Chemical and Genetic Discrimination of Cistanches Herba Based on UPLC-QTOF/MS and DNA Barcoding

Sihao Zheng, Xue Jiang, Labin Wu, Zenghui Wang, Linfang Huang*

Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, People's Republic of China

Abstract

Cistanches Herba (Rou Cong Rong), known as “Ginseng of the desert”, has a striking curative effect on strength and nourishment, especially in kidney reinforcement to strengthen yang. However, the two plant origins of Cistanches Herba, Cistanche deserticola and Cistanche tubulosa, vary in terms of pharmacological action and chemical components. To discriminate the plant origin of Cistanches Herba, a combined method system of chemical and genetic –UPLC-QTOF/MS technology and DNA barcoding–were firstly employed in this study. The results indicated that three potential marker compounds (isomer of campneoside II, cistanoside C, and cistanoside A) were obtained to discriminate the two origins by PCA and OPLS-DA analyses. DNA barcoding enabled to differentiate two origins accurately. NJ tree showed that two origins clustered into two clades. Our findings demonstrate that the two origins of Cistanches Herba possess different chemical compositions and genetic variation. This is the first reported evaluation of two origins of Cistanches Herba, and the finding will facilitate quality control and its clinical application.

Introduction

Cistanches Herba (Rou Cong Rong), known as “Ginseng of the desert”, originates from dried succulent stems of Cistanche deserticola Y.C. Ma and Cistanche tubulosa (Schrenk) Wig according to the Chinese Pharmacopoeia (2010 edition), and is popular for its tonifying the kidney-yin, benefiting life essence and relaxing bowel. Currently, Cistanches Herba is mainly distributed in arid and warm deserts in northwest China, particularly in Xinjiang and Inner Mongolia provinces. However, the two origins of Cistanches Herba differ in terms of their pharmacological activity and chemical components. Tu et al. investigated the decoction of three Cistanches species (C. deserticola, C. tubulosa, Cistanche salsa) and found that C. tubulosa showed the lowest effect in the Yang-deficiency mouse model [1]. Zhang et al. compared pharmacological activity between C. deserticola, C. tubulosa and C. salsa, and found that these species had medicinal functions such as anti-fatigue and hypoxia tolerance, but not on the same extent [2]. Previous research reported the chemical component, and indicated the difference of chemical component and content for plant origins of Cistanches Herba [3]. As for the clinical application and market circulation, as a tonic, C. tubulosa has been traditionally used as a blood circulation-promoting agent and in the treatment of impotence, sterility, lumbago, body weakness in Japan [4–8]. Consequently, it is of great significance to discriminate two origins of Cistanches Herba for the quality control and clinical application. However, there is no research focus on discrimination of two origins of Cistanches Herba. Many researched methods, including microscopy, ultraviolet and infrared detection, intersimple sequence repeats method have been used to identify the genus of Cistanches, but not only for two origins specially [9–20].

Here, we conjunctively utilized chemical and molecular techniques to distinguish two origins of Cistanches Herba, UPLC-QTOF/MS (ultra-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry) and DNA barcoding. UPLC-QTOF/MS provides information more rapidly and efficiently compared with other techniques. The high selectivity and sensitivity of UPLC-QTOF/MS have resulted in its application for both quantitative and qualitative analyses, as well as in metabolite analysis and identification of complex compounds in Traditional Chinese Medicine [21–22]. Principal component analysis (PCA) and orthogonal projection to latent structure-discriminant analysis (OPLS-DA) are also developed to identify potential marker compounds. DNA barcoding, an easier and more universal molecular marker technology, uses a DNA fragment to identify species or genera. It is objective, more accurate, and easier to perform than traditional identification methods and other molecular marker technologies. Moreover, DNA barcoding has successfully been applied to identify animal and plant, including medicinal plants [23–26].

The purpose of this research is to establish a scientific method system, combined UPLC-QTOF/MS and DNA barcoding, for discrimination of two plant origins of Cistanches Herba.

