Phylogenetic analysis of *Uncaria* species based on internal transcribed spacer (ITS) region and ITS2 secondary structure

Shuang Zhu\(^a\), Qiwei Li\(^a\), Shanchong Chen\(^a\), Yesheng Wang\(^a\), Lin Zhou\(^a\), Changqing Zeng\(^b\) and Jun Dong\(^b\)

\(^a\)Center for Bioresources and Drug Discovery and School of Biosciences and Biopharmaceutics Guangdong Province Key Laboratory for Biotechnology Drug Candidates, Guangdong Pharmaceutical University, Guangzhou, People's Republic of China; \(^b\)School of Traditional Chinese Medicine, Guangdong Pharmaceutical University, Guangzhou, People's Republic of China

**ABSTRACT**

**Context:** The plant genus *Uncaria* (Rubiaceae), also known as Gouteng, is the source of an important traditional Chinese medicine. Misidentification and adulteration of Gouteng affect the safety and efficacy of the medication. Phylogenetic relationships among the species of this genus are unknown.

**Objective:** The present study sought to detect the phylogenetic relationships based on internal transcribed spacer (ITS) region of all 12 species of *Uncaria* recorded in the *Flora of China*.

**Materials and methods:** Accession of seven species of *Uncaria* served as reference samples. ITS region was used for polymerase chain reaction (PCR) amplification of the reference samples representing 39 specimens. Distance analysis, species discrimination, and secondary structure of ITS2 were used to assess the ability of ITS sequence in authenticating. The phylogenetic relationships were detected using three methods: Bayesian inference (BI), maximum likelihood (ML), and neighbor joining (NJ).

**Results:** Five species of traditional Chinese medicine Gouteng were well resolved in molecular phylogenetic tree. Besides, *Uncaria lancifolia* Hutch. was closer to *U. rhynchophylla* F.C. How and *U. sessilifructus* Roxb. was closer to *U. laevigata* Wall. within the tree. Further, we also found that ITS2 secondary structure can be a candidate tool in distinguishing two closely related species *U. yunnanensis* K.C.Hsia and *U. lanosa* Wall. For accurate identification of different species of *Uncaria* based on species-specific nucleotide sites, a consensus sequences database with all 12 species is established.

**Discussions and conclusions:** The results are able to discriminate *Uncaria* species and illustrate the phylogenetic relationships, which are essential for the investigation of adulterants and misidentifications of *Uncaria*.

**Introduction**

*Uncaria* (Rubiaceae) is a genus that contains 34 species: 29 in tropical Asia through Australia, three in Africa and Madagascar, two in tropical America and 12 species (five endemics) in China (Ridsdale 1978; Tao et al. 2011). The Chinese pharmacopoeia records the stems with hooks from five *Uncaria* species namely, *Uncaria rhynchophylla* Miq., *U. macrophylla* Wall., *U. hirsute* Havil., *U. sinensis* Havil, and *U. sessilifructus* Roxb., which form an important part of traditional Chinese medicine known collectively as Gouteng (Chinese Pharmacopoeia Commission 2015). Previous studies have reported that the main compounds in Gouteng species (alkaloids, triterpenes, and uncarinic acids) have several beneficial properties, such as being antihypertensive, analgesic, sedative, and antioxidant (Ridsdale 1978; Song et al. 2000; Kang et al. 2015; Pan et al. 2015). The potential pharmacological activity from *Uncaria* continues to develop, particularly in the area of immunomodulatory, anti-inflammatory, and antitumor (Lee et al. 2000; Yuan et al. 2009; Lima-Junior et al. 2013; Zhang et al. 2013). Moreover, Gouteng is a traditional herb used in northeast Asia for the treatment of Parkinson’s disease (Shim et al. 2009; Chen 2016). The chemical composition and bioactivities vary among species: *U. rhynchophylla* and *U. sinensis* have the highest quality of beneficial compounds and are the most important medicine resources of Gouteng in China (Kang et al. 2015). However, due to their extensive collection for medicinal use, plants of high value species are now endangered in some regions of China. Moreover, the medicinal herb Gouteng is being misidentified and adulterated with similar but less valuable species such as *Uncaria laevigata* Wall., *U. lancifolia* Hutch., and *Uncaria scandens* Hutch., which is affecting the safety and efficacy of the medication.

The identification of *Uncaria* (Gouteng) is based primarily on the morphological characteristics, microscopic structures and/or chemical components of the plant (Gao et al. 2015). Previous studies have used restriction fragment length polymorphism (RFLP) and random amplified polymorphic DNA (RAPD) to look for molecular markers in *Uncaria* (Xu et al. 2012; Zhu, Zhou et al. 2012). Zhao et al. (2018) demonstrated that the liquid chromatography–mass spectrometry tandem ion trap-time of flight mass spectrometry (LCMS-IT-TOF) was a feasible method to solve the confusion in *Uncaria* species. Tang et al. (2016)
evaluated five different DNA barcodes and proposed ITS2 and psbA–trnH as suitable markers for identification in this genus, but the phylogenetic relationships of Uncaria are still not fully known, especially among the five medicinal species recorded in the Chinese pharmacopoeia.

