High Levels of Genetic Diversity within One Population of Rheum tanguticum on the Qinghai-Tibet Plateau have Implications for Germplasm Conservation

Xiaochong Ma, Caixiang Xie, Meng Guan, Xiuhai Xu, Eiji Miki, Osami Takeda, Tianyi Xin, Sihao Zheng, Hui Yao, Linchun Shi, Jingyuan Song, and Shilin Chen

1Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, 151 Malianwa North Road, Haidian District, Beijing 100193, China; 2China Meheco Corporation, Beijing 100061, China; 3Botanical Raw Material Research Department, TSUMURA & CO, Japan

Abstract: Rheum tanguticum, a source plant for the traditional Chinese medicine Dahuang (rhubarb), is an endangered species. In this study, we first assessed genetic diversity using 102 internal transcribed spacer (ITS) sequences cloned from one population of R. tanguticum on the Qinghai-Tibet Plateau. In total, 1773 polymorphic sites were identified, including 707 sites in the ITS1 region, 773 sites in the ITS2 region, and 293 sites on the 5.8S gene. Overall, transitions were the most common type of mutation and represented 77% of all base substitutions found within the ITS regions of R. tanguticum. We found that the average intra-specific genetic distance was 0.1260±0.0755 and ranged from 0 to 0.3618. Furthermore, a neighbor-joining phylogenetic tree constructed using these sequences strongly supported the conclusion that abundant intra-specific variation exists within the ITS regions of R. tanguticum. Interestingly, we found that eighty-seven clones of R. tanguticum had high similarity to the ITS sequence of R. rhabarbarum and ten clones had high similarity to the ITS sequence of R. officinale; however, five clones of R. tanguticum had similarity to Fagopyrum esculentum. Therefore, we observed a significant amount of ITS sequence variation at both the intra-genomic and intra-specific levels in R. tanguticum. Based on these results, we suggest that in situ conservation is preferable for R. tanguticum protection.

Keywords: Genetic diversity, intra-specific variation, Rheum tanguticum, internal transcribed spacer, endangered species conservation.

INTRODUCTION

Rheum tanguticum Maxim. Ex Balf. (Polygonaceae) is one of the three main sources of Dahuang (rhubarb), a well-known traditional Chinese medicine (TCM). Its phenolic compounds and polysaccharides have shown some promising pharmacological properties [1, 2]. The dried rhizomes and roots of R. tanguticum have been widely used as laxatives and anti-inflammatory agents in China [3], while the standardized rhubarb extracts, such as emodin, have become available on the international market [4, 5]. R. tanguticum, also known as Tangut rheumor chicken-feet rheum, is native to the Qinghai-Tibet Plateau at altitudes ranging from 2300 to 4200 m [6, 7]. In recent years, the wild resources of R. tanguticum were seriously threatened by over-exploitation and an extensive loss of habitat; thus, this species has been included on the important conservation of wild plants list for the Qinghai province [8]. Some measures have been taken to reduce the difference between the commercial demand and the wild resource depletion, such as cultivation in Guoluo [9]. However, reductions to wild R. tanguticum populations are a major concern for development of superior pharmaceutical crop cultivars of the Tangut rheum. Analysis of genetic variations in the endangered species at the population level is critically important for conservation strategies because sufficient genetic variation can increase the species chance of survival [10]. Therefore, the analysis of genetic variation within and between populations becomes important not only for the conservation of the endangered species [11-15] but also for pharmaceutical crop development.

The internal transcribed spacer (ITS) of nuclear ribosomal DNA (nrDNA) has been used extensively in studies on genetic diversity, phylogenetics, and evolution at the species or genera levels due to its fast rate of evolution and great variation [16-18]. The ITS region is composed of three parts: ITS1, ITS2, and the highly conserved 5.8S rDNA gene, which is located between ITS1 and ITS2 [19]. Compared to organellar DNA, the ITS of nrDNA, which is inherited from both parents, can provide much more information [20]. Furthermore, this region is simple to amplify using PCR and a set of universal primers, which are available for a wide range of organisms, and it can be easily sequenced due to its relatively small size [21]. Importantly, high levels of variation are usually observed within ITS regions and are ideal for revealing genetic diversity and population differentiation in many species [22-24]. However, to date, there have been no reports of ITS sequence variation at the intra-specific level in R. tanguticum. In previous reports, genetic diversity of R. tanguticum was revealed based on ISSR and SSR analysis.
ITS sequences within one population of *R. tanguticum* characterize the distribution of sequence variation based on the ITS sequences within one population of *R. tanguticum*, analyze potential reasons for sequence variations, and propose conservation strategies for this species.

