N, Liu A, Kan L, Li MW, Lam HM. Potential uses of wild germplasms of grain legumes for crop improvement. *International Journal of Molecular Sciences*. 2017;**18**:328. DOI: 10.3390/ijms18020328
- [5] Brozynska M, Furtado A, Henry RJ. Genomics of crop wild relatives: Expanding the gene pool for crop improvement. *Plant Biotechnology Journal*. 2015;**14**:1070-1085
- [6] Zhang H, Mittal N, Leamy LJ, Barazani O, Song BH. Back into the wild—Apply untapped genetic diversity of wild relatives for crop improvement. *Evolutionary Applications*. 2017;**10**:5-24. DOI: 10.1111/eva.12434
- [7] Sanchez PL, Wing RA, Brar DS. The wild relatives of rice: Genomes and genomics. In: *Genetics and Genomics of Rice, Plant Genetics and Genomics: Crop and Models*. New York: Springer Science + Business Media; 2014. pp. 9-25. DOI: 10.1007/978-1-4614-7903-1_2
- [8] Arrigo N, Guadagnuolo R, Lappe S, Pasche S, Parisod C, Felber F. Gene flow between wheat and wild relatives: empirical evidence from *Aegilops geniculata*, *Ae. neglecta* and *Ae. truncialis*. *Evolutionary Application*. 2011;**4**:685-695. DOI: 10.1111/j.1752-4571.2011.00191.x
- [9] Maazou ARS, Qiu J, Mu J, Liu Z. Utilization of wild relatives for maize (*Zea mays* L.) improvement. *African Journal of Plant Science*. 2017;**11**:105-113. DOI: 10.5897/AJPS2017.1521

- [10] Chung G, Singh RJ. Broadening the genetic base of soybean: A multidisciplinary approach. *Critical Reviews in Plant Sciences*. 2008;**5**:295-341. DOI: 10.1080/07352680802333904
- [11] Porch TG, Beaver JS, Debouck DG, Jackson SA, Kelly JD, Dempewolf H. Use of wild relatives and closely related species to adapt common bean to climate change. *Agronomy*. 2013;**3**:433-461. DOI: 10.3390/agronomy3020433
- [12] Sing R, Sharma P, Varshney RK, Sharma SK, Sing SK. Chickpea improvement: Role of wild species and genetic markers. *Biotechnology and Genetic Engineering Reviews*. 2008;**25**:267-314. DOI: 10.5661/bger-25-267
- [13] Guo J, Wang Y, Song C, Zhou J, Qiu L, Huang H, Wang Y. A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): Implications from microsatellites and nucleotide sequences. *Annals of Botany*. 2010;**106**:505-514. DOI: 10.1093/aob/mcq125
- [14] Phylogeographical studies of *Glycine soja*: Implicating the refugium during the quaternary glacial period and large-scale expansion after the Last Glacial Maximum. *Turkish Journal of Agriculture and Forestry*. 2016;**40**:825-838. DOI: 10.3906/tar-1607-30
- [15] Lam HM, Xu X, Liu X, Chen W, Yang G, Wong FL, Li MW, He W, Qin N, Wang B, Li J, Jian M, Wang J, Shao G, Wang J, SSM S, Zhang G. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. *Nature Genetics*. 2015;**42**:1053-1059. DOI: 10.1038/ng.715
- [16] Ohara M, Shimamoto Y. Importance of genetic characterization and conservation of plant genetic resources: The breeding system and genetic diversity of wild soybean (*Glycine soja*). *Plant Species Biology*. 2002;**17**:51-58. DOI: 10.1046/j.1442-1984.2002.00073.x
- [17] Sherman-Broyles S, Bombarely A, Powell AF, Doyle JL, Egan AN, Coate JE, Doyle JJ. The wild side of a major crop: Soybean's perennial cousins from down under. *American Journal of Botany*. 2014;**101**:1651-1665. DOI: 10.3732/ajb.1400121
- [18] Singh RJ, Nelson RL, Chung G. Soybean [*Glycine max* (L.) Merr.]. In: Singh RJ, editor. *Genetic Resources, Chromosome Engineering, and Crop Improvement*. Vol. 4. Boca Raton, Florida: CRC Press, (Taylor & Frances Group); 2007. pp. 13-50