DNA barcoding using short genetic markers or gene regions for species identification has potential for use in the detection and protection of endangered and valuable species (Hebert et al. 2003; CBOL Plant Working Group 2009; Hollingsworth et al. 2011). The Consortium for the Barcode of Life (CBOL) has proposed combined plastid barcoding with matK and rbcL genes as an alternative for species identification among plants (CBOL Plant Working Group 2009). Several chloroplast gene sequences such as psbA–trnH, trnL–trnF, ycf1, and rpcO1 have been evaluated as potential DNA barcodes (CBOL Plant Working Group 2009; Dong et al. 2015; Yu et al. 2016). In addition to plastid barcoding, internal transcribed spacer (ITS) regions of ribosomal genes have been proposed as supplemental barcodes for matK and rbcL. The ITS region includes the ITS1 and ITS2 regions, separated by the 5.8S gene, and is situated between the 18S and 28S genes in the nrDNA repeat unit (Bellemain et al. 2010). Despite problems associated with the ITS region such as incomplete concerted evolution, fungal contamination, and difficulties in amplifying and sequencing in some species, it provides enough variable sites for differentiating among species (Chen et al. 2010; Yao et al. 2010; China Plant BOL Group 2011; Lee et al. 2016; Wang et al. 2017). Compared to the ITS region, the ITS2 region is easy to amplify and sequence, and also provides sufficient information for phylogenetic analysis. Moreover, the secondary structure of the ITS2 region can offer additional information for species identification. The ITS2 RNA transcript contains a core structure of two helices with hallmark characteristics that are important for ribosomal RNA processing (Coleman 2007). This secondary structure allows the detection of sequencing errors and pseudogenes in the ITS2 region (Coleman 2007; Rampersad 2014).

In the present study, we tested the ability of ITS and ITS2 regions to discriminate among Uncaria species, and established a sequence database with species-specific positions to discriminate the 12 species of Uncaria recorded in the Flora of China. We also attempted to illustrate the phylogenetic relationships among 12 species of Uncaria and use ITS2 secondary structure to distinguish potential closely related species.

Materials and methods

DNA extraction, PCR amplification and sequencing

This study was based on a total of 39 Uncaria specimens comprised of 32 fresh or silica gel-dried leaf specimens and 7 stem specimens, which included 7 of 12 Uncaria species (Table 1). They were collected from various locations in China and identified by Professor Changqing Zeng (School of Traditional Chinese Medicine, Guangdong Pharmaceutical University) based on morphological traits. Total DNA was extracted from fresh or silica gel-dried leaf/stem tissue using the Plant Genomic DNA Kit (Biateke Co., Guangzhou, China) according to the manufacturer’s instructions. A primer pair (forward-ITS5: 5'-GGAAATTTAGAAGCTGAAGTGTAAC-3', reverse-ITS4: 5'-TCTCCGCTCTATTGATATGC-3') was used for PCR amplification on an S1000 Thermal Cycler (BIO-RAD, USA). The PCR cycle consisted of 3 min at 95°C, 30 cycles of 1 min at 94°C, 50 s at 56°C, 55 s at 72°C, and 10 min at 72°C. PCR products were purified and bidirectional sequenced using the same primer pair in Beijing Genomic Institute (Guanzhou, China). The ITS sequences obtained by sequencing were edited and refined manually using the software DNAStar version 7.1.

Sequence analysis, distance analysis and species discrimination

The ITS sequences of the genus Uncaria were queried and downloaded from the National Center for Biotechnology Information (NCBI) database. Low quantity and ambiguous sequences were manually checked and deleted. Further, a total of 39 ITS sequences were obtained in this study. Each sequence was individually defined for a complete ITS2 fragment using the “Annotate” feature of the ITS2 database website (http://its2.bioapps.biozentrum.uni-wuerzburg.de/) (Schultz et al. 2006; Keller et al. 2009). DNA sequences generated from this study along with those downloaded from GenBank were assembled and aligned using the software MEGA version 6.06 (Tamura et al. 2013). The ITS dataset was used for nucleotide composition analysis, distance analysis, species discrimination and phylogenetic analysis, while the ITS2 dataset was used for secondary structure analysis. The nucleotide composition in the ITS region of each species was generated using MEGA version 6.06. Intraspecific and interspecific distances were calculated with the Kimura 2-Parameter (K2P) method using the software TaxonDNA version 1.8 (Kimura 1980; Meier et al. 2006). Barcoding gaps (i.e., the distribution of the pairwise intra- and interspecific distances) were illustrated by bar graphs. Species discrimination was calculated using the “Best Match” and “Best Close Match” functions in TaxonDNA, based on the K2P method and a minimum sequence overlap of 100 bp.