Materials and Methods

Ethics statement

We confirm that the field studies did not involve endangered or protected species. GPS coordinates have included in the sample information, please see table 1.”
Plant materials and reagents

Succulent stems of Cistanthe Herba were collected from wild desert region in Inner Mongolia, Qinghai Provinces, Xinjiang Uygur Autonomous Region, People’s Republic of China (Table 1) in May 2012. The samples of the research were all collected in wild desert region, not in private land, where no specific permissions were required. The botanical identities of the stems were confirmed by Dr. Linfang Huang. Voucher specimens were deposited at The Institute of Medicinal Plant Development. High-performance liquid chromatography (HPLC)-grade acetonitrile (Merck KGaA, Darmstadt, Germany) and formic acid (Tedia, USA) were utilized for UPLC analysis. Deionized water was purified using a Milli-Q system (Millipore, Bedford, MA, USA). All other chemicals were of analytical grade.

Sample preparation

Cistanthe Herba samples (1.0 g, 65-mesh) were transferred into a 50-mL conical flask, and 50 mL of 70% methanol was added. After soaking for 30 min, ultrasonication (35 kHz) was performed at room temperature for 30 min. After centrifugation at 10,000 rev/min for 10 min, the supernatant was stored at 4°C and filtered through a 0.22-μm membrane before injection into the UPLC-QTOF/MS system for analysis.

UPLC-QTOF/MS

For UPLC analysis, the following systems/parameters were used: Waters Acquity system (Waters) equipped with a binary solvent delivery pump, auto-sampler and PDA detector connected to a Waters Empower 2 data station; ultrasonication (250 W, 50 kHz, Kunshan Ultrasonic Instrument Co., Zhejiang, China); and an electronic analytical balance model AB155-2 (Mettler-Toledo, Greifensee, Zurich, Switzerland). A Waters Acquity UPLC BEH C18 column (1.7 μm, 2.1×100 mm, Waters) and a Waters C18 guard column (same material, waters) were used and maintained at 30°C. The mobile phase was 0.1% formic acid aqueous solution (A) and acetonitrile (B) with a gradient program as follows: 0–3 min, 10–22% B; 3–4 min, 22–23% B; 4–6 min, 23–35% B; 6–8 min, 35–37% B; 8–11 min, 37–42% B; 11–12 min, 42–46% B; 12–15 min, 48–50% B at a flow rate of 0.3 mL/min. The injection volume was 5 μL.

The UPLC/MS analysis was performed on a QTOF Synapt G2 HDMS system (Waters, Manchester, UK) equipped with an electrospray ionization [ESI] source operated in the negative-ion mode. N2 was used as the desolvation gas. The desolvation temperature was set at 450°C at a flow rate of 800 L/h, and the source temperature was set at 120°C. The capillary and cone voltages were set to 2500 and 40 V, respectively. Data were collected between 50–1200 Da with a 0.1-s scan time and a 0.01-s interscan delay over a 15-min analysis time. Argon was used as the collision gas at a pressure of 7.06661023 Pa. All MS data were collected using the LockSpray system to ensure mass accuracy and reproducibility. The [M-H]− ion of leucine-enkephalin at m/z 554.2615 was used as the lock mass in negative ESI mode.

Data analysis

UPLC-QTOF/MS data for Cistanthe Herba samples were analyzed to identify potential discriminant variables. Peak finding, alignment and filtering of ES raw data were carried out using the Marker Lynx applications manager, version 4.1 (Waters, Manchester, UK). The parameters used were as follows: retention time (tR) of 0–15 min, mass of 50–1200 Da, retention time tolerance of 0.02 min, and mass tolerance of 0.02 Da. Three replicate samples collected from each geographic location were used (n = 3). A total of 6,339 variables were used to create the model.