**MATERIALS AND METHODOLOGY**

**Plant Materials**

Twenty-nine specimens of *R. tanguticum* were collected from the Baihe pasture of Zoige in the Sichuan Province of China (latitude: 33°18′35″N to 33°21′32″N; longitude: 102°27′12″E to 102°28′9″E; altitude: 3436-3459 m) (see the Appendix Table S1). A typical habitat for *R. tanguticum* is shown in (Fig. 1).

![Fig. (1). The habitat of *Rheum tanguticum.*](image)

**Genomic DNA Preparation and PCR Amplification**

Genomic DNA was extracted from 20 mg of silica gel-dried leaves using the Plant Genomic DNA Kit (Tiangen Biotech Co., China) after grinding for 1 min at 30 revolutions/s with a DNA Extraction Grinder (MM400; Retsch, Haan, Germany). Polymerase chain reaction (PCR) amplifications of the ITS region were carried out using a Peltier Thermal Cycler ABI9700 (ABI lab, Inc., USA) and the following pair of universal primers: ITS5F (5′-GGAAGTAAAGTCGTAACAAAGG-3′) [26]. PCR amplification was performed using 20-50 ng of genomic DNA as template and a total volume of 25 μL, which also contained the following: 1×PCR buffer (without MgCl₂), 2.0 mM MgCl₂, 0.2 mM each dNTP, 0.1 μM each ITS primer (synthesized by Sangon Co., China), and 1.0 U Taq DNA Polymerase (Bioron BioScience & Technology Co., China). PCR reaction conditions consisted of an initial denaturation step for 5 min at 94°C, followed by 35 cycles of denaturation for 1 min at 94°C, annealing for 1 min at 50°C and extension for 1.5 min at 72°C (+3s/cycle), with a final extension step at 72°C for 7 min [27]. The PCR products were separated on 1.0% agarose gels in 0.5×TBE buffer for visualization. Direct sequencing of the PCR products from *R. tanguticum* was unsuccessful, indicating the presence of polymorphisms within individual specimens.

**Cloning and Sequencing of the ITS Region**

The PCR products were purified using the TIAN quick Midi Purification Kit (Tiangen Biotech Co., China). The recovered products were ligated into the pMD19-T vector (Takara Biotech Co., China) following the addition of dATP residues to the 3′ ends of the PCR products using the DNA A-Tailing Kit (Takara Biotech Co., China). The cloned products were then transformed into *E. coli* DH5α cells using standard recombinant DNA techniques.

*E. coli* colonies were cultured in LB culture medium containing ampicillin, and following selection with colony PCR, five positive clones from each sample were sequenced using an ABI3730XL sequencer (Majorbio Co., Shanghai, China). In total, we obtained 105 clones from 29 samples (excluding fungal sequences). Three of the clones (including KF514611, KF514546 and KF514613) showed significant differences when compared with the other sequences (~90bp deletions at the start of the ITS region) and were not analyzed in this paper (see the sequence alignment in Appendix Fig. S1, S2, and S3).

**Sequence Alignment and Date Analysis**

The sequences of the ITS regions of *R. tanguticum* were aligned using CodonCode Aligner3.7.1 (CodonCode Co., USA). Genetic variations were analyzed with a Kimura 2-parameter (K2P) distance matrix, which was constructed using the MEGA5.0 software program [28]. The neighbor-joining (NJ) phylogenetic trees were built in the MEGA5.0 software program using the 102 sequenced clones from *R. tanguticum* as well as four other GenBank sequences from other species: *Rheum officinale* (JN187108), *Rheum nobile* (GQ206264), *Rheum rhabarbarum* (JQ288756), and *Fagopyrum esculentum* (EF653685). The 5.8S gene and the ITS1 region were obtained through sequence alignment with *R. rhabarbarum* and *F. esculentum*, and the ITS2 region was obtained using hidden Markov model (HMM)-based annotation methods [29].

