A new nuclear DNA marker from ubiquitin ligase gene region for genetic diversity detection of walnut germplasm resources

Zhili Suoa,*, Lingna Chenb, Dong Peib,*, Xiaobai Jinc, Huijin Zhangc

a State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, 20 Nanxincun, Haidian District, Beijing 100091, China
b State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Yiheyuan Hou, Beijing 100091, China
c Beijing Botanical Garden, Institute of Botany, Chinese Academy of Sciences, 20 Nanxincun, Haidian District, Beijing 100093, China

ABSTRACT

Development of more sensitive nuclear DNA markers for identification of species, particularly closely allied taxa has been a challenging task that has attracted interest from scientists in fields of biotechnological development and genetic diversity detection. In this study, the sequence of the ubiquitin ligase gene (UBE3) region of nuclear DNA was tested for applicability and efficacy in revealing genetic diversity of walnut resources, with an emphasis on inter- and intra-specific levels. Analysis on genetic relationship among the taxa was conducted with the neighbor-joining (NJ) method. The number of variable bases in the UBE3 region was 20 sites. All nine taxa (species/variety/cultivars) were distinguished using the UBE3 sequence. In addition, each taxon was characterized molecularly with a unique nucleotide molecular formula using ten variable base sites derived from the nuclear DNA UBE3 gene sequence. This study presents a good complementary methodology for developing new DNA markers for identification of genus Juglan.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

1. Introduction

The internal transcribed spacer (ITS) sequence of nuclear ribosomal DNA with bi-parental inheritance is currently used widely to determine the genetic diversity of land plants [1–10]. However, one limitation of species identification using the ITS sequence is that the method has limited resolution in identifying species, especially within closely related taxa [1–11]. For instance, the discriminating power of the four recommended DNA markers at the species level in Alnus (Betulaceae) was 10% (rbcL), 31.25% (matK), 63.6% (trnH-psbA), and 76.9% (ITS) [3]. Among the four DNA regions (rbcL, matK, trnH-psbA, and ITS), the ITS sequence has the most variable information, and appears to have limited power to discriminate closely related taxa in Juglandaceae [5]. Hanabusaya and Adenophora sect. Remitiflora of the family Campanulaceae could not be resolved using ITS sequences [6]. As a result of insufficient morphological information and DNA markers, the development of scientific research has been severely hindered in fields such as taxonomy, ecology, and genetic resource evaluation. Development of new and more sensitive nuclear DNA markers for biodiversity detection is highly desirable [11,12].

The ubiquitin–proteasome system, which plays a key role in degradation of proteins, is imperative for maintaining the cellular homeostasis in eukaryotic cells [13]. Three enzymes are required in the ubiquitination and targeting of proteins for degradation: ubiquitin activating enzyme E1, ubiquitin conjugating enzyme E2, and the highly conserved ubiquitin ligase E3. Ubiquitin ligases are key components of the ubiquitin–proteasome system. After a protein has been ubiquitinated, the substrate protein will be located to the proteasome (a cylindrical complex) to be degraded into smaller polypeptides or other molecules with biological activity for reutilization [13].

Ubiquitin plays a vital role not only in protein degradation but also in many cellular functions including DNA repair processes, cell cycle regulation, cell growth, immune system functionality, and hormone-mediated signaling in plants. In recent years, several
types of the ubiquitin ligase E3, such as RING-finger-containing E3s (RBR and TRIM families), cullin-containing E3 complexes, Ubox E3s and HECT E3s, were reported [14]. However, few reports are available concerning development of nuclear DNA markers from the genomic regions in relation to the ubiquitin–proteasome system.

In this study, we used plant material from Juglandaceae to develop a new nuclear DNA marker within the ubiquitin ligase gene (UBE3) region to discriminate the representative samples (species/varieties/cultivars) of the genus Juglans. Our objectives were: (i) to test the applicability of the nuclear DNA marker from the UBE3 gene region; and (ii) to evaluate the resolution ability of the nuclear DNA marker from the UBE3 gene region. The results of this effort show that UBE3 is sensitive for characterizing genetic diversity in the family Juglandaceae.