Phylogenetic analysis

The sequence datasets were analyzed with three phylogenetic methods: Bayesian inference (BI), maximum likelihood (ML) and neighbor joining (NJ). Bayesian inference was performed with the computer program MrBayes ver. 3.2.6 (Ronquist et al. 2012) under the general time reversible (GTR)+G nucleotide substitution models determined by the program Mr Model test ver. 2.3 (Nylander 2004). Four simultaneous chains of Markov Chain Monte Carlo (MCMC) algorithm were run twice with two million generations. Trees were sampled every hundredth generation, whereby the first 5,000 trees were discarded as burn-in. The remaining trees were used to calculate posterior probabilities (PP) of the branching pattern in the 50% majority-rule consensus tree. The ML analyses, including 1,000 nonparametric bootstrap replicates, were carried out in MEGA version 6.06 under the Tamura 3-parameter (T92)+G model. The NJ tree was constructed under the K2P model, which is based on distance substitution, using MEGA version 6.06. Bootstrap support for the NJ tree was estimated with 1000 replicates, while uninformative characters (gaps and missing data) were completely deleted. Moreover, two Nauclea species were downloaded from GenBank database for use as outgroup species, as this genus is a relative of Uncaria (Manns and Bremer 2016).
Table 1. Detailed information about newly obtained ITS and ITS2 sequences from *Uncaria* specimens in the present study.

| Voucher | Tissue | Collection site                          | Putative identification | Molecular identification | ITS GenBank accession | Length (bp) | GC (%) | ITS2 GenBank accession | Length (bp) | GC (%) |
|---------|--------|-----------------------------------------|--------------------------|--------------------------|-----------------------|-------------|--------|------------------------|-------------|--------|
| GT-1    | Leaf   | Dinghu Mountain, Zhaoqing, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033267 | 677 | 61.6 | MF033306 | 220 | 66.4 |
| GT-16   | Stems  | Dinghu Mountain, Zhaoqing, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033268 | 677 | 61.7 | MF033307 | 220 | 66.4 |
| GT-35   | Leaf   | Dinghu Mountain, Zhaoqing, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033269 | 677 | 61.6 | MF033308 | 220 | 66.4 |
| GT-8    | Stems  | Guangzhou University of Chinese Medicine, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033270 | 677 | 61 | MF033309 | 220 | 65 |
| GT-17   | Leaf   | Yangshan, Qingyuan, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033271 | 677 | 61 | MF033310 | 220 | 65 |
| GT-18   | Leaf   | Wengyuan, Shaoguan, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033272 | 677 | 61 | MF033311 | 220 | 65 |
| GT-19   | Leaf   | Guiyang University of Chinese Medicine, Guizhou Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033273 | 677 | 61 | MF033312 | 220 | 65 |
| GT-20   | Leaf   | Meizhou, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033274 | 677 | 61 | MF033313 | 220 | 65 |
| GT-21   | Leaf   | Yangshan, Qingyuan, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033275 | 677 | 60.9 | MF033314 | 220 | 65 |
| GT-22   | Leaf   | Wuzhou, Guangxi Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033276 | 677 | 61 | MF033315 | 220 | 65 |
| GT-24   | Leaf   | Guangxi botanical garden, Guangxi Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033277 | 677 | 61 | MF033316 | 220 | 65 |
| GT-25   | Leaf   | Guangzhou University of Chinese Medicine, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033278 | 677 | 61 | MF033317 | 220 | 65 |
| GT-26   | Leaf   | Dapu, Meizhou, Guangdong Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033279 | 677 | 61 | MF033318 | 220 | 65 |
| GT-27   | Leaf   | Tianzhu, Guizhou Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033280 | 677 | 61 | MF033319 | 220 | 65 |
| GT-28   | Leaf   | Majiang, Guizhou Province, China | *U. rhynchophylla* | *U. rhynchophylla* | MF033281 | 677 | 61.2 | MF033320 | 220 | 65 |