**RESULTS**

**Sequence Length and GC Content**

The average and distribution of the ITS sequence lengths, as well as GC contents, are shown in (Fig. 2). The ITS sequence length ranged from 522-571 bp in the 102 clones from *R. tanguticum*. The average length of the ITS sequences was 559 bp, and the median length was similar to the average. The GC content of the ITS sequences from *R. tanguticum* ranged from 56- 69%, with an average value of 63% and a median of 62%.

**Sequence Variation and Polymorphic Distributions of the ITS Region within Individual Genomes**

We did not analyze insertions and deletions (indels) in the present study, as they are often spread across multiple sites, particularly in long homopolymeric tracts, and are difficult to identify in an accurate and reliable manner. Therefore, we limited ourselves to the detection and comparative
analysis of base substitution-only type polymorphisms (e.g., transitions and transversions).

Sample RTM10 of *R. tanguticum* is a typical representative of the samples (see the Appendix Table S1), and a multiple sequence alignment of five clones of this sample is shown in Fig. (4). There were 105 variable sites between clones, with the majority of variation existing between KF514583 and the three clones KF514579, KF514580, and KF514581. Only one variable site was observed between KF514583 and KF514582 and a small amount of variation was identified among the other three clones.

**Analysis of ITS Sequence Variations by Genetic Distance**

We divided the *R. tanguticum* mutations into two groups, containing intra-specific and intra-genomic variations, which were then analyzed using K2P genetic distances. In particular, we looked at the distribution of divergences within the classes at a resolution 0.02 distance units (Fig. 5). Our results showed that the average intra-specific distance was 0.1260±0.0755 and that *R. tanguticum* had a relatively high level of sequence divergence, ranging from 0 to 0.3618. In addition, K2P intra-genomic distances ranged from 0 to 0.3338 within individual *R. tanguticum* specimens (see the Appendix Table S1). These results demonstrate a high level of variation within the ITS sequences of *R. tanguticum*.

**Analysis of Phylogenetic Relationships Based on ITS Sequences Using NJ Trees**

We constructed a NJ tree based on the 102 ITS sequences of *R. tanguticum* as well as the four GenBank sequences from *R. officinale*, *R. nobile*, *R. rhabarbarum* and *F. esculentum* (Polygonaceae) (Fig. 6). The tree was divided into three groups: (1) ten clones from six samples of *R. tanguticum* (green region) that had high similarity to the ITS sequence of *R. officinale*, (2) a second group containing *R. nobile*, and (3) the remaining eighty-seven clones (including the blue, yellow, and red regions, as well as KF514638) that had high similarity to the ITS sequence of *R. rhabarbarum*. Overall, the NJ tree supported the hypothesis that the ITS sequences from *R. tanguticum* were polyphyletic. Unexpectedly, five clones of *R. tanguticum* (purple region) clustered most closely with *F. esculentum*.

**DISCUSSION**

The present study revealed a high level of intra-specific genetic polymorphisms in the ITS regions of *R. tanguticum*, which was consistent with the molecular evidence provided by ISSR and SSR analysis in previous reports [8, 25, 30]. The high levels of genetic diversity within the population of *R. tanguticum* may be due to several factors. First, mating systems are thought to be one of the most important factors determining genetic diversity in *R. tanguticum*. In the wild, rheum plants are self-incompatible [31], and it has been proven that self-incompatibility is crucial for the maintenance of high levels of genetic variability within many species [32]. Second, the physiological structure of *R. tanguticum* may be another important factor. Indeed, *R. tanguticum* possesses panicles containing numerous small flowers that can produce large amounts of trigonous achenes, which are small, light, and show a high germination rate between plants [33]. Therefore, for developmental reasons, this species may have more opportunity to accumulate genetic diversity.

**Fig. (2).** ITS sequence length and GC content in 102 clones of *Rheum tanguticum*. Box plots show the interquartile range (IQR) of the data, which is defined as the difference between the 75th and 25th percentiles. The solid and dotted lines in the box represent the median and average values, respectively, for length and GC content.

**Fig. (3).** The distribution of polymorphisms (base substitution-only) in the *Rheum tanguticum* rDNA array, as derived from ITS sequencing data and shown in 50-bp bins.
Fig. (4). The complete alignment matrix of the ITS sequence from sample RTM10 of *Rheum tanguticum*.
Fig. (5). The relative distributions of intra-genomic divergence (red) between clones within individual *Rheum tanguticum* samples and intra-specific variation (yellow) for the ITS sequences. x axis, K2P genetic distance; y axis, frequency.