2. Materials and methods

2.1. Plant materials

Nine representative taxa of the genus Juglans and two outgroups (Cyclocarya paliurus and Pterocarya stenoptera in Juglandaceae) were used in this study (Table 1). The eleven taxa were sampled from three places: the resources nursery (N 34° 18′, E 111° 30′) of Forestry Bureau of Lunoning County, Henan Province, China; the Arboretum (N 25° 08′, E 102°45′) of Forestry Academy of Yunnan Province, Kunming City, Yunnan, China; and Beijing Botanical Garden (N 39° 48′, 116° 28′) under the Institute of Botany, Chinese Academy of Sciences, Beijing, China. All necessary permits for the collection of fresh leaves from the trees growing at each place were acquired prior to material collection. All collected material was verified by a taxonomic expert. Fresh leaves of each accession were collected in the spring and dried immediately using silica gel for future DNA extraction.

2.2. DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted using the Plant Genomic DNA Kit (DP305) from Tiangen Biotech (Beijing) Co., Ltd. China. The nuclear DNA UBE3 gene locus was amplified using the primer pair H_UBE3_23F (5′-TGGCCTCAAGTTCAGT-3′) and H_UBE3_838R (5′-CTCCCCATAGGTGTTCTCC-3′). Taq DNA polymerase and PCR buffer (TaKaRa Code: DR100B) were from TaKaRa Biotechnology Co., Ltd. (Dalian, China). The PCR protocol were as follows: preheating at 94 °C for 4 min, 34 cycles at 94 °C for 30 s, 54 °C for 45 s, and final extension at 72 °C for 10 min. PCR amplification of the regions of interest was performed in an Applied Biosystems VeritiTM 96-Well Thermal Cycler (Model#: 9902, made in Singapore). The amplicons were resolved simultaneously on 2% agarose gels (Promega, the USA) run in 1 x TAE buffer at 3 V cm⁻¹ for 2.5 h and were stained with ethidium bromide. The fragments (PCR products) were directly sequenced with the same primer pair mentioned above using a 3730xl DNA analyzer (Applied Biosystems, Foster City, CA, USA).

2.3. Data analysis

The DNA sequences were aligned with ClustalX [15] and then were manually confirmed using Sequencher (v.4.6) software. Sequence haplotype diversity was calculated using DnaSP (DNA Sequences Polymorphism version 5.10.01) software [16]. Sequence datasets were analyzed using Mega 6 software [17]. Tamura’s 3-parameter model was found to be the best nucleotide substitution model for the UBE3 sequences with Mega 6. The analysis on genetic relationship among the taxa was conducted in Mega 6 using neighbor-joining (NJ) method. Bootstrapping was done with 1000 replicates. All positions containing gaps and missing data were treated with pairwise deletion option. The sequences of the accessions were deposited in GenBank (GenBank accession numbers: KF994007-KF994018).

3. Results

3.1. Identification and confirmation of the DNA sequences

The nuclear DNA UBE3 sequences (Fig. S1) obtained in this study were confirmed to be part of the coding region of the UBE3 gene, as shown using BLAST to compare the sequence to that in the GenBank database. Additional comparisons were made using the E3 ubiquitin–protein ligase sequences of Prunus mume (GenBank accession no. XM_003537762.2) and several other plant species (GenBank accession no. XM_007199611.1, XM_003537761.2, XM_004505735.1, and XM_003607148.1). At least three independent samples from each species or cultivar have been sequenced; identical results were obtained. Therefore, only one sample data was used to represent each taxon.

