Scientific Research Publishing

American Journal of Plant Sciences, 2022, 13, 473-480
https://www.scirp.org/journal/ajps
ISSN Online: 2158-2750
ISSN Print: 2158-2742

A Reevaluation of the Status of *Lilium nepalense* Based on Complete Chloroplast Genome Sequences

Genshen Yin[1*], Shilong Wang[2], Shuangshuang Zhang[1], Wenlei Cheng[1], Minghua Dong[1]

[1]School of Agriculture & Life Sciences, Kunming University, Kunming, China
[2]School of Life Sciences, Zhengzhou University, Zhengzhou, China
Email: *yingenshen@126.com

How to cite this paper: Yin, G.S., Wang, S.L., Zhang, S.S., Cheng, W.L. and Dong, M.H. (2022) A Reevaluation of the Status of *Lilium nepalense* Based on Complete Chloroplast Genome Sequences. *American Journal of Plant Sciences*, **13**, 473-480.
https://doi.org/10.4236/ajps.2022.134030

Received: March 3, 2022
Accepted: April 15, 2022
Published: April 18, 2022

Abstract

Lilium nepalense D. Don is an ornamental plant and exhibits high morphological variation across its distribution range in China. There is incongruence between different data sets for the status of it in *Lilium*. In this study, we reported a complete chloroplast genome of *L. nepalense* sampled in Hengduan Mountains, China. The whole chloroplast genome possessed a total length of 152,206 bp with typical circular structure, and contained a large single copy (LSC) of 81,854 bp and a small single copy (SSC) of 17,563 bp, which were separated by a pair of inverted repeats (IRa/IRb) of 26,399 bp. The average GC content among the whole chloroplast genome sequence was 37%, and the GC content in LSC, SSC, IRs regions were 34.8%, 30.6%, and 42.5%, respectively. There were 135 genes detected from the whole chloroplast genome sequence, including 89 protein-coding genes, 38 tRNAs, and 8 rRNAs. Phylogenetic results using maximum likelihood and Bayesian methods showed congruent results that *L. nepalense* together with the other two *L. nepalense* samples collected from different habitats formed a single branch, indicating a close relationship of *L. nepalense* with *L. taliense* belonging to the section *Sinomartagon*. This newly characterized chloroplast genome will provide essential data for the further population genetics research of *L. nepalense*.

Keywords

Lilium nepalense, Chloroplast Genome, Phylogenetic Status, Habitats

1. Introduction

The *Lilium* L., belonging to Liliaceae, includes approximately 120 species, with the central distribution area in Hengduan Mountains and the Himalayas, many

species of *Lilium* are cultivated as ornamental [1] or medicinal plants [2]. In 1949, Comber used the classic taxonomic methods to establish the most widely accepted classification of this genus [3]. However, in the genus *Lilium*, it is difficult to determine the phylogenetic position of some species based on only distinct morphological characters, especially those having great morphological variations across the distribution areas, because some characters, especially flower shape, are possessed jointly by distantly related species. For example, morphologically, *L. nepalense* D. Don could be placed in the section *Sinomartagon* for its revolute tepals, but the characters of its slightly funnelform were somewhat similar to those of the plants within the section *Liriotypus*. Moreover, the results of phylogenetic analyses using rDNA ITS sequences showed that *L. nepalense* was included in section *Sinomartagon* [4]. Other results using the ITS and *psb*A-*trn*H markers showed that *L. nepalense* was nested within *Nomocharis* Clade [5] or with *Nomocharis* formed a single branch and sister to section *Liriotypus*, but distant from section *Sinomartagon* in phylogenetic tree constructed by ITS and *trn*A-*pet*H region [6]. There are two main reasons for these inconsistent results: firstly, *L. nepalense* inhabits a wide distribution range accompanied by heterogeneous ecological environments [7], perhaps resulting in high levels of genetic diversity in natural populations; secondly, it is likely that limited data sampling (e.g., partial chloroplast or nuclear gene fragments) were often insufficient to provide high resolutions for resolving the complicated phylogenetic position of *L. nepalense* within *Lilium*. Therefore, the phylogenetic position of *L. nepalense* varies according to different datasets or methodologies.