Fig. (6). Circle Neighbor-Joining trees based on the ITS sequences of *Rheum tanguticum* and several closely related species. Bootstrap scores (1000 replicates) are shown (≥60%) for each branch.
After *lyptus impensa* formulating conservation strategies. The protection of characteristics and the living conditions of species when possible [44]. In addition, we also consider the physiological method was used to preserve it in as many populations as micropropagation methods, such as this species were difficult to obtain, thus tissue culture and due to frequent fire and disease. Additionally, the seeds of found in *ple*, due to low intra-specific variation and fewer wild re-

In addition, we captured some genetic information from *R. tanguticum* that had not been revealed previously when we studied genetic diversity between clones. From analyses of the phylogenetic tree, ten clones of *R. tanguticum* had high similarity to ITS sequence of *R. officinale*, which is supported by morphological classification [35]. However, eighty-seven clones had high similarity to ITS sequence of *R. rhabarbarum*, which was consistent with palynological research [36, 37]. Pollen morphology demonstrated that *R. tanguticum* and *R. rhabarbarum* were both microechinate-foveolate, while *R. nobile* and *R. officinale* were different. Unexpectedly, five clones of *R. tanguticum* were similar to *F. esculentum*. These two species belong to different genera of the Polygonaceae, which may indicate that the ITS of nrDNA inherited from parents can reveal horizontal gene-transfer and hybridization events in the species evolutionary history. The phenomena of identical variants across genera in nature were also found in the nuclear genome of *Eleutherococcus giraldii* and *Panax ginseng* of Araliaceae [38] and in the nuclear genome of *Dendropanax hainanensis* and *Merillipinax membranifolius* of Araliaceae [39].

We suggest that the preferable choice for protection of *R. tanguticum* is *in situ* conservation because of its high intra-specific genetic variation and its living environment. Genetic diversity is important to consider when formulating conservation strategies for species because it is critical to the capability of a species to adapt to environmental changes [40]. In other words, we must select appropriate protection methods to maintain as much genetic diversity as possible. For example, due to low intra-specific variation and fewer wild resources of *P. ginseng*, tissue culture, somatic embryogenesis, and cultivation were applied [41-43]. However, high genetic variation and highly structured geographic patterns were found in *Lamiophlomis rotate* (Lamiaceae), so *in situ* method was used to preserve it in as many populations as possible [44]. In addition, we also consider the physiological characteristics and the living conditions of species when formulating conservation strategies. The protection of *Euca-

In summary, the present study revealed a high level of genetic diversity in the ITS regions of *R. tanguticum*. In addition, we identified genetic information from *R. tanguticum* that has not been revealed previously. Our results showed that ten clones of *R. tanguticum* had high similarity to the ITS sequence of *R. officinale*, eighty-seven clones had high similarity to the ITS sequence of *R. rhabarbarum*, and five clones of *R. tanguticum* were similar to *F. esculentum*. Based on these results, *in situ* conservation is the preferable choice for protection of *R. tanguticum*.

**CONFLICT OF INTEREST**

The author(s) confirm that this article content has no conflicts of interest.

**ACKNOWLEDGEMENT**

We are thankful for the great support and friendly cooperation from China Meheco Co., Ltd, the Botanical Raw Material Research Department, TSUMURA & CO and *R. tanguticum* Research Establishment and Farm of Baitue Pasture, Zoige County, Sichuan Province. This work was supported by the Conservation on Natural Resources of Rharb, Program for Innovative Research Team in IMPLAD, and the National Natural Science Foundation of China Project (Grant no. 81373922).

**ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
</table>
| ITS      | Internal transcribed spacer of nuclear riboso-
| ITS1     | Internal transcribed spacer 1 of nuclear ribo-
| ITS2     | Internal transcribed spacer 2 of nuclear ribo-
| nrDNA    | Nuclear ribosomal DNA                          |
| TCM      | Traditional Chinese medicine                   |
| PCR      | Polymerase chain reaction                      |
| NJ tree  | Neighbor-joining phylogenetic tree             |
| K2P      | Kimura 2-parameter distance                    |
| MEGA5.0  | Molecular Evolutionary Genetics Analysis      |
| HMM      | Hidden Markov model                            |
| Indels   | Insertions and deletions                       |

**REFERENCES**


Received: August 28, 2013  
Revised: December 16, 2013  
Accepted: December 18, 2013

© Ma et al.; Licensee Bentham Open.

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