3.2. Genetic variation of the UBE3 region among the taxa

According to the UBE3 sequence dataset, eight variable base sites were detected at inter- and intra-specific levels. These variable sites were found at nucleotide positions 46, 125, 205, 227, 459, 562, 595 and 663, and represent 40% of the total variable sites detected. The closely related taxa within each section are 70% of the total between sections and 5% of the total either within each section or cultivar. Therefore, only one sample data was used to represent each taxon.

**Table 1** Materials used in this study.

| No. | Taxon                      | Rank  | Section         | Place of collection |
|-----|----------------------------|-------|-----------------|--------------------|
| 1   | Juglans regia ‘Zha 343’    | Cultivar | Juglans          |                    |
| 2   | Juglans sigillata          | Species | Juglans         |                    |
| 3   | Juglans sigillata ‘Lushui’ | Cultivar | Juglans         |                    |
| 4   | Juglans cathayensis        | Variety | Cardiocaryon     |                    |
| 5   | Juglans mandshurica        | Species | Cardiocaryon     |                    |
| 6   | Juglans hindisi            | Species | Rhysocaryon      |                    |
| 7   | Juglans major              | Species | Rhysocaryon      |                    |
| 8   | Juglans microcarpa         | Species | Rhysocaryon      |                    |
| 9   | Juglans nigra              | Species | Rhysocaryon      |                    |
| 10  | Pterocarya stenoptera      | Species |                 |                    |
| 11  | Cyclocarya paliurus        | Species |                 |                    |

---

* Resources Nursery, Forestry Bureau of Lunoning County, Henan, China.
* Forestry Academy of Yunnan Province, Kunming City, Yunnan, China.
* Beijing Botanical Garden, Institute of Botany, Chinese Academy of Sciences, Beijing, China.
3.3. Genetic relationship revealed by the UBE3 sequence

Neighbor-joining (NJ), maximum likelihood (ML) and maximum parsimony (MP) algorithms were tried. However, the NJ method showed the best resolution, and significant differences existed between the methods. Thus, only NJ trees are presented. Twenty variable sites were detected according to the aligned UBE3 sequences (Table 2; Table S1; Fig. S1). Using NJ analysis and UBE3 sequences, all nine taxa (species/variety/cultivars) within the three sections were uniquely identified, and the outgroups were placed at reasonable positions outside the genus Juglans. Bootstrap support values were more than 50% at each clade (Fig. 1). This topology was identical to the classification based on their morphology and habits [10,18–23]. Pairwise distances are shown in Table S3.

3.4. Construction of molecular taxonomic key based on nucleotide molecular formulae

The UBE3 sequence dataset was employed for construction of the nucleotide molecular formulae (NMF). The 724bp aligned sequence corresponds to the DNA tract from bases 15 to 738 of the entire sequence of the UBE3 fragment from the 5` end and includes all the variable sites of this region (Table 2; Fig. S1). The position number of each variable site used in the formula was determined according to the newly generated 724 bp-length sequence alignment. The ten polymorphic base sites used in the NMF of the taxa for the genus Juglans are No. 42, 85, 125, 205, 227, 322, 457, 459, 595 and 663 (Table 2; Fig. S1). For instance, “Nuclear DNA_UBE3_cds” was used to refer to the coding region of the nuclear UBE3 gene employed in the NMF and “aln_724 bp” refers to the aligned sequence length (724 bp) of the nine representative species/variety/cultivars in Juglans L. As a result, “Nuclear DNA_UBE3_cds_aln_724bp_” can be constructed as an NMF for molecularly characterizing the cultivar Juglans regia ‘Zha 343’, with the figure following the nucleotide character indicating the position of the corresponding polymorphic base site from the 5` end of the aligned sequence [24]. The NMF can be constructed in a similar way for the rest of the samples of the genus Juglans and the outgroups. “Nuclear DNA_UBE3_cds_aln_724bp_” is omitted to save space in the description below. “Type”, for example, in the following taxonomic key, refers to the taxon/taxa with –typed base mutation, i.e., nucleotide A can be detected at base position 42 from the 5` end in the UBE3 region. Other types of base mutation are indicated in the same way.