(continued)
| Voucher | Tissue | Collection site                                      | Putative identification | Molecular identification | ITS GenBank accession | Length (bp) | GC (%) | ITS2 GenBank accession | Length (bp) | GC (%) |
|---------|--------|-----------------------------------------------------|-------------------------|---------------------------|----------------------|-------------|--------|------------------------|-------------|--------|
| GT-31   | Leaf   | Kaiyang, Guizhou Province, China                    | U. rhynchophylla        | U. rhynchophylla          | MF033282 677         | 60.9        |        | MF033321               | 220         | 65     |
| GT-32   | Leaf   | Guangxi botanical garden, Guangxi Province, China  | U. rhynchophylla        | U. rhynchophylla          | MF033283 677         | 60.9        |        | MF033322               | 220         | 64.5   |
| GT-H    | Leaf   | Jinfoshan, Chongqing, China                         | U. rhynchophylla        | U. rhynchophylla          | MF033284 677         | 61.3        |        | MF033323               | 220         | 65     |
| GT-Hunan| Leaf   | Xinhua, Hunan Province, China                       | U. rhynchophylla        | U. rhynchophylla          | MF033285 677         | 61          |        | MF033324               | 220         | 65     |
| GT-GZY  | Leaf   | Guangzhou University of Chinese Medicine, Guangdong| U. rhynchophylla        | U. rhynchophylla          | MF033286 677         | 60.9        |        | MF033325               | 220         | 64.5   |
| GT-MZh  | Leaf   | Meizhou, Guangdong Province, China                  | U. rhynchophylla        | U. rhynchophylla          | MF033287 677         | 61          |        | MF033326               | 220         | 65     |
| GT-4    | Leaf   | Jinfoshan, Chongqing, China                         | U. sinensis             | U. sinensis               | MF033288 677         | 60.9        |        | MF033327               | 220         | 64.1   |
| GT-6    | Stems  | Guangxi botanical garden, Guangxi Province, China  | U. sinensis             | U. sinensis               | MF033289 677         | 60.9        |        | MF033328               | 220         | 64.1   |
| GT-7    | Stems  | Guizhou Province, China                             | U. sinensis             | U. sinensis               | MF033290 677         | 60.9        |        | MF033329               | 220         | 64.1   |
| GT-34   | Leaf   | Anshun, Guizhou Province, China                     | U. sinensis             | U. sinensis               | MF033291 677         | 61          |        | MF033330               | 220         | 64.5   |
| GT-M    | Leaf   | South China botanical garden, Guangdong Province,   | U. hirsuta              | U. homomalla              | MF033292 677         | 61.4        |        | MF033331               | 220         | 65     |
| GT-2    | Leaf   | South China botanical garden, Guangdong Province,   | U. rhynchophylla        | U. hirsuta                | MF033293 676         | 61.4        |        | MF033332               | 220         | 66.4   |
| GT-12   | Leaf   | Guangxi botanical garden, Guangxi Province, China   | U. hirsuta              | U. hirsuta                | MF033294 676         | 61.4        |        | MF033333               | 220         | 66.4   |
| GT-13   | Stems  | Dinghu Mountain, Zhaoqing, Guangdong Province, China| U. hirsuta              | U. hirsuta                | MF033295 676         | 61.4        |        | MF033334               | 220         | 66.4   |
| GT-10   | Leaf   | Guangxi botanical garden, Guangxi Province, China   | U. sessilfructus        | U. sessilfructus          | MF033296 678         | 61.9        |        | MF033335               | 221         | 66.1   |
| GT-15   | Stems  | Guangxi botanical garden, Guangxi Province, China   | U. sessilfructus        | U. sessilfructus          | MF033297 678         | 61.9        |        | MF033336               | 221         | 66.1   |
Prediction of secondary structure

The secondary structure of ITS2 was folded using the Mfold web server using the default temperature (37°C) and ionic conditions [http://unafold.rna.albany.edu] (Zuker 2003). All output was used to predict a consensus secondary structure on the RNAalifold webserver [http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAalifold.cgi], set at default options (Lorenz et al. 2011). Consensus ITS secondary structures were re-drawn as radial view structures and annotated using RNAlogo [http://rnalogo.mbc.nctu.edu.tw/index.php] (Chang et al. 2008). The compensatory base changes (CBCs) were detected using the program 4SALE version 1.7 [http://4sale.bioapps.biozentrum.uni-wuerzburg.de/] (Seibel et al. 2006, 2008).

Results and discussion

Sequences

The dataset used in this study included all 12 species of Uncaria recorded in Flora of China. The primer pair ITS5 and ITS4 effectively amplified the complete ITS sequences of all 39 Uncaria specimens, which were deposited in the NCBI GenBank database under the accession numbers MF033267 to MF033305. The ITS2 sequences (annotated and defined by the ITS2 database) were also submitted to GenBank (MF033306 to MF033344). The total length of newly obtained ITS and ITS2 regions ranged from 676 to 678 bp and from 220 to 221 bp,
Table 3. Species-specific nucleotide variation in the ITS region based on consensus sequences of different Uncaria species.

| Position | U. rhynchophylloides | U. rhynchophylla | U. sinensis | U. homomalla | U. hirsuta | U. sessilifructus | U. macrophylla | U. laeovigata | U. lancifolia | U. lanosa | U. yunnanensis |
|----------|-----------------------|-----------------|-------------|--------------|------------|-----------------|---------------|--------------|-------------|----------|---------------|
| 1        | T                     | G               | A            | C            | T          | A               | C             | R            | C           | G         | T             |
| 2        | A                     | G               | C            | C            | A          | T               | C             | C            | C           | C         | Y             |
| 3        | G                     | T               | G            | T            | T          | G               | C             | A            | A           | C         | Y             |
| 4        | G                     | C               | G            | A            | C          | A               | S             | A            | Y           | C         | T             |
| 5        | C                     | G               | C            | G            | T          | M               | W             | A            | C           | C         | V             |
| 6        | G                     | C               | G            | C            | S          | G               | A             | G            | A           | C         | T             |
| 7        | G                     | G               | C            | C            | G          | S               | G             | A            | T           | C         | A             |
| 8        | G                     | C               | G            | C            | G          | G               | A             | G            | C           | C         | T             |
| 9        | G                     | C               | G            | C            | G          | G               | A             | C            | C           | C         | V             |
| 10       | G                     | C               | G            | C            | G          | G               | A             | G            | A           | C         | T             |