DNA barcoding: DNA extraction, PCR amplification and sequencing

Samples taken from dried fleshy stems of C. deserticola and C. tubulosa (30 mg) were rubbed for 2 minutes at a frequency of 30 r/min. DNA was extracted according to the manufacturer’s instructions (Tiangen). Specifically, the protocol was modified such that chloroform was replaced with a mixture of chloroform: isoamyl alcohol (24:1 in the same volume), and buffer solution GP2 with isopropanol (same volume). The rubbed powder was put into a 1.5 ml eppendorf tubes, added 700 μL isopropanol (same volume), and buffer solution GP2 with isopropanol (same volume). The rubbing powder was put into a 1.5 ml eppendorf tubes, added 700 μL 65°C preheated GP1 and 1 μL β-mercaptoethanol to mix using vortex for 10–20 s, and incubated for 60 minutes at 65°C. Adding 700 μL mixture of chloroform: isoamyl alcohol (24:1), centrifuge for 5 minutes at 12000 rpm (~13400 x g); Pipette supernatant to a new tube, adding 700 μL isopropanol, blending for 15–20 minutes; Piping all the mixture into spin column CB3 and centrifuge for 40 s at
12000 rpm; Discarding the filtrate and adding 500 μL GD (adding quantitative anhydrous ethanol before use), centrifuge at 12000 rpm for 40 s; discarding the filtrate and adding 700 μL PW (adding quantitative anhydrous ethanol before use) to wash the membrane, centrifuge for 40 s at 12000 rpm; Discarding the filtrate and adding 500 μL PW, centrifuge for 40 s at 12000 rpm; Discarding the filtrate and centrifuge for 2 minutes at 12000 rpm to remove residual wash buffer PW; Transferring the spin column CB3 into a clean 1.5 ml eppendorf tube, and drying at room temperature for 3–5 minutes; Centrifuge for 2 minutes at 12000 rpm to obtain the total DNA. Primers for polymerase chain reaction (PCR) were based on sequences reported previously [4,5]. PCR reaction mixtures contained 2-μL DNA template, 8.5-μL ddH2O, 12.5-μL 2 × Taq PCR Master Mix (Beijing TransGen Biotech Co., China), 1/1-μL forward/reverse (F/R) primers (2.5 μM), in a final volume of 25 μL. PCR amplification was conducted as described by Kress et al. [4]. The primer of PCR reaction were fwd PA: GTTATGCATGAACGTAATGCTC (5′-3′) and rev TH: CGCGCATGGTGATTCACAATCC (5′-3′). PCR products were separated and detected by 1% agarose gel electrophoresis. PCR products were purified following the manufacturer’s protocol and directly subjected to sequencing.

Sequence alignment and analysis

ITS and ITS2 sequences were collected from the GenBank database. Sequences from sequencing of the samples were submitted to GenBank database (Accession numbers were listed in table 1), assembled with CodonCode Aligner 3.7.1 (CodonCode Co., USA) and aligned using ClustalW. Kimura 2-Parameter (K2P) distances, GC content of base and Neighbor-joining (N-J) trees were calculated and constructed using the MEGA 5.05 with the Bootstrap method (1000 resampling) and K2P model [27]. Barcoding gap (spacer region that was formed between intra- and inter-specific genetic variations) and identification efficiency (the ability of identification for comparing different barcodes) were drawn and calculated based on the method reported by Meyer and Paulay [28].

Results

Tentative peak assignment by UPLC-QTOF/MS

Representative chromatograms of C. deserticola and C. tubulosa from different producing areas are shown in Figure 1. The fingerprint chromatogram indicated similarities among Cistanches Herba samples. A total of 23 qualified mass peaks were detected and 16 peaks were identified by matching the retention times and mass spectra with those reported previously (Table 2) [29–35]. Peaks 2, 3, 5, 6, 7, 8, 9, 11, 12, 15, 16, 17, 20, 21, 22, and 23 were tentatively identified as cistanoside F, mussaenoside acide, cistanbuloside C1/C2, campneoside II, isomer of campneoside II, echinacoside, cistanoside A, acteoside, isoacteoside, syringalide A-3′-α-L-rhamnopyranoside, cistanoside C, 2′-acetylcampneoside, osmanthuside B, cistanoside D, tubuloside B, and cistancinenside A, respectively. Chemical constituents were determined to be primarily phenylethanoid glycosides (PhGs), while one compound, mussaenoside acide, was an iridoid polysaccharide. PhGs are the main active compounds in terms of treatment of kidney deficiency, and antioxidant and neuroprotective effects [36].