As shown in Fig. 2, a novel taxonomic key based on nucleotide molecular formulae is constructed by which the molecular feature of each taxon is given.

4. Discussion

Plants of Juglans sect. Cardioacaryon are precious tree species for high quality wood production. J. mandshurica and J. cathayensis are closely related taxa in Juglans sect. Cardioacaryon. J. mandshurica is mostly distributed in provinces of North and Northeast China, where the climate is colder. J. cathayensis is mainly distributed in warmer provinces of South and Southwest China [19,20,23]. The four black walnut species of Juglans sect. Rhysoacaryon are closely related with each other, with some presence in North America as well [18–23]. Members of Juglans sect. Juglans are economically important tree species for edible walnut production. The distribution of Juglans sigillata and J. sigillata ‘Lushui 1 Hao’ is limited to Southwest China (mainly Yunnan Province) [19,20,23]. J. sigillata ‘Lushui 1 Hao’ is a traditional local cultivar with an annual nut production of
Fig. 1. The neighbor-joining (NJ) tree generated based on the UBE3 gene sequences showing a success in discrimination at species, variety and cultivar levels. Numbers above the branches are bootstrap support values (%) for each clade with 1000 replicates.

1a. Type A22
2a. Type A27
3a. Type G123: NMF = A42G195128T203 A227A32C457C457C457C395C663 .......... Juglans regia ‘Zha 343’
3b. Type A123: NMF = A42G195 A123T203 A227A32C457C457C457C395C663 ............... Juglans sigillata
2b. Type T227: NMF = A42G195 R123T203T277 A32C457C457C457C395C663 .......... Juglans sigillata ‘Lushui 1 Hao’
1b. Type T22
4a. Type A85
5a. Type T203: NMF = T42A85 A123 T203 A227A32C457T457C395C663 .......... Juglans cathayensis
5b. Type C203: NMF = T42A85 A123 C203 A227A32C457T457C395C663 .......... Juglans mandshurica
4b. Type G85
6a. Type G123: NMF = T42G85 A123 T305 A227A32C457C457C395C663 ................. Juglans hindii
6b. Type A227
7a. Type T395: NMF = T42G85 A123 T203 A227 G123 A457C457T395T663 ............... Juglans major
7b. Type C955
8a. Type Y455: NMF = T42G85 A123 T203 G322A457Y455C395T663 .......... Juglans microcarpa
8b. Type C455: NMF = T42G85 A123 T203 A222 G322A457C457C395T663 ............... Juglans nigra
9a. Type C18 ........................................ Cyclocarya paliurus
9b. Type T18 ............................... The taxa excluding Cyclocarya paliurus in this key
10a. Type A274 ........................................ Pterocarya stenoptera
10b. Type G274 ............................... The taxa excluding Pterocarya_stenoptera in this key

Fig. 2. Molecular taxonomic key compiled based on nucleotide molecular formulae derived from the UBE3 sequence.
more than $1.0 \times 10^6$ kg. In contrast, the annual nut production of *J. sigillata* is around $4.0 \times 10^6$ kg. They are the major walnut trees cultivated in Yunnan Province, China. *J. sigillata* ‘Lushui 1Hao’ prefers a warmer climate with higher humidity for normal growth compared to *J. sigillata*. Fruit maturation time of *J. sigillata* ‘Lushui 1Hao’ is about 15 days earlier than that of *J. sigillata*. There is almost no difference in floral morphology between them. *J. sigillata* ‘Lushui 1Hao’ possesses 9–11 leaflets in the odd-pinnate leaf without obvious degradation of the terminal leaflet, whereas *J. sigillata* has 9–13 leaflets in its odd-pinnate leaf whose terminal leaflet degraded significantly ([19,20,23]).