- [19] Nawaz MA, Yang SH, Rehman HM, Baloch FS, Lee JD, Park JH, Chung G. Genetic diversity and population structure of Korean wild soybean (*Glycine soja* Sieb. and Zucc.) inferred from microsatellite markers. *Biochemical Systematics and Ecology*. 2017;**71**: 87-96. DOI: 10.1016/j.bse.2017.02.002
- [20] Van K, Kim MY, Shin JH, Kim KD, Lee YH, Lee SH. Molecular evidence for soybean domestication. In: Tuberosa R, Graner A, Frison E, editors. *Genomics of Plant Genetic Resources*. Dordrecht: Springer; 2014. pp. 465-481
- [21] Kim MY, Lee S, Van K, Kim TH, Jeong SC, Choi IY, Kim DS, Lee YS, Park D, Ma J, Kim WY, Kim BC, Park S, Lee KA, Kim DH, Kim KH, Shin JH, Jang YE, Kim KD, Liu WX,

- Chaisan T, Kang YJ, Lee YH, Kim KH, Moon JK, Schmutz J, Jackson SA, Bhak J, Lee SH. Whole-genome sequencing and intensive analysis of the undomesticated soybean (*Glycine soja* Sieb. and Zucc.) genome. PNAS. 2010;**107**:22032-22037. DOI: 10.1073/pnas.1009526107
- [22] Kim MY, Van K, Kang YJ, Kim KH, Lee SH. Tracing soybean domestication history: From nucleotide to genome. Breeding Science. 2012;**61**:445-452. DOI: 10.1270/jsbbs.61.445
- [23] Chung WH, Jeong N, Kim H, Lee WK, Lee YG, Lee SH, Yoon W, Kim JH, Choi IY, Choi HK, Moon JK, Kim N, Jeong SC. Population structure and domestication revealed by high-depth resequencing of Korean cultivated and wild soybean genomes. DNA Research. 2014;**21**:153-167. DOI: 10.1093/dnares/dst047
- [24] Qi X, Li MW, Xie M, Liu X, Ni M, Shao G, Song C, Yim AKY, Tao Y, Wong FL, Isobe S, Wong CF, Wong KS, Xu C, Li C, Wang Y, Guan R, Sun F, Fan G, Ziao Z, Zhou F, Phang TH, Liu X, Tong SW, Chang TF, Yiu SM, Tabata S, Wang J, Xu X, Lam HM. Identification of a novel salt tolerance gene in wild soybean by whole-genome sequencing. Nature Communications. 2014;**5**:4340. DOI: 10.1038/ncomms5340
- [25] Li YH, Zhou G, Ma J, Jiang W, Jin LG, Zhang Z, Guo Y, Zhang J, Sui Y, Zheng L, Zhang SS, Zuo Q, Shi XH, Li YF, Zhang WK, Hu Y, Kong G, Hong HL, Tan B, Song J, Liu ZX, Wang Y, Ruan H, Yeung CKL, Liu J, Wang H, Zhang LJ, Guan RX, Wang KJ, Li WB, Chen SY, Chang RZ, Jiang Z, Jackson SA, Li R, Qiu LJ. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. Nature Biotechnology. 2014;**32**:1045-1052. DOI: 10.1038/nbt.2979
- [26] Stupar RM. Into the wild: The soybean genome meets its undomesticated relative. PNAS. 2010;**107**:21947-21948. DOI: 10.1073/pnas.1016809108
- [27] Sundaramoorthy J, Park GT, Chang JH, Lee JD, Kim JH, Seo HS, Chung G, Song JT. Identification and molecular analysis of four new alleles at the W1 locus associated with flower colour in soybean. PLoS One. 2016;**11**:e0159865. DOI: 10.1371/journal.pone.0159865
- [28] Kim JH, Bae DN, Park SK, Jeong N, Lee K, Kang H, Kang ST, Moon JK, Park E, Jeong SC. Molecular genetic analysis of a novel recessive white flower gene in wild soybean. Crop Science. 2017;**57**:3027-3034. DOI: 10.2135/cropsci2017.03.0163
- [29] Li MW, Munoz NB, Wong CF, Wong FL, Wong KS, Wong JWH, Qi X, Li KP, Ng MS, Lam HM. QTLs regulating the contents of antioxidants, phenolics and flavonoids in soybean seeds share a common genomic region. Frontiers in Plant Sciences. 2016;**7**:854. DOI: 10.3389/fpls.2016.00854
