



Research article

Evidence of genome duplication revealed by sequence analysis of multi-loci expressed sequence tag–simple sequence repeat bands in *Panax ginseng* Meyer

Nam-Hoon Kim, Hong-Il Choi, Kyung Hee Kim, Woojong Jang, Tae-Jin Yang*

Department of Plant Science, Plant Genomics and Breeding Institute and Research Institute for Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul, Korea

ARTICLE INFO

Article history:

Received 8 October 2013
 Received in Revised form
 30 December 2013
 Accepted 31 December 2013
 Available online 18 January 2014

Keywords:

genome duplication
 multi-band pattern
Panax ginseng
 paralogs
 SSR marker

ABSTRACT

Background: *Panax ginseng*, the most famous medicinal herb, has a highly duplicated genome structure. However, the genome duplication of *P. ginseng* has not been characterized at the sequence level. Multiple band patterns have been consistently observed during the development of DNA markers using unique sequences in *P. ginseng*.

Methods: We compared the sequences of multiple bands derived from unique expressed sequence tag–simple sequence repeat (EST–SSR) markers to investigate the sequence level genome duplication.

Results: Reamplification and sequencing of the individual bands revealed that, for each marker, two bands around the expected size were genuine amplicons derived from two paralogous loci. In each case, one of the two bands was polymorphic, showing different allelic forms among nine ginseng cultivars, whereas the other band was usually monomorphic. Sequences derived from the two loci showed a high similarity, including the same primer-binding site, but each locus could be distinguished based on SSR number variations and additional single nucleotide polymorphisms (SNPs) or InDels. A locus-specific marker designed from the SNP site between the paralogous loci produced a single band that also showed clear polymorphism among ginseng cultivars.

Conclusion: Our data imply that the recent genome duplication has resulted in two highly similar paralogous regions in the ginseng genome. The two paralogous sequences could be differentiated by large SSR number variations and one or two additional SNPs or InDels in every 100 bp of genic region, which can serve as a reliable identifier for each locus.

Copyright © 2014, The Korean Society of Ginseng, Published by Elsevier. All rights reserved.

1. Introduction

Korean ginseng (*Panax ginseng*) is a renowned perennial herb that has long been used for medicinal purposes in East Asia [1]. *P. ginseng* has a large genome estimated to be more than 3 Gbp in size [2] and $2n = 48$ chromosomes [3]. Species belonging to the genus *Panax* have $2n = 24$ chromosomes or 48 chromosomes, so that the species with $2n = 48$ chromosomes have been regarded as tetraploids [4,5]. Genome duplication creates extra copies of the chromosome set, with the duplicated pairs in the genome defined as having paralogous relationships [6]. Our recent study using comparative analysis of expressed sequence tags (ESTs) [7] showed that *P. ginseng* and American ginseng (*Panax quinquefolius* L.)

concurrently experienced two rounds of genome duplication events based on the number of substitutions per synonymous site (K_s) of paralogous gene pairs. The more recent event is estimated to have occurred at $K_s = 0.02–0.04$, which corresponds to about 1.6–3.3 million years ago based on adopting a synonymous substitution rate of 6.1×10^{-9} substitutions/synonymous site/year [8]. However, genomic sequence-based clues and features have not yet been described to uncover the duplicated genome structure for *P. ginseng*.

We have developed large numbers of simple sequence repeat (SSR) markers designed from ESTs and genomic sequences for mapping and cultivar authentication. When we amplified ginseng genomic DNA with SSR markers, we observed multiple bands from

* Corresponding author. Department of Plant Science, Plant Genomics and Breeding Institute and Research Institute for Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, 200-4119, 599 Gwanangno, Gwangak-gu, Seoul 151-921, Korea.

E-mail address: tjyang@snu.ac.kr (T.-J. Yang).