Species |
|---------|
| U. rhynchophylloides |
| U. rhynchophylla |
| U. sinensis |
| U. homomalla |
| U. hirsuta |
| U. sessilifructus |
| U. macrophylla |
| U. laeovigata |
| U. lancifolia |
| U. lanosa |
| U. yunnanensis |

Distance analysis and species discrimination

The mean haplotype diversity and the mean nucleotide diversity of the ITS region of all Uncaria species were 0.954% and 1.967%, respectively. For the ITS2 region, the respective diversities were 0.932% and 2.932% (Table 4). U. laevigata, U. lancifolia, and U. sessilifructus had high mean nucleotide diversity in the ITS and ITS2 region. For a suitable barcode for species discrimination, the intraspecific distance is required to be lower than the interspecific distance, i.e., it should show a barcode gap (Meier 2008; Hartvig et al. 2015). The barcode gap analysis, based on uncorrected p-distance histograms, revealed a partial overlap between the intraspecific and interspecific distances of ITS and ITS2. The interspecific distances of sequences for ITS and ITS2 were considerably higher than the intraspecific distances (Figure 1). Moreover, the mean interspecific distances were significantly higher than the corresponding intraspecific distances for both ITS and ITS2 regions (Table 2).

Species discrimination ability of the ITS and ITS2 region were tested by the best match and best close match methods (Table 2). These two methods are based on direct sequence comparison instead of tree-based identification. Best match assigns sequences with the smallest distance to the query sequence, while best close match requests best match sequences within 95% of the intraspecific distance (Meier et al. 2006). We found that we achieved a high rate of species discrimination in genus Uncaria using these two methods. The ITS region was able to correctly
identify 128 out of 132 individuals (96.97%) using both best match and best close match. The correct identification rates of the ITS2 region were slightly lower than for the ITS region, yet still had a high discrimination ability (>94%). U. lanosa had only one sequence in the dataset, thus it could not be correctly identified due to the lack of conspecific matching.

Previous studies indicated that the ITS and ITS2 regions evolve rapidly, leading to genetic changes similar to those in plastid DNA barcodes (such as matK, rbcL, psbA-trnH, et al.) (Kress et al. 2005; China Plant BOL Group 2011). However, the well-documented concerns and limitations of ITS include: (1) divergent paralogous copies within individuals lead by incomplete concerted evolution, (2) difficulties in amplifying and sequencing in some sample sets such as gymnosperms, and (3) fungal contamination (China Plant BOL Group 2011; Hollingsworth et al. 2011). Although they are imperfect, we believe that the ITS and ITS2 regions have sufficient ability to identify Uncaria species using best match and best close match techniques.

Phylogenetic analysis within Uncaria species

The ITS region is one of the most frequently utilized barcode in phylogenetic analysis at genus and species levels in eukaryotes (Coleman 2003; CBOL Plant Working Group 2009). We analyzed the ITS and ITS2 dataset with three different phylogenetic methods (BI, ML and NJ). The phylogenetic relationships that we identified among 12 species of Uncaria are shown in Figure 2 and Supplementary Figure S1. Phylogenetic trees inferred from the ITS and ITS2 dataset exhibited a structure similar to the branching patterns of the respective phylogenetic trees. The taxa U. hirsute Havil., U. homomalla Miq., and U. scandens were each clustered into separated clades, in both the ITS and ITS2 phylogenetic trees.

The phylogenetic relationships of five important medicinal species of Uncaria were shown in the present study (Figure 2). Three medicinal Uncaria plants, U. rhynchophylla, U. sinensis and U. hirsuta, are reportedly similar in the morphological characteristics of their hook-bearing stems, and therefore difficult to identify (Zhu, Chen, et al. 2012; Chinese Pharmacopoeia Commission 2015). In the present study, the taxa U. rhynchophylla and U. sinensis each clustered as monophyletic and as sister species in the ITS tree, while U. sinensis was nested with U. rhynchophylla in the ITS2 tree, and then clustered with three cladest represented U. hirsuta, Uncaria homomalla and U. scandens. The other two medicinal plants were U. macrophylla and U. sessilifructus. U. sessilifructus was clustered as monophyletic and then clustered with U. laevigata. Except one individual (voucher GT-14), U. macrophylla was not clustered into a clade but separated from others species.

Ridsdale (1978) concluded that U. rhynchophylloides is the same species as U. rhynchophylla even though they appear morphologically distinct (Wang et al. 2007) and have different chemical constituents. U. rhynchophylloides has a high proportion of hirsutine but very low content of rhynchophylline and isorhynchophylline, which are important chemical compounds in medicinal Gouteng (Zhong and Feng 1996). In contrast, phylogenetic analysis in the present study showed that U. rhynchophylloides was not a synonym of U. rhynchophylla, as they were located far apart (Figure 2). On the other hand, U. rhynchophylloides was depicted as a separate clade and nested in U. lancifolia, which is highly similar to prior description of U. rhynchophylloides based on morphological characteristics (Wang et al. 2007), yet U. lancifolia did not cluster into a monophyletic clade in two trees. Similar branching patterns were also found between U. sessilifructus and U. laevigata (Figure 2), i.e., U. sessilifructus was clustered into a separate clade but U. laevigata failed to cluster into a monophyletic group. High nucleotide diversity may explain why U. lancifolia and U. laevigata were not monophyletic (Table 4).