PCA of C. deserticola and C. tubulosa

PCA was employed to distinguish samples of different plant species. PCA is an unsupervised multivariate data analysis method that aims to visualize the similarities and/or differences within multivariate data of secondary metabolite composition [37]. The two-component PCA model cumulatively accounted for 46.04% of the variation (PC1, 36.43%; PC2, 9.61%). Figure 2 shows that 24 samples were clustered into two groups in the PCA scores plotted according to species origin, indicating that the chemical composition of C. deserticola and C. tubulosa differed significantly.

Figure 1. The representative chromatograms of C. deserticola and C. tubulosa. The left side were the chromatograms of C. deserticola collected from different locations; the right side were the chromatograms of C. tubulosa collected from different locations. This figure shows the differences between these two origins in chemical profiles.

doi:10.1371/journal.pone.0098061.g001
## Table 2. Tentatively identified compounds from *C. deserticola* and *C. tubulosa.*

| Peak no. | t<sub>r</sub> (min) | Name                                | Formula      | [M-H] m/z | Mean measured mass (Da) | Theoretical exact mass (Da) ppm | Fragments m/z                      | Reference |
|----------|----------------------|-------------------------------------|--------------|-----------|-------------------------|---------------------------------|-----------------------------------|-----------|
| 1        | 0.84                 | unknown                             | C₁₂H₂₂O₁₁    | 341.1078  | 341.1084  -1.8           | 195.0512, 179.0585               |                                   |           |
| 2        | 3.69                 | cistanoside F                        | C₂₁H₂₈O₁₃    | 487.1454  | 487.1452  0.4            | 179.0355                        |                                   | [13]      |
| 3        | 4.00                 | mussaenoside acide                   | C₁₆H₂₈O₁₀    | 375.1289  | 375.1291  -0.5           | 235.0561, 213.0758, 191.0689, 169.0860 |                                   |           |
| 4        | 4.87                 | unknown                             | C₁₇H₂₈O         | 691.2054  | 691.2062  -1.2           | 307.0252, 161.0243               |                                   |           |
| 5        | 5.89                 | cistanbuloside C1/C2                | C₃₅H₄₆O₂₁    | 801.2449  | 801.2453  -0.5           | 621.2043, 201.0146, 179.0353, 161.0230 |                                   | [15]      |
| 6        | 6.58                 | campneoside II                      | C₂₉H₃₈O₃      | 639.1947  | 639.1960  -2.0           | 621.2183, 179.0359, 161.0235     |                                   | [16]      |
| 7        | 6.66                 | Isomer of campneoside II            | C₂₉H₃₈O₃      | 639.1947  | 639.1960  -2.0           | 621.2183, 179.0359, 161.0235     |                                   | [16]      |
| 8        | 6.89                 | echinacoside                         | C₁₆H₂₆O₁₁    | 785.2517  | 785.2504  1.7            | 623.2201, 477.1612, 461.1643, 161.0244, 135.0445 |                                   | [16–18]   |
| 9        | 7.64                 | cistanoside A                       | C₃₆H₄₈O₂₀    | 801.2449  | 801.2453  0.9            | 621.2043, 201.0146, 179.0353, 161.0230 |                                   | [16–18]   |
| 10       | 7.51                 | unknown                             | C₂₂H₃₈O₁₂    | 493.2274  | 493.2285  -2.2           | –                                |                                   |           |
| 11       | 8.28                 | acteoside                           | C₂₀H₃₈O₁₅     | 623.2189  | 623.1976  2.1            | 421.1646, 319.0787, 161.0237     |                                   | [16–18]   |
| 12       | 9.48                 | isoacteoside                        | C₂₀H₃₈O₁₅     | 623.2189  | 623.1976  2.1            | 421.1646, 319.0787, 161.0237     |                                   | [16–18]   |
| 13       | 8.78                 | unknown                             | C₁₈H₂₆O₈      | 345.1552  | 345.1549  0.9            | 205.0836, 179.0338, 161.0226     |                                   |           |
| 14       | 9.36                 | unknown                             | C₁₈H₂₆O₆      | 345.1548  | 345.1549  -0.3           | 247.0938, 205.0836, 165.0902, 161.0236 |                                   |           |
| 15       | 9.57                 | syringalide A-3′-α-L-rhamnopyranoside | C₂₉H₄₂O₁₄    | 607.2023  | 607.2027  -0.7           | 445.1716, 299.1123, 201.0165, 161.0235 |                                   | [16]      |
| 16       | 9.89                 | cistanoside C                       | C₂₀H₃₈O₁₅     | 637.2137  | 637.2132  0.8            | 475.1777, 461.1655, 315.1072, 179.0344, 161.0245 |                                   | [16]      |
| 17       | 10.35                | 2′-acetylacteosid                    | C₂₀H₃₈O₁₅     | 665.2086  | 665.2082  0.6            | 645.1805, 623.2011, 503.1779, 461.1668, 315.1079, 161.0242 |                                   | [16–18]   |
| 18       | 11.09                | unknown                             | C₂₀H₃₈O₈      | 347.1076  | 347.1706  0              | 369.1531, 347.1715, 207.1012, 161.0245 |                                   | [16–18]   |
| 19       | 11.31                | unknown                             | C₂₀H₃₈O₈      | 367.2135  | 367.2132  0.5            | 637.2114, 591.2069, 487.3010, 451.3267 |                                   |           |
| 20       | 11.43                | osmanthuside B                      | C₂₀H₃₈O₁₃     | 591.2076  | 591.2078  -0.3           | 467.1423, 45.1707, 161.0224, 145.0284 |                                   | [16]      |
| 21       | 11.74                | cistanoside D                       | C₂₁H₄₀O₁₅     | 651.2287  | 651.2289  -1.7           | 677.4901, 527.1442, 351.1073, 175.0393 |                                   | [19]      |
| 22       | 12.17                | tubuloside B                        | C₂₁H₄₀O₁₅     | 665.2086  | 665.2082  0.6            | 461.1709, 161.0244               |                                   | [16]      |
| 23       | 12.38                | Cistancinenside A                   | C₂₁H₄₀O₁₅     | 679.2253  | 679.2238  2.2            | 677.4927, 659.1908, 179.0355, 161.0236 |                                   | [15]      |
OPLS-DA and marker identification