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

almost all of the primer pairs [9,10]. These phenomena cannot be abolished by changing polymerase chain reaction (PCR) conditions and extending primer length. In other reports on ginseng SSR markers, the number of alleles ranged from two to nine and the observed heterozygosity of markers is usually greater than 0.5 [11–13]. These results show that multiple bands are consistently generated with ginseng genomic DNA; whether the multiple bands originate from different loci or the same locus can be confusing. For instance, two bands appearing in one cultivar could be misinterpreted as representing a heterozygous form even though they were derived from two independent loci. Meanwhile, chloroplast genome sequence-based markers produced clear single bands from ginseng genomic DNA [14], which may indicate that the recently duplicated nuclear genome causes multiple bands to be coincidentally amplified by the same primer set.

This study was conducted to examine whether the multiple band patterns of PCR products are associated with the genome duplication of *P. ginseng*. We sequenced SSR bands produced by five EST-SSR markers that were previously selected as the best and most clearly polymorphic SSR markers to authenticate ginseng cultivars in a screening of more than 200 SSR markers [10]. Sequence comparisons of SSR bands derived from multiple loci and multiple alleles showed the sequence level differences in the duplicated genome and thus promoted our understanding of genomics and whole genome sequencing of *P. ginseng*.

2. Materials and methods

2.1. Sample preparation and preliminary PCR

Leaf samples of six ginseng cultivars (Chunpoong, Yunpoong, Sunpoong, Sunone, Sunun, and Gopoong) were collected from a research field of Seoul National University, Suwon, Korea. The total DNA of the samples was extracted by modified cetyltrimethylammonium bromide methods [15].

Five EST-SSR markers (gm47n, gm45n, gm129, gm175, and gm184) that have shown clear polymorphism among Korean ginseng cultivars in previous work [9,10] were used for amplification in several cultivars showing different genotypes.

2.2. Sequencing of each individual band among multiple bands

Amplicons were visualized on polyacrylamide gels by UV lamp and individual bands from cultivars were excised and eluted. PCR was conducted again using the same primer set with the eluted product from each band as the DNA template. The final PCR product for each band was purified with the Inclone Gel & PCR purification kit (IN1002-0200, Seoul, Korea) and sequenced using an ABI3730xl DNA analyzer at the National Instrumentation Center for Environmental Management, Seoul, Korea. A high quality sequence for each band was determined by alignment of more than two duplicated forward and reverse sequences.

2.3. Sequence analysis and design of locus-specific primer for the gm47n marker

High-quality sequences of Band-A (the smallest band in Fig. 1) and Band-B (the second smallest band in Fig. 1) derived from different cultivars were aligned for every marker using the CLUSTALW program with default setting in MEGA5 [16]. Sequence differences such as SSRs, SNPs, and InDels were manually inspected based on multiple sequence alignments with the original EST.

A locus-specific left primer was newly designed from the region showing an SNP between Band-A and Band-B sequences of the gm47n marker by a modified method with an additional base

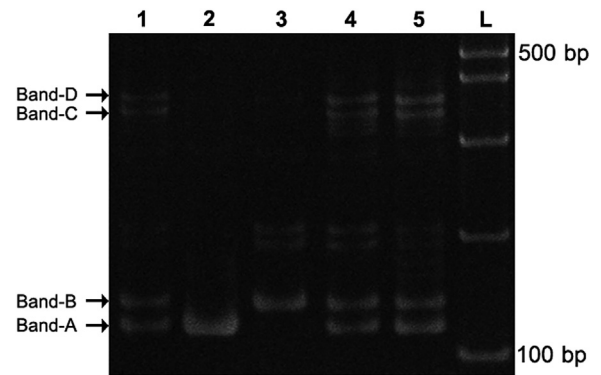


Fig. 1. Band profiles of the gm47n simple sequence repeat (SSR) marker using genomic DNA and eluted DNA from each band. Each of the four bands amplified using the gm47n primer pair from Chunpoong genomic DNA (lane 1) was eluted from the acrylamide gel. Bands-A–D are named based on their relative size and the eluted DNA for each was used as a template for reamplification (lanes 2–5, respectively). L, 100 bp DNA ladder.

change [17]. The SNP was common to all cultivars. With the new left and the original right primer, PCR was performed using genomic DNA from nine cultivars (Chunpoong, Yunpoong, Sunpoong, Gumpoong, Gopoong, Sunun, Cheongsun, Sunhyang, and Sunone). One individual plant was analyzed for each cultivar. These primer pairs were also applied to 11 individual plants of F₂ populations between Yunpoong and Chunpoong. Electrophoresis was conducted using a fragment analyzer (Advanced Analytical Technologies, Marco Island, FL, USA).