Nearly 2.0 \times 10^6 kg of the annual walnut production in China is provided by *J. regia*. In fact, *J. regia* ‘Zha 343’ is a major walnut cultivar in Xinjiang Uyghur Autonomous Region, China. In the Yunnan Province, the growth of *J. regia* gradually becomes weaker after planting because the local climate averages lower temperature and higher humidity than what is required by the species. Thus, in China, *J. regia* is mainly cultivated in the walnut distribution area outside the Southwest, although plants of *J. regia* can be seen in Yunnan Province.

Generally, the greater the number of informative base sites available, the higher discrimination efficiency should be achieved during genetic diversity detection. One of the important tasks in DNA marker development is to seek DNA regions with a large number of variable base sites ([19,20,23]). However, when compared to researches on genetic variations at the family, genus, or section level, development of nuclear DNA marker covering lower taxa is complementary source for development of nuclear DNA markers for genetic diversity detection, covering both inter-specific and intra-specific levels, and will promote evaluation, conservation, and utilization of plant resources and other organisms.

### Conflict of interest

The authors declare that they have no conflict of interest.

### Acknowledgements

This study was financially supported by the Chinese Special Fund Project for the Scientific Research of the Forest Public Welfare Industry (Project No. 2010040408) from the State Forestry Administration of China and by the National Natural Science Foundation of China (Grant No. 30972412). The expert who instructed the identification of samples used in this study is Prof. Runquan Dong and Yu Zhang of the Forestry Academy of Yunnan Province, Huzhi Xu, professor and former director of Forestry Bureau of Luoning County, Henan, China. The authors thanks Wengmao Ma, Chenguang Wang, Fengmin Li, Peng Wang, Zhihong Ding, Weiwai Gao, Hao Liu, Qingguo Ma, Xianlan Li, Bin Lu, and Ping Zhao for their kind help in field investigations, material collections and discussions. We are sincerely grateful to three anonymous referees for their thoughtful and meaningful comments.

### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.btre.2014.11.003.

### References

[1] W.J. Kress, K.J. Wurdack, E.A. Zimmerman, L.A. Weigt, D.H. Janzen, Use of DNA barcodes to identify flowering plants, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 8369–8374.

[2] C. Sass, D.P. Little, D.W. Stevenson, C.D. Specht, DNA barcoding in the cyadals: Testing the potential of proposed barcoding markers for species identification of cyadals, PLoS One 2 (2007) e1154, doi:http://dx.doi.org/10.1371/journal.pone.0013102.

[3] B.Q. Ren, X.G. Xiang, Z.D. Chen, Species identification of *Alnus* (Betulaceae) using *nr*DNA and cpDNA genetic markers, Mol. Ecol. Resour. 10 (2010) 594–605, doi:http://dx.doi.org/10.1111/j.1755-0998.2009.00815.x.

[4] H. Yao, J.Y. Song, C. Liu, K. Luo, J.P. Han, Y. Li, X. Pang, H. Xu, Y. Zhu, F. Xiao, S. Chen, Use of ITS, region as the universal DNA barcode for plants and animals, PLoS One 5 (2010) e13102, doi:http://dx.doi.org/10.1371/journal.pone.0013102.

[5] X.G. Xiang, J.B. Zhang, A.M. Lu, R.Q. Li, Molecular identification on species in *Juglandaceae*: a tiered method, J. Syst. Evol. 49 (3) (2011) 252–260, doi:http://dx.doi.org/10.1111/j.1759-6831.2011.00133.x.

[6] K.S. Cheon, K.O. Yoo, Phylogeny of *Hanburya* (Cycadales), PLoS One 2 (2007) e1154, doi:http://dx.doi.org/10.1371/journal.pone.0013102.

[7] K.S. Cheon, K.O. Yoo, Phylogeny of *Hanburya* (Cycadales), PLoS One 8 (2) (2013) 1–16, doi:10.1371/journal.pone.0057934.

