# Development of Highly Transferable Microsatellites for Panax ginseng (Araliaceae) Using Whole-Genome Data

Authors: Jiang, Peng, Shi, Feng-Xue, Li, Ya-Ling, Liu, Bao, and Li, Lin-Feng

Source: Applications in Plant Sciences, 4(11)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1600075

Applications in Plant Sciences 2016 4(11): 1600075

PRIMER NOTE

# DEVELOPMENT OF HIGHLY TRANSFERABLE MICROSATELLITES FOR *PANAX GINSENG* (ARALIACEAE) USING WHOLE-GENOME DATA[1]

PENG JIANG[2], FENG-XUE SHI[3], YA-LING LI[2], BAO LIU[2], AND LIN-FENG LI[2,4,5]

[2]Key Laboratory of Molecular Epigenetics of the Ministry of Education, Northeast Normal University, Changchun 130024, People's Republic of China; [3]Museum of Natural History, Northeast Normal University, Changchun 130117, People's Republic of China; and [4]Department of Biology, Washington University, St. Louis, Missouri 63130 USA

- *Premise of the study:* Highly transferable expressed sequence tag (EST) microsatellites were developed for *Panax ginseng* (Araliaceae), one of the most celebrated traditional Chinese medicines and an endangered species in East Asia, using whole-genome data.
- *Methods and Results:* Twenty-one EST microsatellites were characterized from next-generation sequencing and were composed of di- and trinucleotide repeats. Polymorphisms and genetic diversity were evaluated for 45 accessions of three ginseng landraces. The number of alleles for each locus ranged from one to five among the landraces, and the polymorphism information content varied from 0.0000 to 0.6450. These microsatellites were also tested for congeneric amplification with *P. notoginseng*, *P. stipuleanatus*, *P. quinquefolius*, *P. bipinnatifidus*, and the closely related species *Aralia elata*.
- *Conclusions:* These novel EST-derived microsatellite markers will facilitate further population genetic studies of the genera *Panax* and *Aralia*.

**Key words:** Araliaceae; microsatellite; *Panax ginseng*; polyploidy; traditional Chinese medicine.

*Panax* L. (Araliaceae) is a medicinally important genus, which consists of seven well-recognized species and one species complex, and is widely distributed in East Asia and North America (Lee and Wen, 2004). *Panax ginseng* C. A. Mey. is very popular in traditional Chinese medicine and has been used as an herbal remedy in East Asia for thousands of years (Liu and Xiao, 1992). Although the pharmacology and medical effects of *P. ginseng* have been investigated extensively, only a few studies have focused on the genetic diversity and population structure of this species (Li et al., 2015), primarily owing to the limitations of molecular markers. In our recent studies, we employed single-copy nuclear genes to investigate the genetic diversity of cultivated and wild ginseng (Li et al., 2015). However, the relatively low mutation rate of single-copy nuclear genes and the recent domestication of cultivated ginseng largely limited the application of our selected nuclear genes in assessment of its population structure and domestication history. In addition, we have recently discovered multiple rounds of whole-genome duplication within the genus *Panax* (Shi et al., 2015). *Panax ginseng* is an allotetraploid species and has undergone two rounds of whole-genome duplication, which makes it difficult to obtain all alleles through traditional Sanger sequencing. In this regard, as codominant markers with high mutation rates (Jarne and Lagoda, 1996), nuclear microsatellites might provide novel insights into the polyploidy of *Panax* and domestication of *P. ginseng*. In previous studies, other microsatellites have been developed (Kim et al., 2007; Yang et al., 2008; Ahn et al., 2009; Van Dan et al., 2010; Choi et al., 2011; Reunova et al., 2014). Nevertheless, most of these studies developed microsatellites based on traditional Sanger sequencing, and the transferability of most of these microsatellites to congeneric species remains unknown. We employed next-generation sequencing to develop microsatellites from whole-genome data and to test the transferability of these identified microsatellites in other *Panax* and *Aralia* L. species.