3. Results

3.1. Separation and reamplification of individual bands

In previous work, five EST-SSR markers (gm47n, gm45n, gm129, gm175, and gm184) that showed clear polymorphism among Korean ginseng cultivars were identified [9,10]. However, all five markers produced more than two bands for each cultivar. Therefore these same five markers were selected for this study and used for amplification in several cultivars showing different genotypes. The PCR products amplified by the five markers exhibited four bands in gel electrophoresis. Among the four bands, two lower bands (Band-A and Band-B in Fig. 1) were similar to the expected size, whereas the upper two bands (Band-C and Band-D in Fig. 1) were much larger than the expected size [10]. After elution and reamplification of each band, the two lower bands each produced a single amplicon that was the same size as the original band (lanes 2 and 3 in Fig. 1), whereas the amplicons from the upper bands appeared as multiple bands including Band-A and Band-B (lanes 4 and 5 in Fig. 1). This result indicates that these unexpected larger bands are modified forms of Band-A and Band-B. This phenomenon was common to all five markers and we conclude that only the two lower bands of the expected size (Band-A and Band-B) were *bona fide* PCR amplicons. Therefore all Band-A and Band-B products appearing in different cultivars were selected for further analysis.

3.2. Sequence comparison of Band-A and Band-B

We sequenced the two lower bands, Band-A and Band-B, derived from different cultivars showing different genotypes for each of the five markers. Two representative cultivars, Chunpoong and Yunpoong, were sequenced for all five markers and other cultivars were also sequenced, including Sunpoong for the gm47n marker, Sunun for the gm129 marker, Sunone for the gm175 marker, and Sunpoong, Sunone, and Gopoong for the gm184

marker. A total of 34 high-quality sequences derived from individual bands was obtained.

Multiple sequence comparison allowed us to classify the multiple bands as representing different loci in the same cultivar (paralogs) or allelic forms of the same locus in different cultivars (alleles; Fig. 2). The bands close to the expected size (Band-B of gm45n, gm47n, and gm175 and Band-A of gm129 and gm184) were derived from same locus as the reported EST. The other bands (Band-B of gm45n, gm47n, and gm175 and Band-A of gm129 and gm184) amplified from a paralogous locus showed relatively different sizes from those expected.

The paralogous sequences were characterized by SNP or InDel variations as well as much larger variations in SSR unit number. For example, the gm175 marker showed polymorphism for both loci among cultivars. Each of the two bands showed one or two copy differences of the AGG SSR motif among cultivars. There was a maximum copy number difference of four for the AGG SSR motif as well as a 21 bp InDel variation between Band-A and Band-B (Table 1). The Band-B sequence of Chunpoong corresponded to the EST, indicating that the EST is derived from the locus of Band-B (Fig. 2A). The gm45n marker showed a maximum copy number difference of five

for the TGG SSR motif, (TGG)₅ and (TGG)₁₀, as well as two SNPs between Band-A and Band-B. The allelic form of Band-B showed only a two-copy difference for the TGG SSR motif, (TGG)₈ and (TGG)₁₀, in Chunpoong and Yunpoong cultivars, respectively (Table 1). By contrast, Band-A showed no variation among the different cultivars. Similarly, only one of the two bands, Band-B, was polymorphic among cultivars, except for the gm175 marker. Among the five markers, four had SNPs and the other had an InDel between Band-A and Band-B that served as a signature to distinguish paralogous sequences (Fig. 2, Table 1).