Two highly similar species U. yunnanensis and U. lanosa clustered together into a monophyletic clade with high support rates in the ITS and ITS2 phylogenetic tree (Figure 2; Supplementary Figure S1). The only nucleotide insertion that occurred in U. yunnanensis (mentioned above) was deleted during construction of the phylogenetic tree.

To summarize, the ITS region has more variable sites to discriminate Uncaria species and thus has better discriminating
performance than does the ITS2 region. The phylogenetic relationships of most Uncaria species were able to be resolved, except for U. yunnanensis and U. lanosa. In additional, the chloroplast DNA regions (matK, rbcL, psbA-trnH, et al.) may be candidate barcodes to resolve the phylogenetic relationship between U. yunnanensis and U. lanosa.

Secondary structure modeling and CBC analysis

To predict ITS2 secondary structure, the minimum free energy method was used to form a structure synonymous with a natural-mode structure (Tinoco et al. 1971). The consensus secondary structure of most Uncaria taxa shared a similar folding pattern, i.e., four helices surrounding a central loop (Figure 3) which generated by the interaction of 5.8S-LSU (5.8S rRNA-28S rRNA) (Schultz et al. 2005; Coleman 2007). In order to discriminate the two closely related species U. yunnanensis and U. lanosa, we compared their secondary structure (Figure 3). The secondary structure of U. lanosa had a similar fold pattern as the consensus structure of others species, while U. yunnanensis showed a different fold pattern in helix IV (consisting of helix IVa and helix IVb). This difference may be caused by the species-specific insertion at 181 nt in the ITS2 molecule (at 602 nt in ITS region), which changed the fold pattern based on the minimum free energy method. Intraspecific mutation and prediction models can also induce the change of secondary structure. Hence, we also predicted the U. yunnanensis secondary structure using other programs or web servers [RNAstructure version 5.8.1; Reuter and Mathews 2010; RNAalifold webserver, and LocARNA webserver (http://rna.informatik.uni-freiburg.de/LocARNA/Input.jsp); Will et al. 2007; Smith et al. 2010; Will
et al. 2012], and obtained similar results for the fold pattern (Supplementary Figure S2). Thus, we conclude that the secondary structure is a candidate method to discriminate *U. yunnanensis* and *U. lanosa*.

CBCs in the ITS2 molecule correlate with sexual incompatibility, so is an important molecular indicator for discriminating closely related species (Müller et al. 2007; Wolf et al. 2013; Shazib et al. 2016). However, it was reported that CBCs were not able to effectively discriminate among species (Caisová et al. 2013; Shazib et al. 2016). In the present study, we did not observe any CBCs in helices of the ITS2 molecule among *Uncaria* species, based on the 4SALE program (Supplementary Table S2). Thus, we conclude that CBC analysis is not an effective method to discriminate among *Uncaria* species.

**Conclusions**

A comprehensive phylogenetic analysis including all the 12 species of *Uncaria* recorded in the *Flora of China* were concluded.
Initially, all Uncaria medicinal species could be clustered clearly in the molecular phylogeny tree. Secondly, we established the ITS database with all of obtained sequences and found that ITS sequences have appropriate variable sites for discrimination most of species in Uncaria. Finally, the ITS2 secondary structure can be used as candidate method for distinguishing the two closely related species U. yunnanensis and U. lanosa.

Disclosure statement

The authors declare that they have no competing interests.

Funding

This research was supported by the National Natural Science Foundation of China under Grants of (41271263), the Science and Technology Plan Projects of Zhongshan City, China (No. 20101H1017).

ORCID

Qwei Li https://orcid.org/0000-0002-7674-5565

References

Bellemain E, Carlsen T, Brochmann C, Coissac E, Taberlet P, Kauserud H. 2010. ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. BMC Microbiol. 10:189.

Caisová L, Marín B, Mellkonian M. 2013. A consensus secondary structure of ITS2 in the chlorophya identified by phylogenetic reconstruction. Protop. 164:482–496.

CBOL Plant Working Group. 2009. A DNA barcode for land plants. Proc Natl Acad Sci USA. 106:12794–12797.

Chang T, Horng J, Huang H. 2008. RNALogo: a new approach to display structural RNA alignment. Nucleic Acids Res. 36:W91–W96.

Chen L. 2016. The molecular mechanism of Chinese medicine Uncaria rhynchophylla (gouteng) for inducing autophagy and protecting neurons in Parkinson’s disease. China: Hong Kong Baptist University.

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, et al. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS One. 5:e8613.

China Plant BOL Group. 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc Natl Acad Sci USA. 108:19641–19646.

Chinese Pharmacopoeia Commission. 2015. The Pharmacopoeia of the People’s Republic of China. Beijing: China Medical Science Press.

Coleman AW. 2006. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. Trends Genet. 19:370–375.