## METHODS AND RESULTS

***Plant material and DNA extraction*—**A total of 63 wild and cultivated ginseng accessions were collected from North Korea, eastern Russia, and northeastern China (Appendix 1). The related species *P. notoginseng* (Burkill) F. H. Chen ex C. Y. Wu & K. M. Feng, *P. stipuleanatus* H. T. Tsai & K. M. Feng, *P. bipinnatifidus* Seem., *P. quinquefolius* L., and *Aralia elata* (Miq.) Seem. were sampled from Yunnan, Jilin, and Sichuan provinces of China and from Wisconsin, USA (Appendix 1). We collected cultivated ginseng accessions from multiple geographic localities, while wild *P. ginseng* accessions and related species were generously provided by our collaborators. Of these *P. ginseng* accessions, 18 were subjected to whole-genome sequencing and 45 were used to test the polymorphisms of the identified microsatellites. Total genomic DNA was extracted from silica gel–dried leaves for each accession separately using the TIANGEN Plant Mini Kit (TIANGEN Biotech, Beijing, China).

***Microsatellite development*—**The construction of DNA libraries of the 18 ginseng accessions was conducted by Novogene Corporation (Beijing, China),

[1]Manuscript received 27 June 2016; revision accepted 11 August 2016.

The authors thank the editor and an anonymous reviewer for their constructive comments that greatly improved the manuscript, Dr. Zhenhui Wang for the de novo assembling of *Panax notoginseng*, and Dr. Hongxing Xiao and Dr. Yuezhi Pan for providing the wild *P. ginseng* samples and related *Panax* species. This work was supported by the National Natural Science Foundation of China (no. 31470010).

[5]Author for correspondence: lilf241@nenu.edu.cn

doi:10.3732/apps.1600075

 

TABLE 1. Characteristics of 21 microsatellite loci developed in *Panax ginseng*.