3.3. Locus-specific PCR amplification for the gm47n marker

We next tried to develop locus-unique markers to amplify selectively single bands derived from one of two paralogous regions. We focused on the SNP regions between paralogous sequences. The gm47n marker showed a more than four SSR unit difference as well as one SNP between Band-A and Band-B (Fig. 2B). The SNP was identified at the position 51 bp as “C” and “T” for Band-A and Band-B, respectively (Table 1). For the polymorphic Band-B-specific primer, we designed an additional left primer, 5’-

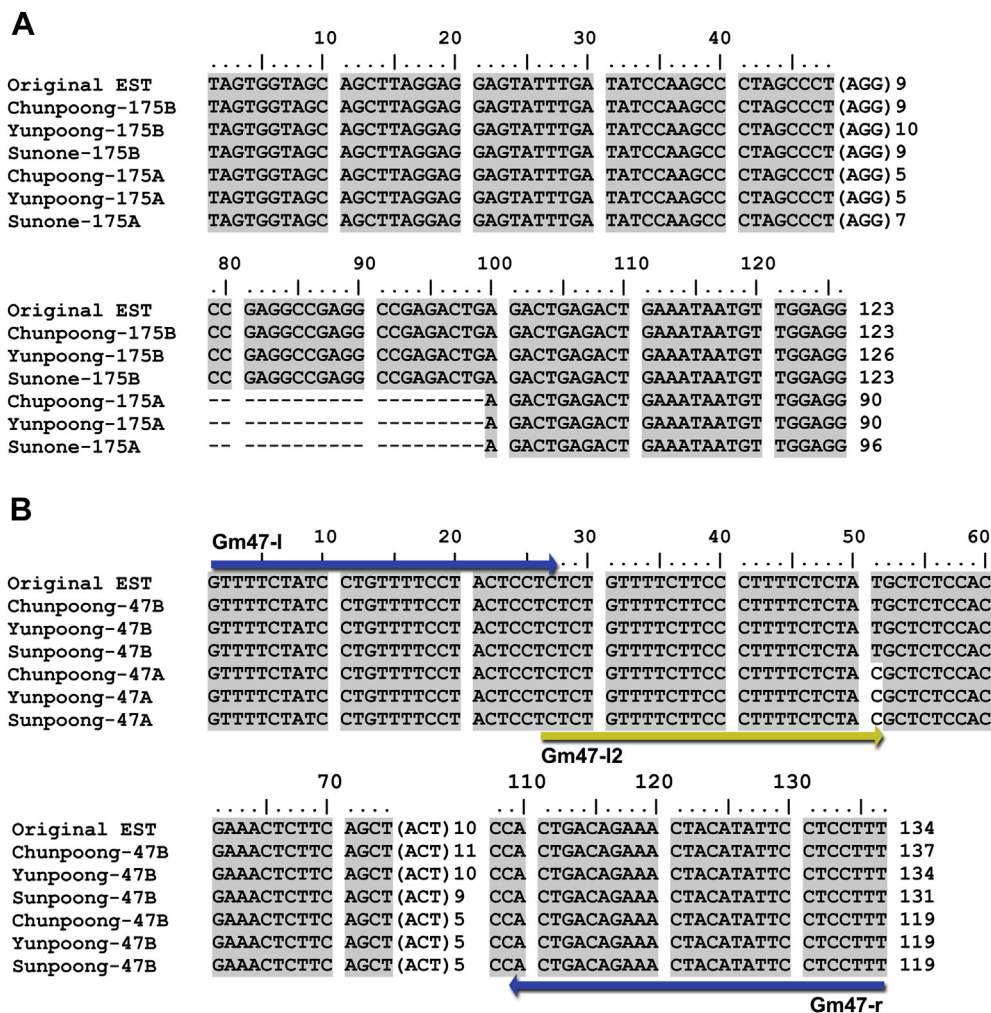


Fig. 2. Multiple sequence alignment of eluted bands derived from different cultivars. (A) Alignment of the original expressed sequence tag (EST) sequence and six sequences, two bands from three cultivars, for the gm175 marker. Band-A sequences show a copy number variation of the simple sequence repeat motif: (GGA)₄ for Chunpoong and Yunpoong and (GGA)₆ for Sunone. Band-B sequences also show a copy number variation: (GGA)₈ for Chunpoong and Sunone and (GGA)₉ for Yunpoong. All Band-A sequences show an additional 21 bp deletion compared with the EST and Band-B sequences. (B) The gm47n markers show repeat number differences ranging from four to six and one single nucleotide polymorphism at position 51 between Band-A and Band-B sequences of three cultivars. To amplify Band-B-specific polymerase chain reaction products, a new left primer (Gm47-l2, indicated with a yellow bar) was designed with a single nucleotide polymorphism at the end. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Summary of the Variation Among Expressed Sequence Tags and Amplified Sequences