To identify potential chemical markers for discrimination of the two species, the S-plot of OPLS-DA was generated (Fig. 3). In the S-plot, each point represents one \( t_R - m/z \) ion pair. The X and Y axes represent the contribution and confidence of the ion, respectively; the farther the distance the ion \( t_R - m/z \) pair points from zero, the larger the contribution/confidence of this ion is to the difference between the two groups. Thus, the \( t_R - m/z \) ion pointing to the two ends of the ‘S’ represent the characteristic markers with the highest confidence in each group.

The OPLS-DA results showed that UPLC-QTOF/MS could be used to distinguish *C. deserticola* from *C. tubulosa* (Fig. 3). A total of six credible and significant markers were determined to facilitate discrimination of these groups (Table 3). The identities of three potential markers were tentatively assigned. The components correlated with these three ions were tentatively identified as isomers of campneoside II, cistanoside C and cistanoside A. The marker compounds \( a \), \( b \) and \( c \) could be used to distinguish the two plant species, as the ion intensities of \( a \) and \( b \) in *C. deserticola* was higher than in *C. tubulosa* (Fig. 4A, 4B), and marker \( c \) could be detected in *C. tubulosa*, but not in *C. deserticola* (Fig. 4C).
DNA barcoding: sequence information and identification efficiency