| Locus | Primer sequences (5′–3′) | Repeat motif | Allele size (bp) | $T_a$ (°C) | PIC | *A* | Putative function | GenBank accession no. |
|---|---|---|---|---|---|---|---|---|
| S2 | F: CTTGCTGCTTCTACATCC<br>R: GGTCTTGCTAATCCCAT | $(CCA)_4$ | 470 | 56 | 0.0000 | 1 | regulation of nuclear pre-mRNA domain protein | KU879256 |
| S6 | F: CCCAACCTACTAACATCC<br>R: GGTTTAGCTGCTCTGTACT | $(TCT)_6$ | 316 | 56 | 0.3749 | 2 | subtilisin-like protease | KU879260 |
| S11 | F: AATTGTACCTCCATAAACC<br>R: AGAGCCCGAGATAACCTA | $(TTC)_6$ | 370 | 56 | 0.0696 | 2 | DEAD-box ATP-dependent RNA helicase | KU879265 |
| S12 | F: GCGACGAATTAGACGATG<br>R: ATTGATTTCTCCTGCTGA | $(AG)_5$ | 358 | 48 | 0.0947 | 2 | cyclic dof factor 2 | KU879266 |
| S13 | F: TATTCCAATTCGGCAAAG<br>R: GGAGTGTTTGGGAGCATC | $(GGA)_6$ | 301 | 56 | 0.4441 | 5 | uncharacterized protein | KU879267 |
| S16 | F: ATGAAGCCGATGGTGGAG<br>R: TTCTCCAATACTTCTCCC | $(AAG)_5$ | 484 | 56 | 0.3748 | 2 | translation initiation factor IF-2 | KU879270 |
| S15 | F: TGAACTACTCCAGCTTCG<br>R: ACGGTGATGGCTGGTGGT | $(AGA)_6$ | 272 | 54 | 0.0000 | 1 | C-type lectin receptor-like tyrosine-protein kinase | KU879269 |
| S17 | F: ATTCCCGACAATAATGAG<br>R: TTGAGGCAAGCAAGGTGA | $(CT)_5$ | 329 | 54 | 0.3645 | 2 | probable RNA helicase SDE3 | KU879271 |
| S19 | F: GGGATGCCCTTACCCTTTG<br>R: CGTGTTGGCGTTGTCGTG | $(GGC)_4$ | 429 | 56 | 0.3744 | 2 | scarecrow-like protein 27 | KU879273 |
| S20 | F: GTGCTTTATGGCATCTTT<br>R: AACAGTGGTGCTTGAGT | $(AAG)_6$ | 285 | 56 | 0.5943 | 4 | septin and tuftelin-interacting protein 1 | KU879274 |
| S22 | F: AAACCTTCTCCCTTATCT<br>R: GGTTCGTTTGGACCTTTT | $(CTC)_4$ | 203 | 54 | 0.3738 | 2 | uncharacterized protein | KU879276 |
| S23 | F: CTCAAATCTTACGCATCT<br>R: GGTATTGTCCCATTGAGT | $(TC)_4$ | 282 | 56 | 0.0000 | 1 | receptor-like protein 12 | KU879277 |
| S24 | F: GTAGAAGAAGAGCAGCACA<br>R: CGGAGTAACTGAAGGGAG | $(CGC)_3$ | 384 | 56 | 0.6447 | 4 | uncharacterized protein | KU879278 |
| S25 | F: GCTGCTGTTCTGTTACGC<br>R: ATCTATCATCCACCTCCC | $(GAT)_3$ | 376 | 56 | 0.0434 | 2 | methionine–tRNA ligase | KU879279 |
| S26 | F: CTGTCCCAACTCCCAATA<br>R: GGGTAGGCTAAATAACTGA | $(CT)_6$ | 416 | 56 | 0.0000 | 1 | D-xylose-proton symporter-like 3 | KU879280 |
| S27 | F: AAAGACAATCCCAGAAG<br>R: CAAACTTGCTCTTCCTCC | $(AG)_4$ | 257 | 54 | 0.5480 | 4 | uncharacterized protein | KU879281 |
| S30 | F: CTCACAGATGTTTCCACCCA<br>R: TCCTACCCATTTCGCTCC | $(ACC)_4$ | 450 | 56 | 0.0000 | 1 | uncharacterized protein | KU879284 |
| S31 | F: TCAGGGTTCTCAGCATAA<br>R: AACCATCAGTGAGCCAA | $(TC)_5$ | 257 | 56 | 0.0000 | 1 | uncharacterized protein | KU879285 |
| S32 | F: AGGAAAGCGAACACGAAC<br>R: TAAATCCCAATCCAGCA | $(TG)_5$ | 366 | 56 | 0.2150 | 2 | 4-coumarate–CoA ligase 2 | KU879286 |
| S33 | F: AAGATTGAGCGTTATGTG<br>R: CTTACTTATGGAAGCACC | $(TGA)_6$ | 411 | 56 | 0.0000 | 1 | ribosomal RNA processing protein 1 B | KU879287 |
| S38 | F: AACGGCTCCAGTGATGTA<br>R: TGAAACAGGTGGTTGAGTA | $(CTG)_5$ | 283 | 56 | 0.6450 | 4 | ENTH/VHS family protein | KU879292 |

*Note*: *A* = number of alleles; PIC = polymorphism information content; $T_a$ = annealing temperature.

and then the libraries were sequenced using an Illumina HiSeq 2000 system (Illumina, San Diego, California, USA). We also downloaded the whole-genome data of one South Korean ginseng accession from GenBank (GenBank accession no.: SRR1181600). Because *P. ginseng* is an allotetraploid species with a large genome size (ca. 3 Gb), we were not able to perform de novo assembly in this study. Instead, we downloaded transcriptome data of the diploid species *P. notoginseng* (GenBank accession no.: SRX378873, SRX378878, and SRX378880) and performed de novo assembly using Trinity (Grabherr et al., 2011). These assembled transcripts were then used as references to perform the short read alignment for *P. ginseng* using Burrows-Wheeler Aligner's Smith-Waterman Alignment (BWA-SW; Li and Durbin, 2010). The insertions/deletions were reported using SAMtools (Li et al., 2009). Thereafter, we developed a series of Perl scripts to identify the polymorphic microsatellites from the obtained variant call format (VCF).