Marker	Sequence source	SSR (pos ¹ /ssr motif)	Polymorphic cultivars ²	SNP (pos/base)	InDel (pos/InDel)	Size (bp)
gm45n	Original EST ³	65/(TGG) ₁₀		37/G, 42/A		120
	Band-A	65/(TGG) ₅	No polymorphism	37/A, 42/T	—	99
	Band-B ⁴	65/(TGG) _{8,10}	CP, YP	37/G, 42/A	—	114–120
gm47n	Original EST	75/(ACT) ₁₀		51/T		131
	Band-A	75/(ACT) ₅	No polymorphism	51/C	—	110
	Band-B	75/(ACT) _{9,10,11}	SP, YP, CP	51/T	—	128–134
gm129	Original EST	42/(TAG) ₆		55/A		99
	Band-A	42/(TAG) ₆	No polymorphism	55/A	—	99
	Band-B	42/(TAG) _{12,13,14}	YP, CP, SU	55/T	—	115–121
gm175	Original EST	50/(AGG) ₉			74/21 bp Ins	123
	Band-A	50/(AGG) _{5,7}	CP, SO	—	74/21 bp Del	90–96
	Band-B	50/(AGG) _{9,10}	CP, YP	—	74/21 bp Ins	123–126
gm184	Original EST	33/(GAA) ₆		32/G, 51/G		122
	Band-A	33/(GAA) ₆	No polymorphism	32/G, 51/G	—	122
	Band-B	33/(GAA) _{14,17,18,20,29}	GO, CP, YP, SP, SO	32/A, 51/C	—	148–166

CP, Chunpoong; EST, expressed sequence tag; GO, Gopoong; SNP, single nucleotide polymorphism; SO, Sunone; SP, Sunpoong; SSR, simple sequence repeat; SU, Sunun; YP, Yunpoong.

¹ Nucleotide position based on amplified region of the original EST sequence.

² Cultivars are listed in the same order as in the SSR column. "No polymorphism" indicates that all cultivars had same number of SSRs.

³ Original EST sequence that was the source of marker development.

⁴ Underlining indicates the same locus as the original EST.

CTCTGTTTTCTCCCTTTTCTCTGT-3', which has the Band-B specific nucleotide "T" at the end and an additional modified nucleotide "G" (Fig. 2B). The Band-B-unique primer set amplified one clear single band of 105–111 bp that showed 3 bp and 6 bp variation among nine cultivars (Fig. 3). This result is in accord with the sequence analysis showing one and two copy variations in ACT SSR motifs, (ACT)₉, (ACT)₁₀, and (ACT)₁₁, for Chunpoong, Yunpoong, and Sunpoong cultivars, respectively (Fig. 2B). In addition, the locus-specific marker was applied to the F₂ population of a cross between Chunpoong and Yunpoong, and the two 3 bp different polymorphic bands segregated according to a Mendelian single gene pattern (Fig. 4).

4. Discussion

4.1. PCR amplification of two paralogous loci derived from genome duplication in *P. ginseng*

Multiple bands were always obtained when we tried to develop SSR markers using over 1000 EST and genome sequence-based SSR primers in *P. ginseng*. Although long sequence-specific primers with a minimum length of 24 nucleotides to enhance their specificity were designed, the multiple bands were consistently produced in PCR amplification from ginseng genomic DNA [9,10], hindering our progress in genetic mapping because of unclear genotyping. In this study, we demonstrated that the multiple bands were derived from

simultaneous amplification of paralogous loci. Sequence analysis revealed that two bands near the expected size for a given marker were amplified from two different loci with a paralogous relationship. The recent genome duplication event in *P. ginseng* was estimated at $K_s = 0.02–0.04$, which is markedly later than that of other plant species such as soybean ($K_s = 0.10–0.15$) [18], maize ($K_s = 0.15–0.20$) [18], apple ($K_s = 0.15–0.20$) [19], and poplar ($K_s = 0.20–0.30$) [20]. It appears that the recentness of the genome duplication in the ginseng genome has resulted in the paralogous loci being very similar to one another and able to be amplified by the same primer pair.