- [30] Yano R, Takagi K, Takada Y, Mukaiyama K, Tsukamoto C, Sayama T, Kaga A, Anai T, Sawai S, Ohyama K, Saito K, Ishimoto M. Metabolic switching of astringent and beneficial triterpenoid saponins in soybean is achieved by a loss of function mutation in cytochrome P450 72A69. The Plant Journal. 2017;**89**:527-539

- [31] Rehman HM, Nawaz MA, Shah ZH, Yang SH, Chung G. Functional characterization of naturally occurring wild soybean mutant (sg-5) lacking astringent saponins using whole genome sequencing approach. *Plant Science*. 2017; in press. DOI: 10.1016/j.plantsci.2017.11.014
- [32] Park J, Kim JH, Krishnamurthy P, Tsukamoto C, Song JT, Chung G, Shannon G, Lee JD. Characterization of new allele of the saponin-synthesizing gene Sg-1 in soybean. *Crop Science*. 2016;**56**:1-7. DOI: 10.2135/cropsci2015.04.0250
- [33] Luo X, Bai X, Sun X, Zhu D, Liu B, Ji W, Cai H, Cao L, Wu J, Hu M, Liu X, Tang L, Zhu Y. Expression of wild soybean WRKY20 in *Arabidopsis* enhances drought tolerance and regulates ABA signalling. *Journal of Experimental Botany*. 2013;**64**:2155-2169. DOI: 10.1093/jxb/ert073
- [34] Li D, Pfeiffer TW, Cornelius PL. Soybean QTL for yield and yield components associated with *Glycine soja* alleles. *Crop Science*. 2007;**48**:571-581. DOI: 10.2135/cropsci2007.06.0361
- [35] Pantalone VR, Rebetzke GJ, Burton JW, Wilson RF. Genetic regulation of linolenic acid concentration in wild soybean *Glycine soja* accessions. *Journal of the American Oil Chemists' Society*. 1997;**74**:159-163
- [36] Zhang WK, Wang YJ, Luo GZ, Zhang JS, He CY, Wu XL, Gai JY, Chen SY. QTL mapping of ten agronomic traits on the soybean (*Glycine max* L. Merr.) genetic map and their association with EST markers. *Theoretical and Applied Genetics*. 2004;**108**:1131-1139
- [37] Iquira E, Humira S, François B. Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. *BMC Plant Biology*. 2015;**15**:5. DOI: 10.1186/s12870-014-0408-y
- [38] Zhang H, Li C, Davis EL, Wang J, Griffin JD, Kofsky J, Song BH. Genome-wide association study of resistance to soybean cyst nematode (*Heterodera glycines*) HG Type 2.5.7 in wild soybean (*Glycine soja*). *Frontiers in Plant Science*. 2016;**17**:1214. DOI: 10.3389/fpls.2016.01214
- [39] Manavalan LP, Prince SJ, Musket TA, Chaky J, Deshmukh R, Vuong TD, Song L, Cregan PB, Nelson JC, Shannon JG, Specht JE, Nguyen HT. Identification of novel QTL governing root architectural traits in an interspecific soybean population. *PLoS One*. 2015;**10**:e0120490. DOI: 10.1371/journal.pone.0120490
- [40] Ha BK, Vuong TD, Velusamy V, Nguyen HT, Shannon JG, Lee JD. Genetic mapping of quantitative trait loci conditioning salt tolerance in wild soybean (*Glycine soja*) PI 483463. *Euphytica*. 2013;**193**:79-88
- [41] Asekova S, Kulkarni KP, Patil G, Kim M, Song JT, Nguyen HT, Shannon JG, Lee JD. Genetic analysis of shoot fresh weight in a cross of wild (*G. soja*) and cultivated (*G. max*) soybean. *Molecular Breeding*. 2016;**36**:103
- [42] Chinese Crop Germplasm Information system. http://www.cgris.net/cgris_english.html [Accessed: 2017-11-20]

- [43] National Institute of Agrobiological Sciences Genebank. http://www.naro.affrc.go.jp/archive/nias/index_e.html. [Accessed: 2017-11-20]
- [44] United States Department of Agriculture. Soybean Germplasm Collection. <http://grscicoll.org/institutional-collection/usda-soybean-germplasm-collection>. [Accessed: 2017-11-20]
- [45] N. I. Vavilov Institute of Plant Genetic Resources. <http://www.vir.nw.ru/>. [Accessed: 2017-11-20]