***Microsatellite marker data analysis*—**A total of 60 multiallelic microsatellites were detected from the whole-genome data of the 19 ginseng accessions (including SRR1181600). To determine the function of microsatellite-associated unigenes, these assembled transcripts were searched against the GenBank nonredundant protein database using BLASTX (Altschul et al., 1997) with an expected value $<10^{-7}$. The putative functions of these microsatellite-associated genes are listed in Table 1. We chose the candidate microsatellites according to the following criteria: (1) more than 20 bp at the flanking regions of the microsatellite repeat, and (2) no large intron (<500 bp in length) within the target region used to design primers. Forty-one microsatellites were selected to design the PCR primers, and 38 of them showed clean PCR amplifications in all five *P. ginseng* accessions tested. The PCR amplifications were conducted in a 25-μL volume with 2.5 mM $MgCl_2$, 0.1 μM forward and reverse primers, 400 μM dNTPs, 1 unit r*Taq* (TaKaRa Biotechnology Co., Dalian, Liaoning, China), and 20–50 ng DNA. PCRs were performed for each microsatellite under the following conditions: an initial denaturation step of 5 min at 95°C; followed by 35 cycles of 94°C for 30 s, annealing temperature (Table 1) for 30 s, and 72°C for 30 s; and a final step of 8 min at 72°C. After amplification, the PCR products were sequenced on an ABI730 sequencer (Applied Biosystems, Foster City, California, USA). The obtained genomic DNA sequences of *P. ginseng* were then compared with the assembled transcripts of *P. notoginseng*. As expected, all microsatellites were confirmed in the genomic sequences of *P. ginseng*. All DNA sequences of *P. ginseng* obtained from this study were submitted to GenBank (accession no.: KU879255–KU879294; Table 1).

To further evaluate the polymorphisms in cultivated ginseng, we amplified these microsatellites with 45 accessions of three major ginseng landraces. Twenty-one microsatellites yielded abundant PCR products across the 45 ginseng accessions. Fluorescently labeled PCR products were resolved to genotype on an ABI 3730 sequencer (Applied Biosystems). The number of alleles

TABLE 2. Allelic diversity in 21 microsatellites for three major ginseng landraces.[a]

| | SHIZHU | | | | | BIANTIAO | | | | | COMMON | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Locus | *N* | *A* | $H_o$ | $H_e$ | PIC | *N* | *A* | $H_o$ | $H_e$ | PIC | *N* | *A* | $H_o$ | $H_e$ | PIC |
| S2 | 15 | 1 | 0.000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S6 | 15 | 2 | 0.9375 | 0.4978 | 0.3741 | 15 | 2 | 1.0000 | 0.5000 | 0.3750 | 15 | 2 | 1.0000 | 0.5000 | 0.3750 |
| S11 | 15 | 2 | 0.1667 | 0.1528 | 0.1364 | 15 | 2 | 0.0714 | 0.0689 | 0.0739 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S12 | 15 | 2 | 1.0000 | 0.3750 | 0.3047 | 15 | 2 | 0.1429 | 0.1327 | 0.1217 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S13 | 15 | 4 | 1.0000 | 0.5531 | 0.4528 | 15 | 4 | 1.0000 | 0.5272 | 0.4288 | 15 | 4 | 0.9091 | 0.5695 | 0.4794 |
| S15 | 15 | 1 | 0.000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S16 | 15 | 2 | 0.9000 | 0.4950 | 0.3725 | 15 | 2 | 1.0000 | 0.5000 | 0.3750 | 15 | 2 | 1.0000 | 0.5000 | 0.3750 |
| S17 | 15 | 2 | 0.9000 | 0.4550 | 0.3515 | 15 | 2 | 1.0000 | 0.5000 | 0.3685 | 15 | 2 | 1.0000 | 0.4888 | 0.3685 |
| S19 | 15 | 2 | 0.8333 | 0.4861 | 0.3685 | 15 | 2 | 1.0000 | 0.4861 | 0.3701 | 15 | 2 | 1.0000 | 0.4988 | 0.3746 |
| S20 | 15 | 4 | 1.0000 | 0.6693 | 0.5995 | 15 | 3 | 1.0000 | 0.6505 | 0.5764 | 15 | 4 | 1.0000 | 0.6632 | 0.6000 |
| S22 | 15 | 2 | 0.8571 | 0.4841 | 0.3668 | 15 | 2 | 1.0000 | 0.5000 | 0.3750 | 15 | 2 | 1.0000 | 0.4994 | 0.3746 |
| S23 | 15 | 1 | 0.000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S24 | 15 | 3 | 0.4000 | 0.5600 | 0.4992 | 15 | 3 | 0.7500 | 0.6563 | 0.5786 | 15 | 4 | 0.7500 | 0.6939 | 0.6388 |
| S25 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 2 | 0.5000 | 0.1528 | 0.1364 |
| S26 | 15 | 1 | 0.000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S27 | 15 | 4 | 0.4667 | 0.5450 | 0.4757 | 15 | 4 | 0.9333 | 0.6339 | 0.5880 | 15 | 4 | 0.9091 | 0.6301 | 0.5673 |
| S30 | 15 | 1 | 0.000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S31 | 15 | 1 | 0.000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S32 | 15 | 2 | 0.2857 | 0.1327 | 0.1217 | 15 | 2 | 0.7857 | 0.3157 | 0.2688 | 15 | 2 | 0.8333 | 0.2698 | 0.2327 |
| S33 | 15 | 1 | 0.000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 | 15 | 1 | 0.0000 | 0.0000 | 0.0000 |
| S38 | 15 | 4 | 1.0000 | 0.7188 | 0.6675 | 15 | 4 | 1.0000 | 0.6806 | 0.6205 | 15 | 4 | 1.0000 | 0.7071 | 0.6559 |