[8] A.M. Stanford, R. Harden, C.R. Parks, Phylogeny and biogeography of *Juglans* (*Juglandaceae*) based on matK and ITS sequence data, Am. J. Bot. 87 (2000) 837–842.

[9] D. Potter, R.Y. Gao, S. Baggett, J.R. Mckenna, G.H. McGranahan, Defining the sources of paradox: DNA sequence markers for North American walnut (*Juglans* L.) species and hybrids, Sci. Hort. 94 (2002) 157–170.

[10] D.E. Stone, S.H. Oh, E.A. Trupp, G.L.E. Rios, P.S. Manus, Natural history distribution, phylogenetic relationships, and conservation of Central American black walnuts (*Juglans* sect. *Rhysocaryon*), J. Torrey Bot. Soc. 136 (2009) 1–25.

[11] L.F. Ciarrello, G. Pontecorvo, P. Piccirillo, A.D. Luca, P. Carillo, I. Kafantaris, F. Woodrow, Use of nuclear and mitochondrial single nucleotide polymorphisms to characterize Chinese walnut (*Juglans regia* L.) genotypes,
Plant Mol Biol Rep. 31 (2013) 1116–1130, doi:http://dx.doi.org/10.1007/s11105-013-0575-2.

[12] S. Cosmulescu, M. Botu, Walnut biodiversity in south-western Romania-resource for perspective cultivars, Pak. J. Bot. 44 (1) (2012) 307–311.

[13] A. Ganoth, Y. Tsfadia, R. Wiener, Ubiquitin: molecular modeling and simulations, J. Mol. Graphics Modell. 46 (2013) 29–40, doi:http://dx.doi.org/10.1016/j.jmgm.2013.09.006.

[14] I. Marin, Evolution of plant hect ubiquitin ligases, PLoS One 8 (7) (2013) e68536, doi:http://dx.doi.org/10.1371/journal.pone.0068536.

[15] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal X version 2.0, Bioinformatics 23 (2007) 2947–2948, doi:http://dx.doi.org/10.1093/bioinformatics/btm404.

[16] Librado, J. Rozas, DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, Bioinformatics 25 (2009) 1451–1452, doi:http://dx.doi.org/10.1093/bioinformatics/btp187.

[17] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol. 30 (2013) 2725–2729, doi:http://dx.doi.org/10.1093/molbev/mst197.

[18] L.E. Gunter, G. Kochert, D.E. Giannasi, Phylogenetic relationships of the Juglandaceae, Plant Syst. Evol. 192 (1994) 11–29.

[19] A.M. Lu, The geographical dispersal of Juglandaceae, Acta Phytotaxon. Sin. 20 (1982) 257–271.

[20] A.M. Lu, D.E. Stone, I.J. Grauke, Juglandaceae, in: Z.Y. Wu, P.H. Raven (Eds.), Flora of China, Science Press, Beijing, and Missouri Botanical Garden Press, St. Louis, Missouri, USA, 1999, pp. 277–285.

[21] M.K. Aradhya, B. Potter, F. Gao, C.J. Cinon, Molecular phylogeny of Juglans (Juglandaceae): abigoeographic perspective, Tree Genet Genomes 3 (2007) 363–378, doi:http://dx.doi.org/10.1007/s11295-006-0078-5.

[22] W.E. Manning, The classification within the Juglandaceae, Ann. Mo. Bot. Gard. 65 (1978) 1058–1087.

[23] D. Pei, X.Z. Lu, Walnut Germplasm Resources In China (In Chinese), China Forestry Publishing House, Beijing, 2011, pp. 1–208.

[24] Z.L. Suo, C.H. Zhang, Y.Q. Zheng, L.X. He, X.B. Jin, B.X. Hou, J.J. Li, Revealing genetic diversity of tree peonies at micro-evolution level with hyper-variable chloroplast markers and floral traits, Plant Cell Rep. 31 (2012) 2199–2213, doi: http://dx.doi.org/10.1007/s00299-012-1330-0.