Coleman AW. 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure of the internal transcribed spacer region. Mol Phylogenet Evol. 102:128–136.

Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, Cheng T, Guo J, Zhou S. 2015. yefl, the most promising plastid DNA barcode of land plants. Sci Rep. 5:8348.

Gao X, Liu S, Chen X, Yan H, Zhong Z, Zhang W, Zhu S. 2015. Studies on the botanical sources and molecular identification of Aquilaria liguim resin biomass based on ITS2 DNA barcode. Lishizhen Med Mater Med Res. 26:13–16.

Hartvig I, Czako M, Kjaer ED, Nielsen LR, Theilade I. 2015. The use of DNA barcoding in identification and conservation of rosewood (Dalbergia spp.). PLoS One. 10:e0138231.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR. 2003. Biological identifications through DNA barcodes. Proc Biol Sci. 270:313–321.

Hollingsworth PM, Graham SW, Little DP. 2011. Choosing and using a plant DNA barcode. PLoS One. 6:e19254.

Kang BK, Kim MK, Kim SY, Lee SJ, Choi YW, Choi BT, Shin HK. 2015. Anti-neuroinflammatory effects of Uncaria sinensis in LPS-stimulated BV2 microglia cells and focal cerebral ischemic mice. Am J Chin Med. 43:1099–1115.

Keller A, Schleicher T, Schultz J, Müller T, Dandekar T, Wolf M. 2009. 5.8S RNA interaction and HMM-based ITS2 annotation. Gene. 430:50–57.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 16:111–120.

Kress WJ, Wurdack KJ, Zimmerman EA, Weigt LA, Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci USA. 102:8369–8374.

Lee JS, Kim J, Kim BY, Lee HS, Ahn JS, Chang YS. 2000. Inhibition of phosphorylase b-gamma1 and cancer cell proliferation by triterpene esters from Uncaria rhynchophylla. J Nat Prod. 63:753–756.

Lee SY, Ng WL, Mahat MN, Nazre M, Mohamed R. 2016. DNA barcoding of the endangered Aquilaria (Thymelaeaceae) and its application in species authentication of agarwood products traded in the market. PLoS One. 11:e0154631.

Lima-Junior RS, Mello C da S, Siani AC, Valente LMM, Kubelka CF. 2013. Uncaria tormentosa alkaloid fraction reduces paracellular permeability, IL-8 and NS1 production on human microvascular endothelial cells infected with dengue virus. Nat Prod Commun. 8:1547–1550.

Lorenz R, Bernhart SH, Höner C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. Alg Mol Biol. 6:26.

Manns U, Bremer B. 2010. Towards a better understanding of intertribal relationships and stable tribal delimitations within Cinchonoidae s.s. (Rubiaceae). Mol Phylogenet Evol. 56:21–39.

Meier R. 2008. DNA sequences in taxonomy: opportunities and challenges. In: Wheeler QD, editors. The new taxonomy. Chapter 7, New York: CRC Press; p. 95–127.

Meier R, Shiyan K, Vaidya G, Ng PKL. 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst Biol. 55:715–728.

Müller T, Philipp H, Dandekar T, Schultz J, Wolf M. 2007. Distinguishing species. RNA (New York, NY), 13:1469–1472.

Nylander JAA. 2004. MrModeltest v2.3 Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Pan H, Yang W, Zhang Y, Yang M, Feng R, Wu W, Guo D. 2015. An integrated strategy for the systematic characterization and discovery of new indole alkaloids from Uncaria rhynchophylla by UHPLC/DAD/LTQ-Orbitrap-MS. Anal Bioanal Chem. 407:6057–6070.

Rampersad SN. 2014. ITS1, 5.8S and ITS2 secondary structure modelling for intra-specific differentiation among species of the Colletotrichum gloeosporioides ssp. lato species complex. SpringerPlus. 3:684.

Reuter JS, Mathews DH. 2010. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinform. 11:129.

Ridsdale CE. 1978. A revision of Mitragyna and Uncaria (Rubiaceae). Blumea. 24:43–100.

Rognquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Lartef B, Liu, Suchard MA, Huesenbeck JP. 2012. MrBayes 3:2. efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 61:539–542.

Schultz J, Maisel S, Gerlach D, Müller T, Wolf M. 2005. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. RNA (New York, NY). 11:361–364.

Schultz J, Müller T, Achtziger M, Seibel PN, Dandekar T, Wolf M. 2006. The internal transcribed spacer 2 database—a web server for (not only) low level phylogenetic analyses. Nucleic Acids Res. 34:W704–W707.

Seibel PN, Müller T, Dandekar T, Schultz J, Wolf M. 2006. 4SALE-A tool for synchronous RNA sequence and secondary structure alignment and editor. BMC Bioinform. 7:43.

Seibel PN, Müller T, Dandekar T, Wolf M. 2008. Synchronous visual analysis and editing of RNA sequence and secondary structure alignments using 4SALE. BMC Res Notes. 1:91.