*Note*: *A* = number of alleles; $H_e$ = expected heterozygosity; $H_o$ = observed heterozygosity; *N* = number of individuals used; PIC = polymorphism information content.

[a]Locality and voucher information are provided in Appendix 1.

of these microsatellites varied from one to five, and the polymorphism information content ranged from 0.0000 to 0.6450 (Table 1). We also evaluated polymorphisms of these microsatellites in each of the three ginseng landraces (Table 2). We noted that the polymorphic microsatellite S25 was monomorphic in the landraces SHIZHU and BIANTIAO. Similarly, the polymorphic microsatellites S11 and S12 were monomorphic in the landrace COMMON. The transferability of these primers was tested with *P. notoginseng*, *P. stipuleanatus*, *P. bipinnatifidus*, *P. quinquefolius*, and *A. elata*. All 21 primers amplified successfully in the four congeneric species except *P. bipinnatifidus* (Table 3). All of these microsatellites also yielded clear PCR products in the closely related species *A. elata*, suggesting the high transferability of these primers (Table 3). Notably, we found that seven microsatellites were monomorphic across all 45 ginseng accessions, but five of them were polymorphic in related species (Table 3).

## CONCLUSIONS

Development of molecular markers from nonmodel species has been increasing in recent years. In this study, we identified polymorphic microsatellite markers from the nonmodel species *P. ginseng* using whole-genome data. These polymorphic microsatellites provide useful molecular markers to assess the

TABLE 3. Cross-species amplification information for 21 microsatellite loci in closely related *Panax* and *Aralia* species.[a]