Sequence information was shown in Table 4. The average genetic distance of *psbA-trnH* (0.1732) was larger than other two regions (0.0740, 0.1197) significantly. The average GC content of *psbA-trnH* (20.64%) was smaller than other two regions (55.00%, 55.00%). Though the success rate of ITS and ITS2 was not obtained in this study, the *psbA-trnH* region performed well in PCR amplification and sequencing (100%, 87.23%). Identification efficiency was achieved by BLAST1 analysis and the nearest-distance method, and mainly reflected the success rate of the barcodes. The *psbA-trnH* region was clearly higher than the other two barcodes in identification efficiency based on two methods. The shortage of sequences is most likely the reason that ITS region exhibited 100% identification efficiency based on BLAST1 method, and 0 based on the nearest-distance method.

Analysis of genetic divergence using six parameters

Six parameters were used to analyze intra-specific variation and inter-specific divergence using three barcodes (Table 5). The significant difference between inter- and intra-specific variations was indicative of the utility of the DNA barcodes. Here, the minimum interspecific distance of three barcodes was all higher than the maximum intraspecific distance. Moreover, *psbA-trnH* region had larger maximum intraspecific distance and average interspecific distance than the other two barcodes, indicated that *psbA-trnH* region performed well in discrimination of two origins of Cistanches Herba.

Table 3. Marker *t*<sub>r</sub>–m/z ion pairs in the S-plot.

| Species      | Marker ions                                      |
|--------------|--------------------------------------------------|
| *C. deserticola* | 7.64–799.2667<sub>a</sub>; 9.88–637.2135<sub>b</sub>; 0.84–341.1081 |
| *C. tubulosa*    | 6.66–639.1929<sub>c</sub>; 8.15–493.2284; 4.27–405.0499 |

doi:10.1371/journal.pone.0098061.t003
Chemical and Genetic Discrimination of Cistanches Herba

Table 4. Identification efficiency of three loci using different methods for species identification.

| Markers          | ITS     | ITS2    | psbA-trnH |
|------------------|---------|---------|-----------|
| Number of sequences | 4       | 18      | 37        |
| Length range/bp   | 617–622 | 235–236 | 358–558   |
| Average GC content | 55.00%  | 55.00%  | 20.64%    |
| Efficiency of PCR amplification/% | /      | /       | 100       |
| Success rate of sequencing/% | /      | /       | 87.23     |
| Genetic distance  |         |         |           |
| Min               | 0.0016  | 0.0000  | 0.0000    |
| Max               | 0.1280  | 0.2379  | 0.5279    |
| Average           | 0.0740  | 0.1197  | 0.1732    |
| Identification efficiency/% | BIAST1 | 5.56    | 0        |
| Nearest Distance  | 100     | 5.56    | 100       |

doi:10.1371/journal.pone.0098061.t004

Analysis of barcoding gap to identify C. deserticola and C. tubulosa

The barcoding gap presents the remarkable variation of inter-and intra-species, and demonstrates that separate, non-overlapping distributions between intra- and inter-specific samples. In this study (Fig. 5), the distance range was set to 0–0.45, because the greatest K2P distance of psbA-trnH between C. deserticola and C. tubulosa was close to 0.45. The three barcodes exhibited distinct gaps in the distributions of intra- and inter-specific variation. Furthermore, the gap of psbA-trnH was significantly larger than other two barcodes. Therefore, psbA-trnH region could be an ideal barcode for discriminating two origins of Cistanches Herba.

Neighbor-joining (NJ) tree

An NJ tree illustrates the relationship among species and facilitates determination of their clustering. In this study, NJ tree of three barcodes were built based on K2P model (Fig. 6). The results demonstrated that two origins of Cistanches Herba clustered into two clades separately. Thus, the NJ tree clearly distinguished between C. deserticola and C. tubulosa.