4.2. Sequence variation between paralogous bands

The paralogous sequences between Band-A and Band-B products were characterized by SNP or InDel variation as well as much larger variation SSR unit numbers. SNPs or InDels were clearer signatures to distinguish individual loci than SSR unit differences because the SNP or InDel variations were observed only between paralogous loci (Table 1). This implies that a minor sequence variation can serve as a major index to discriminate each paralogous locus. Comparison of five paralogous sequences revealed that one or two SNPs or InDel variations existed within every 100 bp of genic region between paralogous loci. Sequence variation among different alleles of different cultivars was identified only in the SSR length, with no additional SNP or InDel variation. This indicates that there is very low genetic variation among Korean ginseng cultivars, probably because of its short breeding history [21]. Our analysis was conducted for small PCR fragments of around 100–200 bp derived from genic regions. Further comparison of sequence-level syntenic relationships between long paralogous sequences will be required for clarifying the overall structure of the duplicated ginseng genome.

4.3. Additional artifact bands are caused by the coexistence of two paralogous bands

At least four bands were usually obtained in PCR amplification from genomic DNA: two lower genuine bands and two upper bands. The upper bands were comparatively larger than expected and were revealed to be artifacts in PCR amplification (Band-C

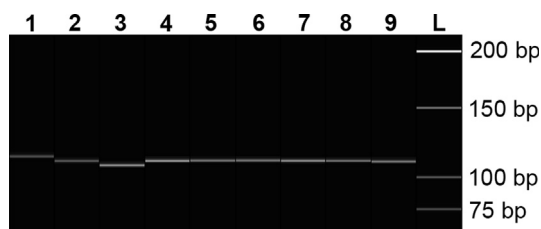


Fig. 3. Polymerase chain reaction amplification with the gm47n Band-B locus-specific primer. Genomic DNAs of nine cultivars were used for the polymerase chain reaction and capillary electrophoresis was conducted. Single bands were amplified in all nine samples and clear polymorphism were detected among cultivars. 1, Chunpoong; 2, Yunpoong; 3, Sunpoong; 4, Gumpoong; 5, Gopoong; 6, Sunun; 7, Cheongsun; 8, Sunhyang; 9, Sunone; L, DNA ladder.

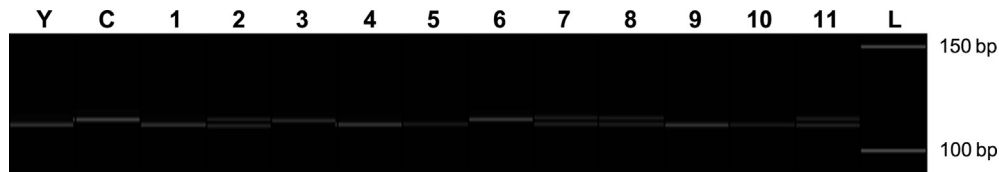


Fig. 4. Application of locus-specific primer of the gm47n marker into F_2 population crossed by Yunpoong and Chunpoong. Polymerase chain reaction amplification of F_2 individual plants between Yunpoong and Chunpoong was conducted with gm47n Band-B locus-specific primer. Lanes with Y, C, and L denote Yunpoong, Chunpoong, and 100 bp ladder, respectively, and 1–11 are individual plants of F_2 population.

and Band-D in Fig. 1). Multiple band artifacts are common in the amplification of SSR sequences and can generally be removed by modifying the PCR conditions or the number of cycles [22–24]. However, the examples described here are fundamentally different from the previous reports for two reasons: firstly, the larger bands could not be abolished by altering the PCR conditions; and, secondly, the reamplification of the larger bands showed the same band patterns as that of preliminary PCR amplification. Based on all these findings together, it appears that the artifact bands are derived from heteroduplexes created by the combining and interruption between coexisting different amplicons.