| | *P. quinquefolius* (*N* = 7) | | | *P. notoginseng* (*N* = 7) | | | *P. stipuleanatus* (*N* = 6) | | | *P. bipinnatifidus* (*N* = 1) | *A. elata* (*N* = 3) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Locus | *A* | $H_e$ | $H_o$ | *A* | $H_e$ | $H_o$ | *A* | $H_e$ | $H_o$ | *A* | *A* |
| S2 | 1 | 0.0000 | 0.0000 | 3 | 0.5612 | 1.0000 | 2 | 0.5000 | 1.0000 | 1 | 1 |
| S6 | 2 | 0.5000 | 1.0000 | 2 | 0.1327 | 0.1429 | 2 | 0.3200 | 0.0000 | 3 | 3 |
| S11 | 4 | 0.6020 | 0.8571 | 2 | 0.5000 | 1.0000 | 2 | 0.2449 | 0.3333 | 2 | 1 |
| S12 | 1 | 0.0000 | 0.0000 | 2 | 0.5000 | 1.0000 | 1 | 0.0000 | 0.0000 | NA | 2 |
| S13 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 2 | 3 |
| S15 | 2 | 0.5000 | 1.0000 | 1 | 0.0000 | 0.0000 | 3 | 0.5694 | 0.6667 | 2 | 1 |
| S16 | 3 | 0.4965 | 0.6667 | 2 | 0.4082 | 0.5714 | 2 | 0.4688 | 0.7500 | 3 | 2 |
| S17 | 3 | 0.6607 | 1.0000 | 2 | 0.5000 | 1.0000 | 2 | 0.3750 | 0.5000 | 2 | 2 |
| S19 | 1 | 0.0000 | 0.0000 | 2 | 0.5000 | 1.0000 | 2 | 0.2188 | 0.2500 | 3 | 2 |
| S20 | 3 | 0.5612 | 1.0000 | 3 | 0.6633 | 1.0000 | 1 | 0.0000 | 0.0000 | 1 | 2 |
| S22 | 2 | 0.0689 | 0.1429 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 2 | 2 |
| S23 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 2 | 0.2778 | 0.0000 | 1 | 1 |
| S24 | 2 | 0.5000 | 1.0000 | 5 | 0.7245 | 1.0000 | 2 | 0.5000 | 1.0000 | 4 | 3 |
| S25 | 1 | 0.0000 | 0.0000 | 2 | 0.4898 | 0.5714 | 1 | 0.0000 | 0.0000 | 3 | 2 |
| S26 | 1 | 0.0000 | 0.0000 | 2 | 0.1327 | 0.1429 | 2 | 0.2778 | 0.0000 | 1 | 2 |
| S27 | 2 | 0.0689 | 0.1429 | 2 | 0.4592 | 0.7143 | 3 | 0.5000 | 1.0000 | 2 | 3 |
| S30 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 1 | 3 |
| S31 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 1 | 1 |
| S32 | 2 | 0.2934 | 0.5714 | 2 | 0.4592 | 0.7143 | 2 | 0.5000 | 1.0000 | 2 | 3 |
| S33 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 1 | 1 |
| S38 | 3 | 0.5332 | 1.0000 | 1 | 0.0000 | 0.0000 | 1 | 0.0000 | 0.0000 | 1 | 1 |

*Note*: *A* = number of alleles; $H_e$ = expected heterozygosity; $H_o$ = observed heterozygosity; *N* = number of individuals used; NA = no PCR products.

[a]Locality and voucher information are provided in Appendix 1.

germplasm resources of *P. ginseng*. In particular, the high transferability of these microsatellites provides reliable molecular markers to investigate the population genetics and polyploid evolution of *Panax* and *Aralia*.

## LITERATURE CITED

Ahn, C. H., B. Kim, E. Yoon, and Y. Choi. 2009. Development of microsatellite markers to distinguish South Korean and Chinese ginseng. *Journal of Korean Forestry Society* 98: 568–575.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. H. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* 25: 3389–3402.

Choi, H. I., N. H. Kim, J. H. Kim, B. S. Choi, I. O. Ahn, J. S. Lee, and T. J. Yang. 2011. Development of reproducible EST-derived SSR markers and assessment of genetic diversity in *Panax ginseng* cultivars and related species. *Journal of Ginseng Research* 35: 399–412.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

Jarne, P., and P. J. Lagoda. 1996. Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* 11: 424–429.

Kim, J., B. H. Jo, K. L. Lee, E. Yoon, G. H. Ryu, and K. W. Chung. 2007. Identification of new microsatellite markers in *Panax ginseng*. *Molecules and Cells* 24: 60–68.

Lee, C., and J. Wen. 2004. Phylogeny of *Panax* using chloroplast *trnC–trnD* intergenic region and the utility of *trnC–trnD* in interspecific studies of plants. *Molecular Phylogenetics and Evolution* 31: 894–903.

Li, H., and R. Durbin. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics (Oxford, England)* 26: 589–595.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics (Oxford, England)* 25: 2078–2079.

Li, M. R., F. X. Shi, Y. X. Zhou, Y. L. Li, X. F. Wang, C. Zhang, X. T. Wang, et al. 2015. Genetic and epigenetic diversities shed light on domestication of cultivated ginseng (*Panax ginseng*). *Molecular Plant* 8: 1612–1622.

Liu, C. X., and P. G. Xiao. 1992. Recent advances on ginseng research in China. *Journal of Ethnopharmacology* 36: 27–38.

Reunova, G. D., O. G. Koren, T. I. Muzarok, and Y. N. Zhuravlev. 2014. Microsatellite analysis of *Panax ginseng* natural populations in Russia. *Chinese Medicine* 5: 231–243.