Discussion and Conclusions

Cistanches Herba is an important medicinal material commonly used to nourish in the Asian community [30]. However, the two origins of Cistanches Herba, C. deserticola and C. tubulosa, have different chemical compositions and pharmacological activities respectively. Concurrently, the two origins differ in clinical and commodity market. The classification of Cistanches Herba is confused and massive substitute and adulterants flood the market due to the shortage of resources and special growing environment for Cistanches Herba. Genus of Cistanche is accepted to include four species and one variant: C. deserticola, C. tubulosa, C. sinensis, C. salsa, and C. salsa var. albiflora [39]. Researchers in Japan considered the origin of Cistanches Herba as C. salsa [40–43], while it was identified as C. deserticola by Tu [44–46]. Therefore, it is confused in classification of Cistanches, and it is hard to discriminate the two origins of Cistanches Herba.

Traditional methods for quality control of Cistanches Herba are morphological identification [47,48], microscopic identification [49] and TLC (Thin-Layer Chromatography) [50,51], FTIR (Fourier Transform Infrared Spectroscopy) [14], HPLC (High Performance Liquid Chromatography) [52,53]. Morphological and microscopic method can easily differentiate species from different genera or families that possess big difference in morphological and microscopic characteristics, while it is hard to distinguish sibling species. TLC and FTIR can clearly discriminate species that possess different kind of chemical compositions, whereas it is difficult to determine the chemical component and content. HPLC is mainly used for differentiating species with different chemical elemente contents, nevertheless, the time of analysis is longer and the sensitivity is relatively lower compared to UPLC. Correspondingly, UPLC-QTOF/MS technology was faster and more accurate in determining chemical composition than other chemical methods. Molecular identification methods exhibit well in discrimination based on the genetic variation, such as SDS-PAGE (Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis) [54], AFLP (Amplified Fragment Length Polymorphism) [55,56]. However, these molecular methods are not easy to operate and are not universal. Correspondingly, DNA

Table 5. Analysis of interbygenus-specific divergence and intra-specific variation of three barcodes.

| Markers                      | ITS     | ITS2    | psbA-trnH |
|------------------------------|---------|---------|-----------|
| Theta (avg_intra_avg)        | 0.0225  | 0.0000  | 0.0091±0.0019 |
| coalescent depth (avg_intra_max) | 0.0338  | 0.0000  | 0.0539±0.0013 |
| All intraspecific distance (avg_between_intra-species) | 0.0225±0.0181 | 0.0000 | 0.0102±0.0100 |
| Theta prime (avg_interbyG_avg) | 0.1258±0.0181 | 0.2381 | 0.4090±0.0100 |
| minimum interspecific distance (avg_interbyG_min) | 0.1235±0.0181 | 0.2381 | 0.2345±0.0100 |
| all inter-specific distance (avg_between_interbyG) | 0.1258±0.0025 | 0.0000 | 0.4090±0.0476 |

doi:10.1371/journal.pone.0098061.t005
barcoding could discriminate species more universally, quickly and accurately than other molecular methods. For the species from same genus and close genetic relationship, those methods alone may not perform well in identification. Here, we combined UPLC-QTOF/MS and DNA barcoding in identifying *C. deserticola* and *C. tubulosa*, and evaluated the chemical and molecular markers.
Figure 6. NJ tree of *C. deserticola* and *C. tubulosa* with three barcodes. NJ tree of *C. deserticola* and *C. tubulosa* with three barcodes were built. The bootstrap scores (1000 replicates) are shown (≥50%) for each branch. *C. deserticola* and *C. tubulosa* were clustered into two clades clearly.

doi:10.1371/journal.pone.0098061.g006
that would allow them to be discriminated. 23 qualified mass peaks were detected and 16 were identified by using UPLC–QTOF/MS, and three potential marker compounds were firstly found to facilitate the discrimination of two origins by PCA and OPLS-DA analysis. Furthermore, four indicators were assessed by DNA barcoding technology in terms of their ability to discriminate two origins: Identification efficiency, genetic efficiency, barcoding gap, and NJ tree analysis. The psbA-trnH region was supported as a suitable DNA barcode for discriminating C. deserticola and C. tubulosa.