4.4. Development of locus-specific markers and future applications

The appearance of multiple bands in PCR products has been regarded as one of the more serious obstacles to marker development and genetic mapping for recently duplicated plant genomes such as rapeseed (*Brassica napus*) [25] and *P. ginseng* [9] because they hinder genotyping against the mapping population as well as the authentication of cultivars. In this study, a clear single band was successfully amplified by using a locus-specific primer designed on the basis of sequence variation between the two paralogous loci. The locus-specific primer was based on the SNP sequence of the polymorphic band of the gm47n marker. In addition to the SNP, T/C in Band-B and Band-A, we added another modified nucleotide, “G” instead of “A”, that resulted in a clear single band of PCR product in ginseng, as suggested by a previous report [17]. The clear single band was polymorphic between two cultivars, Chunpoong and Yunpoong, and segregated with a Mendelian single gene pattern in their F_2 population (Fig. 4). These results support our assumption that Band-A and Band-B are not heterozygous forms, but instead are derived from different loci created by the recent genome duplication of *P. ginseng* [7]. Our method can be applied to other markers to overcome the genotyping difficulty caused by multiple bands in *P. ginseng*. Most plant species have undergone a few rounds of genome duplication [26,27]. We suggest that this approach should be considered as an efficient method to avoid the misinterpretation of multiple band appearances in genome research on wild plant resources that may have undergone recent genome duplication.

Utilization of upcoming ginseng genome sequence information will be a powerful tool for the development of indisputable and reliable markers and genetic mapping in *P. ginseng*. We are conducting whole-genome sequencing for the cultivar Chunpoong using the Illumina platform [28] and have identified many long paralogous genome sequence pairs from the draft sequence assembly. Each of the paralogous sequences can be mapped by developing paralogous locus-specific markers as suggested in this study. Although this study was based on very short sequences, our findings promote the ongoing ginseng whole-genome sequence assemblies.

Conflicts of interest

All contributing authors declare no conflicts of interest.

Acknowledgments

This study was supported by the Next-Generation BioGreen21 Program (No. PJ008202), Rural Development Administration, Korea.