Shi, F. X., M. R. Li, Y. L. Li, P. Jiang, C. Zhang, Y. Z. Pan, B. Liu, et al. 2015. The impacts of polyploidy, geographic and ecological isolations on the diversification of *Panax* (Araliaceae). *BMC Plant Biology* 15: 297.

Van Dan, N., N. Ramchiary, S. R. Choi, T. S. Uhm, T. J. Yang, I. O. Ahn, and Y. P. Lim. 2010. Development and characterization of new microsatellite markers in *Panax ginseng* C.A. Meyer from BAC end sequences. *Conservation Genetics* 11: 1223–1225.

Yang, C. J., J. Wang, L. Q. Mu, S. C. Li, G. J. Liu, and C. Q. Hu. 2008. Development of an EST-SSR marker in *Panax ginseng*. *Chinese Journal of Agricultural Biotechnology* 5: 175–181.

Appendix 1. Locality information for *Panax ginseng* and the other species sampled in this study.

| Species | Accession | No. of accessions | Locality | Geographic coordinates | Voucher no.[a] |
|---|---|---|---|---|---|
| *Panax ginseng* C. A. Mey. | GL1 | 1 | North Korea | NA | |
| | GL2 | 1 | South Korea | NA | |
| | CB | 1 | Changbai, Jilin, China | 41°39′442″N, 127°35′229″E | |
| | WH | 1 | Dunhua, Jilin, China | 43°30′181″N, 127°54′193″E | |
| | TQ | 1 | Yanji, Jilin, China | 43°36′129″N, 129°35′807″E | |
| | FS | 1 | Fusong, Jilin, China | 42°24′216″N, 127°12′186″E | NENU20110718004 |
| | SZ1 | 1 | Kuandian, Liaoning, China | 40°45′595″N, 125°20′863″E | |
| | SZ2 | 1 | Kuandian, Liaoning, China | 40°47′157″N, 125°23′036″E | |
| | BT1 | 1 | Tonghua, Jilin, China | 41°18′492″N, 125°49′954″E | |
| | BT2 | 1 | Tonghua, Jilin, China | 41°05′245″N, 125°55′337″E | |
| | SJ1 | 1 | Songjiang, Jilin, China | NA | |
| | SJ2 | 1 | Songjiang, Jilin, China | NA | |
| | SJ3 | 1 | Songjiang, Jilin, China | NA | |
| | DP | 1 | Dapuchai, Jilin, China | NA | |
| | CY | 1 | Caiyuan, Jilin, China | NA | |
| | XL | 1 | Fusong, Jilin, China | NA | |
| | HR | 1 | Huanren, Liaoning, China | NA | |
| | KD | 1 | Kuandian, Liaoning, China | NA | |
| | EL | 1 | Russia | NA | |
| | BIANTIAO | 15 | Tonghua, Jilin, China | 41°18′492″N, 125°49′954″E | |
| | SHIZHU | 15 | Kuandian, Liaoning, China | 40°45′595″N, 125°20′863″E | |
| | COMMON | 15 | Dunhua, Jilin, China | 43°30′181″N, 127°54′193″E | |
| *Panax notoginseng* (Burkill) F. H. Chen ex C. Y. Wu & K. M. Feng | PN | 7 | Wenshan, Yunnan, China | NA | KUN0560433 |
| *Panax quinquefolius* L. | PQ | 7 | Wisconsin, USA | NA | NENU20110713001 |
| *Panax stipuleanatus* H. T. Tsai & K. M. Feng | PS | 6 | Pingbian, Yunnan, China | NA | |
| *Panax bipinnatifidus* Seem. | PB | 1 | Sichuan, China | 27°31′839″N, 101°42′569″E | NENU20120801001 |
| *Aralia elata* (Miq.) Seem. | M5 | 3 | Changbai, Jilin, China | NA | |

*Note*: NA = exact locations of these samples are unknown.

[a]The vouchers were deposited in the herbaria of Northeast Normal University (NENU), Changchun, Jilin, China, and Kunming Institute of Botany (KUN), Chinese Academy of Sciences, Kunming, Yunnan, China.