Genuine or Fake? A DNA Barcoding Study to Identify Ginseng Species in Traditional Chinese Medicine Market

Yifan Chen, Ningbing Xue, Lanxin Lei, Junhan Zhang
Ulink College of Shanghai, Wuxi Big Bridge Academy, Cold Spring Harbor Asia DNA Learning Center

Abstract

Recent studies have shown that counterfeit drugs are frequently found in the Traditional Chinese Medicine (TCM) market. It is especially difficult for taxonomists to distinguish between ginseng species due to their similar morphological characteristics. Some scientific papers inspired us to use DNA barcoding with both rbcL and ITS2 primers to identify the ginseng species we bought in different pharmacies, and we found that none of the samples were mislabelled.

Introduction

The traditional identification methods of traditional Chinese medicine are basic identification, trait identification, microscopic identification, and physical & chemical identification[1]. The principle is based on the analysis of the morphological characteristics of taxonomy. However, owing to the gradual diversification of traditional Chinese medicine, it is now difficult to accurately and effectively identify species based only on their phenotypes. Several studies indicated that mislabelling situations have been occurring in Traditional Chinese Medicine (TCM) markets, suggesting that the medicines made from stems and roots are more likely to be counterfeited[2][3]. In fact, the various external traits are due to the different genotypes, which means a difference in DNA sequences. With the development of DNA barcoding, the Chinese government has started systematizing the identification of traditional Chinese medicine and has proposed a universal barcode sequence using the ITS2 as main identified region. [1][4][5][7]

In this study, our aims are:

- Using DNA barcoding to examine whether the pharmacies had mislabelled the ginseng species
- To compare the efficiency of the rbcL and ITS2 primers according to the BLAST results, alignments and phylogenetic trees
- To evaluate the quality of DNA isolation using dried roots and fresh leaves

Materials and Methods

- A total of 23 samples were used.
- 5 fresh leaf samples were collected from a herb garden in Wuxi, and the rest are all dry root samples, including 8 samples bought from a pharmacy in Suzhou, 1 from a different pharmacy in Suzhou, 8 from a pharmacy in Shanghai and 1 from a pharmacy in Wuxi.
- The root samples contain polyphenol which may degrade DNA after oxidation and also polysaccharides and proteins which may interfere with PCR. Therefore, we used Qiagen DNeasy Plant Kit to extract DNA to minimize the effects of these impurities.
- A Polymerase Chain Reaction (PCR) was carried out to amplify the barcoding regions using plant ITS2 and rbcL primers.
- We used gel electrophoresis to confirm the success of PCR in amplifying the genes. Positive PCR products were sent to Genewiz for sanger sequencing.
- The DNA sequences were analyzed using DNA subway. We identified the species of the samples with the program BLAST and used MUSCLE to create alignments as well as phylogenetic trees between the samples.

Results

PCR Success

- 12/23 samples were successfully amplified using rbcL and 12/23 samples were successfully amplified using ITS2 primer.
- 9/23 samples were successfully amplified using both primers
- However, the rbcL bands of sample 005 and 016 on the gel were quite faint maybe due to either failure in DNA isolation or unsuccessful amplification, so they were not sent for sequencing.

Sequence Results

- The rbcL and ITS2 BLAST results of 5 samples are consistent with each other. (004, 010, 011, 013, 019)
- The rbcL and ITS2 BLAST results of 4 samples are not consistent with each other although they are same in genus level. (001, 003, 006, 020)
- Three samples are only successfully amplified by one primer so only one BLAST result is obtained

MUSCLE Results

The rbcL alignment results scatter across while the ITS2 alignment results cluster closely in the middle region.