References

- [1] Yun TK. Brief introduction of *Panax ginseng* C.A. Meyer. J Korean Med Sci 2001;16:S3–5.
- [2] Hong CP, Lee SJ, Park JY, Plaha P, Park YS, Lee YK, Choi JE, Kim KY, Lee JH, Lee J, et al. Construction of a BAC library of Korean ginseng and initial analysis of BAC-end sequences. Mol Genet Genomics 2004;271:709–16.
- [3] Waminal NE, Park HM, Ryu KB, Kim JH, Yang TJ, Kim HH. Karyotype analysis of *Panax ginseng* C.A.Meyer, 1843 (Araliaceae) based on rDNA loci and DAPI band distribution. Comp Cytogen 2012;6:425–41.
- [4] Choi HK, Wen J. A phylogenetic analysis of *Panax* (Araliaceae): integrating cpDNA restriction site and nuclear rDNA ITS sequence data. Plant Syst Evol 2000;224:109–20.
- [5] Choi HW, Koo DH, Bang KH, Paek KY, Seong NS, Bang JW. FISH and GISH analysis of the genomic relationships among *Panax* species. Genes Genomics 2009;31:99–105.
- [6] Doyle JJ, Egan AN. Dating the origins of polyploidy events. New Phytol 2010;186:73–85.
- [7] Choi HI, Kim NH, Lee JK, Choi BS, Kim KD, Park JY, Lee SC, Yang TJ. Evolutionary relationship of *Panax ginseng* and *P. quinquefolius* inferred from sequencing and comparative analysis of expressed sequence tags. Genet Resour Crop Evol 2013;60:1377–87.
- [8] Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. Science 2000;290:1151–5.
- [9] Choi HI, Kim NH, Kim JH, Choi BS, Ahn IO, Lee JS, Yang TJ. Development of reproducible EST-derived SSR markers and assessment of genetic diversity in *Panax ginseng* cultivars and related species. J Ginseng Res 2011;35:399–412.
- [10] Kim NH, Choi HI, Ahn IO, Yang TJ. EST-SSR marker sets for practical authentication of all nine registered ginseng cultivars in Korea. J Ginseng Res 2012;36:298–307.
- [11] Kim J, Jo BH, Lee KL, Yoon ES, Ryu GH, Chung KW. Identification of new microsatellite markers in *Panax ginseng*. Mol Cell 2007;24:60–8.
- [12] Ma KH, Dixit A, Kim YC, Lee DY, Kim TS, Cho EG, Park YJ. Development and characterization of new microsatellite markers for ginseng (*Panax ginseng* C.A. Meyer). Conserv Genet 2007;8:1507–9.
- [13] Dan NV, Ramchiary N, Choi SR, Uhm TS, Yang TJ, Ahn IO, Lim YP. Development and characterization of new microsatellite markers in *Panax ginseng* (C.A. Meyer) from BAC end sequences. Conserv Genet 2010;11:1223–5.
- [14] Kim JH, Jung JY, Choi HI, Kim NH, Park JY, Lee Y, Yang TJ. Diversity and evolution of major *Panax* species revealed by scanning the entire chloroplast intergenic spacer sequences. Genet Resour Crop Evol 2013;60:413–25.
- [15] Allen GC, Flores-Vergara MA, Krasnyanski S, Kumar S, Thompson WF. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat Protoc 2006;1:2320–5.
- [16] Kumar S, Nei M, Dudley J, Tamura K. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 2008;9:299–306.
- [17] Kwok S, Kellogg DE, McKinney N, Spasic D, Goda L, Levenson C, Sninsky JJ. Effects of primer template mismatches on the polymerase chain-reaction—human-immunodeficiency-virus type-1 model studies. Nucleic Acids Res 1990;18:999–1005.
- [18] Blanc G, Wolfe KH. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell 2004;16:1667–78.
- [19] Sanzol J. Dating and functional characterization of duplicated genes in the apple (*Malus domestica* Borkh.) by analyzing EST data. BMC Plant Biol 2010;10:87.
- [20] Sterck L, Rombauts S, Jansson S, Sterky F, Rouzé P, Van de Peer Y. EST data suggest that poplar is an ancient polyploid. New Phytol 2005;167:165–70.

- [21] Lee JS, Lee SS, Lee JH, Ahn IO. Effect of seed size and cultivars on the ratio of seed coat dehiscence and seedling performance in *Panax ginseng*. *J Ginseng Res* 2008;32:257–63.
- [22] Walsh PS, Fildes NJ, Reynolds R. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res* 1996;24:2807–12.
- [23] Fernando P, Evans BJ, Morales JC, Melnick DJ. Electrophoresis artefacts—a previously unrecognized cause of error in microsatellite analysis. *Mol Ecol Notes* 2001;1:325–8.
- [24] Bovo D, Rugge M, Shiao YH. Origin of spurious multiple bands in the amplification of microsatellite sequences. *Mol Pathol* 1999;52:50–1.
- [25] Wells R, Trick M, Fraser F, Soumpourou E, Clissold L, Morgan C, Pauquet J, Bancroft I. Sequencing-based variant detection in the polyploid crop oilseed rape. *BMC Plant Biol* 2013;13:111.
- [26] Bowers JE, Chapman BA, Rong JK, Paterson AH. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 2003;422:433–8.
- [27] Yang TJ, Kim JS, Kwon SJ, Lim KB, Choi BS, Kim JA, Jin M, Park JY, Lim MH, Kim HI, et al. Sequence-level analysis of the diploidization process in the triplicated *Flowering Locus C* region of *Brassica rapa*. *Plant Cell* 2006;18:1339–47.
- [28] Choi HI, Kim NH, Jung JY, Park HM, Park HS, Park JY, Lee JH, Park JS, Lee JK, Choi BS, et al. Current status of Korean ginseng (*Panax ginseng*) genome mapping and sequencing. In: Yang DC, Kim SK, Oh HI, editors. *Advances in ginseng research—proceedings of the 10th international symposium on ginseng: September 13–16. Seoul, Korea: The Korean Society of Ginseng; 2010. pp. 762–78.*