Shahzib SUA, Vďačný P, Kim JH, Jang SW, Shin MK. 2016. Molecular phylogeny and species delimitation within the ciliate genus Spirostomum (Ciliophora, Postciliodesmatophora, Heterotrichea), using the internal transcribed spacer region. Int J Mol Sci. 17:1–14.

Smith JS, Kim HG, Ju MS, Choi JG, Jeong SY. Oh MS. 2009. Effects of the hook of Uncaria rhynchophylla on neurotoxicity in the 6-hydroxydopamine model of Parkinson’s disease. J Ethnopharmacol. 126:361–365.

Smith C, Heyne S, Richter AS, Will S, Backofen R. 2010. Freiburg RNA tools: a web server integrating IntaRNA, ExpRNA and LocARNA. Nucleic Acids Res. 38 W373–W377.

Song C, Fan Y, Huang W, Wu D, Hu Z. 2000. Different hypotentive effects of various active constituents isolated from Uncaria rhynchophylla. Chin Trad Herb Drugs. 31:762–764.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30:2725–2729.

Tang Y, Wu Y, Huang R, Chao N, Liu Y, Xu P, Li K, Cai D, Luo Y. 2016. Molecular identification of Uncaria (Gouteng) through DNA barcoding. Chin Med. 11:14.

Tao C, Hua Z, Jiariu C, Taylor CM, Ehrendorfer F, Lantz H, Funston AM, Puff C. 2011. Uncaria. In: Wu Z, Peter HR, editors. Flora of China. Beijing, China: Science Press; p. 348–353.

Tinoco I, Uhlenbeck OC, Levine MD. 1971. Estimation of secondary structure in ribonucleic acids. Nature. 230:362–367.

Wang X, Chen X, Yang P, Wang L, Han J. 2017. Barcoding the Dendrobium (Orchidaceae) species and analysis of the intragenomic variation based on the internal transcribed spacer 2. Biomed Res Int. 2017:1–10.

Wang Y, Nevling LI, Gilbert MG, Lamarck A, Loureiro O. 2007. Aquilaria. In: Flora of China. Beijing, China: Science Press; p. 214–215.

Will S, Joshi T, Hofacker IL, Stadler PF, Backofen R. 2012. LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. RNA (New York, NY). 18:900–914.

Will S, Reiche K, Hofacker IL, Stadler PF, Backofen R. 2007. Inferring non-coding RNA families and classes by means of genome-scale structure-based clustering. PLoS Comput Biol. 3:e65–691.

Wolf M, Chen S, Song J, Ankenbrand M, Müller T. 2013. Compensatory base changes in ITS2 secondary structures correlate with the biological species concept despite intragenomic variability in ITS2 sequences—a proof of concept. PLoS One. 8:e66726.

Xu P, Wu Y, Huang R, Zhou J, Luo Y, Li K, Tang Y, Liu Y. 2012. Polymorphism of three kinds of medicinal plants Uncaria from Guangxi by RAPD. Lishizhen Med Mater Med Res. 23:703–705.

Yao H, Song J, Liu C, Luo K, Han J, Li Y, Pang X, Xu H, Zhu Y, Xiao P, et al. 2010. Use of ITS2 region as the universal DNA barcode for plants and animals. PLoS One. 5:e13102.

Yu M, Liu K, Zhou L, Zhao L, Liu Y. 2016. Testing three proposed DNA barcodes for the wood identification of Dalbergia odorifera T. Chen and Dalbergia tonkinensis Prain. Holzforschung. 70:1–10.

Yuan D, Ma B, Yang J, Xie Y, Yuan Wang L, Zhang L, Jia Kano Y, Wu C. 2009. Anti-inflammatory effects of rhynchophylline and isorhynchophylline in mouse N9 microglial cells and the molecular mechanism. Int Immunopharmacol. 9:1549–1554.

Zhang L, Koyyalamudi SR, Jeong SC, Reddy N, Bailey T, Longvah T. 2013. Immunomodulatory activities of polysaccharides isolated from Taxillus chinensis and Uncaria rhyncophylla. Carbohydr Polym. 98:1458–1465.

Zhao B, Huang Y, Chen Q, Chen Q, Miao H, Zhu S, Zeng C. 2018. Characteristic component profiling and identification of different Uncaria species based on high-performance liquid chromatography-photodiode array detection tandem ion trap and time of flight mass spectrometry coupled with rDNA ITS sequence. Biomed Chromatogr. 32:e4119.

Zhong Y, Feng R. 1996. A comparison study between medicinal plant Uncaria rhynchophylla (MIQ.) MIQ. Ex Havil. and Uncaria rhynchophylloides How. Bull Bot Res. 16:67–76.

Zhu S, Chen J, Liu J, Zhou L, Sun Y, Zeng C. 2012. ITS sequence RFLP analysis of two Uncaria macrophylla Wall from different regions. Strait Pharm J. 24:42–44.

Zhu S, Zhou L, Yang M, Huang H, Sun Y, Zeng C. 2012. Molecular identification of one species of Uncaria sinensis (Oliv.) Havil. J Guangdong Pharm Univ. 28:29–32.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31:3406–3415.