In conclusion, we firstly established a new molecular and chemical analysis-combined method for discriminating and quality control in two origins of Cistanches Herba. DNA barcoding can discriminate two origins in genetic variation and authenticate species universally and accurately; UPLC-QTOF/MS technology can analyze chemical composition to evaluate the quality of medicinal materials rapidly and accurately. The combined method of DNA barcoding and UPLC-QTOF/MS technology guarantees the identification in multiple sources of medicinal materials more accurately and scientifically, and may serve as a method for identifying other confusing species or genus in classification.

Author Contributions
Conceived and designed the experiments: LFH SHZ LBW. Performed the experiments: SHZ LBW. Analyzed the data: SHZ LBW. Contributed reagents/materials/analysis tools: LFH. Wrote the paper: SHZ XJ LBW.

Check the manuscript: ZHW.
41. Kobayashi H, Karasawa H, Miyase T, Fukushima S (1984) Studies on the constituents of Cistanchis herba III. Chem Pharm Bull 32: 3009–3014.
42. Kobayashi H, Karasawa H, Miyase T, Fukushima S (1984) Studies on the constituents of Cistanchis herba IV. Chem Pharm Bull 32: 3880–3885.
43. Kobayashi H, Karasawa H, Miyase T, Fukushima S (1985) Studies on the constituents of Cistanchis herba VI. Chem Pharm Bull 33: 3645–3650.
44. Moriya A, Tu PF, Karasawa D, Arima H, Deyama T, et al. (1995) Pharmacognostical studies of Cistanchis Herba (J). Nat Med 49: 383–393.
45. Moriya A, Tu PF, Karasawa D, Arima H, Deyama T, et al. (1995) Pharmacognostical studies of Cistanchis Herba (II). Nat Med 49: 394–400.
46. Liu XM, Jiang Y, Sun YQ, Xu XW, Tu PF (2011) Study on chemical constituents of Cistanche deserticola. Chin Pharm J 46: 1053–1058.
47. Gu XY, Wang X (2013) Identification between the Original Plants of RouCongRong and Confused Varieties. Western Journal of Traditional Chinese Medicine 26:17–18.
48. Zhang HJ, Ding XF Identification of Cistanche deserticola and adulterants. Shizhen Medicine and Materia Research, 183.
49. Xu XQ, Ni H, Wei HY, Jia XG, Zhang BG, et al. (2013) microscopic identification research of three kinds of Cistanche originated from Xinjiang province. Shizhen Medicine and Materia Research 24: 881–883.
50. Huang M, Liu G (2004) Identification research of Cistanche deserticola and Cynomorium songaricum. HuRei journal of traditional Chinese medicine. 26:54–55.
51. Liu YG, Xu R, Wang W, Liu TN, Chen J (2009) Research progress on quality control and assessment of Cistanche deserticola Y. C. Ma. World Science and Technology- Modernization of Traditional Chinese Medicine. 11:439–444.
52. Ma ZG (2011) Differentiation of processed Cistanche sinensis G. Beck and Herba Cistanches by HPLC fingerprints. Chin. Pharm. J. 46: 899–902.
53. Huang LX, Wang YE, Shi MH, Jia XG, Xie X, et al. (2011) HPLC fingerprints research of Cistanche tubulosa. Xinjiang journal of traditional Chinese medicine. 29: 54–56.
54. Chen P, Guo YH, Wang BM, Zhai ZX (2006) SDS-PAGE of soluble proteins in four plants of Cistanche Hoffing, Et Link. Chinese Traditional and Herbal Drugs 37: 1399-1402.
55. Xu R, Chen J, Chen SL, Liu TN, Na R (2007) AFLP Analysis on diversity of germplasm resource in cultivated and wild Cistanche deserticola. Chinese Traditional and Herbal Drugs 38: 1703-1707.
56. Gan XY, Li M, Ma HA, Song YX (2011) Study on intraspecific variation of Cistanche deserticola Y.C.Ma using AFLP molecular markers. Northern Horticulture, 141–143.