Sample Number	Sample Picture	RbcL BLAST Result	ITS2 BLAST Result	Rbcl and ITS2 BLAST results match or not	Market Name	Correctly labelled OR Mislabelled
001 西洋参		Panax japonicus	Panax quinquefolius	No	Xiyangshen/ Panax quinquefolius	Correctly labelled
003 党参		Codonopsis pilosula	Codonopsis clematidea	No	Dangshen/ Codonopsis pilosula	Correctly labelled
004 白参		Panax ginseng	Panax ginseng	Yes	Baishen/ Panax ginseng	Correctly labelled
006 太子参		Pseudostellaria heterantha	Pseudostellaria heterophylla	No	Taizishen/ Pseudostellaria heterophylla	Correctly labelled
010 西洋参		Panax quinquefolius	Panax quinquefolius	Yes	Xiyangshen/ Panax quinquefolius	Correctly labelled
011 苦参		Sophora flavescens	Sophora flavescens	Yes	Kushen/ Sophora flavescens	Correctly labelled
013 白参		Panax ginseng	Panax ginseng	Yes	Baishen/ Panax ginseng	Correctly labelled
016 北沙参		N/A	Glehnia littoralis	N/A	Beishashen/ Glehnia littoralis	Correctly labelled
019 苦参		Sophora flavescens	Sophora flavescens	Yes	Kushen/ Sophora flavescens	Correctly labelled
020 玄参		Scrophularia ningpoensis	Scrophularia buergeriana	No	Xuanshen/ Scrophularia ningpoensis	Correctly labelled
022 桔梗		N/A	Platycodon grandiflorus	N/A	Jiegeng/ Platycodon grandiflorus	Correctly labelled
023 土参		Commenlina benghalensis	N/A	N/A	Tushen/ Talinum paniculatum	Mislabelled

Figure 1. Chart showing the BLAST results and market names of the samples and whether they are mislabelled

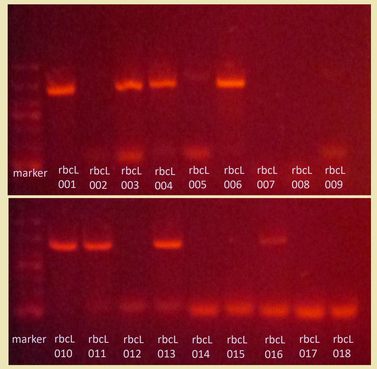

Figure2. Gel results of PCR products of sample 001 to 018 using rbcL primer. The bands of sample 005 and 016 PCR product was very faint so they were not sent for sequencing

Criteria

In order to determine which samples are labelled correctly, we set our own criteria:
if one or both of BLAST results match the market name, then the market name is labelled correctly
if the sample is only successfully amplified by one primer and the BLAST result mismatches the market name, then the sample is mislabelled
if both BLAST results mismatch the market name, then the sample is mislabelled

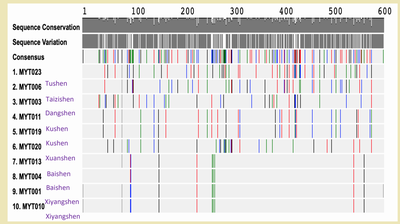

Figure 3. MUSCLE alignments of sequences of rbcL region

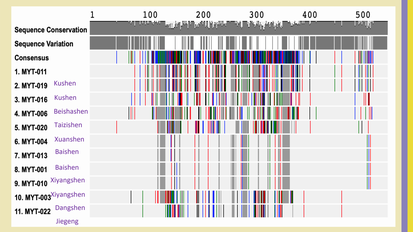

Figure 4. MUSCLE alignments of sequence of ITS2 region

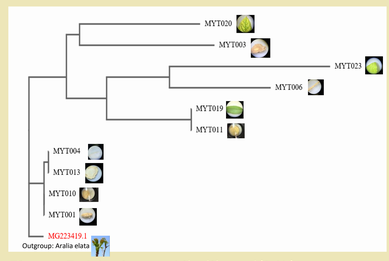

Figure 5. Maximum Likelihood(ML) phylogenetic tree of rbcL region (Outgroup: Aralia elata)

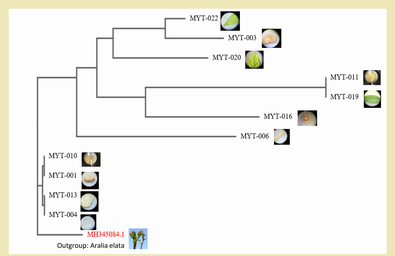

Figure 6. Maximum Likelihood(ML) Phylogenetic tree of ITS2 region (Outgroup: Aralia elata)

Discussion

- We did not use 75% ethanol to clean the samples or use only the inner part due to the samples being dry, tough and brittle. This may affect the efficiency of DNA extraction.
- We established our criteria so that we can find out whether mislabeling exists in our samples and meanwhile determine if ITS or rbcL primer is more effective in identifying ginseng species. However, further experiments are still required to improve the criteria.
- According to our criteria, apart from sample 23, all the other samples were not mislabelled, which suggests that the three pharmacies we went to are reliable in selling these medicines
- The rbcL BLAST result for sample 23 contradicts with the name labelled by the herb garden. However, we cannot conclude that the plant was mislabelled since there were two types of plants in the same pot with that label. We think we might collect the leaf from the wrong plant, so collection and experiments should both be carried out again to confirm our hypothesis.
- We cannot decide whether ITS2 or rbcL primer is more effective, because the BLAST results of both primers show approximately the same number of contradictions with the market names; yet, the ITS2 method seems to identify the samples to the species level more accurately. Sample 001 and 010 are different species from sample 004 and 013, but they are grouped as the same species in the rbcL phylogenetic tree, while their relationship is correctly demonstrated in the ITS tree. According to our results and previous studies, we also think it is better to combine two or more primers, though more samples and further experiments are needed to confirm these two assumptions.
- The bars in the rbcL MUSCLE alignment is more evenly distributed, while the ones in the ITS2 MUSCLE alignment are clustered. It is possibly because that the rbcL region is a coding region, which is unlikely to have substitutions and indels. The ITS region is a non-coding region, so it may have frequent mutations. [6]

Acknowledgements

We want to thank all the staff members in Cold Spring Harbor Asia, DNA Learning Center for assistance with providing the sorts of facilities for us to continue the experiment, and Dr. Pepenella for comments that greatly improved the manuscript. We would also like to show our gratitude to the Genewiz for providing the DNA sequencing service.

References

[1]陈士林,庞晓慧,姚辉, 等.中药DNA条形码鉴定体系及研究方向[J].世界科学技术-中医药现代化,2011,13(5):747-754. DOI:10.3969/j.issn.1674-3849.2011.05.001.
[2]孙涛, 滕少娜, 孔德英, et al. DNA条形码技术应用于人参鉴定[J]. 中国生物工程杂志, 2013, 33(4):143-148.
[3]Han, J., Pang, X., Liao, B., Yao, H., Song, J., & Chen, S. (2016). An authenticity survey of herbal medicines from markets in China using DNA barcoding. Scientific reports.
[4]Zuo, Y., Chen, Z., Kondo, K., Funamoto, T., Wen, J., & Zhou, S. (2011). DNA barcoding of Panax species. Planta medica, 77 2, 182-7.
[5] 任保青, 陈之端. 植物DNA条形码技术[J]. 植物学报, 2010, 45(1):1-12. DOI:10.3969/j.issn.1674-3466.2010.01.001
[6]Kim, W. J., Ji, Y., Choi, G., Kang, Y. M., Yang, S., & Moon, B. C. (2016). Molecular identification and phylogenetic analysis of important medicinal plant species in genus Paeonia based on rDNA-ITS, matK, and rbcL DNA barcode sequences. Genet. Mol. Res, 15(10.4238).
[7]罗志勇, 周钢, 陈湘晖, et al. 高质量植物基因组DNA的分离[J]. 湖南医科大学学报, 2001, 26(2):178-180.
[8]]https://www.dnabarcoding101.org