The chloroplast genome owns a conservative genome structure [8], but there are significant differences in length, and gene sequence between species, which has great advantages in phylogeny, and species delimitation in closely related species [9]. In recent years, with the rapid development of next-generation sequencing technology, an increasing number of complete chloroplast genomes were reported, and provided a promising method for the study of phylogeny and population genomics [10].

In this work, the chloroplast genome of *L. nepalense* collected from the wet forest in Lushui County, Yunnan Province in Hengduan Mountains was newly sequenced. Combining our data with previously published chloroplast genome data, the phylogenetic tree was constructed to reevaluate the status of *L. nepalense*. Species used to build the phylogenetic tree include 2 samples of *L. nepalense* (one from the dry valley in Yuanyang County of Southeast Yunnan [11], and one from an unknown location), 2 species of section *Sinomartagon*, 3 species of section *Liriotypus* and 3 species of *Nomocharis* Clade, and *Fritillaria cirrhosa* D. Don was used as the outgroup.

2. Materials and Methods

2.1. Plant Material, DNA Extraction, and Sequencing

Samples of *L. nepalense* were collected from Lushui county in Yunan, China (27˚01'42.85"N, 100˚11'49.87"E) (**Table 1**). Specimens of they were kept in Her-

barium of Kunming University under the collection No. YGS1907020. Genomic DNA was isolated using Super Plant Genomic DNA Kit (DP360) from 50 mg silica gel dried leaves by a modified method and DNA sample quality was examined with agarose-gel electrophoresis, and the concentration was measured by ultra-micro nucleic acid analyzer. The qualified samples were sequenced by the IIinuina Hiseq X platform from Novogene Inc.

2.2. Assembly and Annotation of the Chloroplast Genome

GetOrganelle [12] and CLC Genomics Workbench v8.0 were used to assemble the sequence from raw reads. Gene annotations were carried out by using the web program GeSeq [13] in MPI-MP CHLOROBOX (https://chlorobox.mpimp-golm.mpg.de/index.html) with BLAST search for protein coding genes through the references by cultivated species *L. nepalense* (GenBank no.: MW136391), tRNAscan-SE server for tRNAs [14], HEMMER for rRNAs [15]. The physical structure of plastomes was drawn by OGDRAW version 1.3.1 [16].

2.3. Phylogenetic Tree Constuction

In addition to two published chloroplast genome data of *L. nepalense* (MK493301, MW136391), genomes of 9 Sequenced species in section *Sinomartagon* [*L. taliense* Franchet (KY009938) and *L. primulinum* var. *ochraceum* (Franchet) Stearn (KY748298)], section *Liriotypus* [*L. bakerianum* Collett & Hemsley (KY748301), *L. amoenum* E.H. Wilson ex Sealy (MT880912) and *L. souliei* (Franchet) Sealy (MW085076)], and *Nomocharis* Clade [*L. gongshanense* (Y.D. Gao & X.J. He) Y.D. Gao (MK493297), *L. meleagrinum* (Franchet) Y.D. Gao (MK493299) and *L. pardalinum* (Franchet) Y.D. Gao (MH029495)] in *Lilium* were also downloaded from NCBI used to build phylogenetic tree, *Fritillaria cirrhosa* (MH593346) was outgroup.

Table 1. Comparison of chloroplast genome among two datas of the *L. nepalense* D. Don.

Characteristics	MW136391	MW853784
Source	Wu *et al.* 2021	this study
sampling place	Yuanyang, Yunan	Lushui, Yunan
Habitat	dry valley	wet forest
Total size (bp)	152,956	152,206
LSC size (bp)	82,573	81,854
SSC size (bp)	17,527	17,563
IR size (bp)	26,428	26,399
Number of genes	131	135
GC content (%)	37	37
rRNA gene number	8	8
tRNA gene number	38	38
Protein-coding gene	85	89

Maximum likelihood (ML) estimation with 1000 bootstrap replications was conducted using IQ-TREE 2 [17]. For Bayesian Inference (BI) analysis, MrBayes 3.2.6 [18] was used to BI tree.

3. Results and Discussion

3.1. Genome Content and Organization

The chloroplast genome of *L. nepalense* was a circular molecule of double-stranded DNA and possessed a total length of 152,206 bp (**Figure 1**, **Table 1**). The whole genome sequence contained a large single copy (LSC) of 81,854 bp

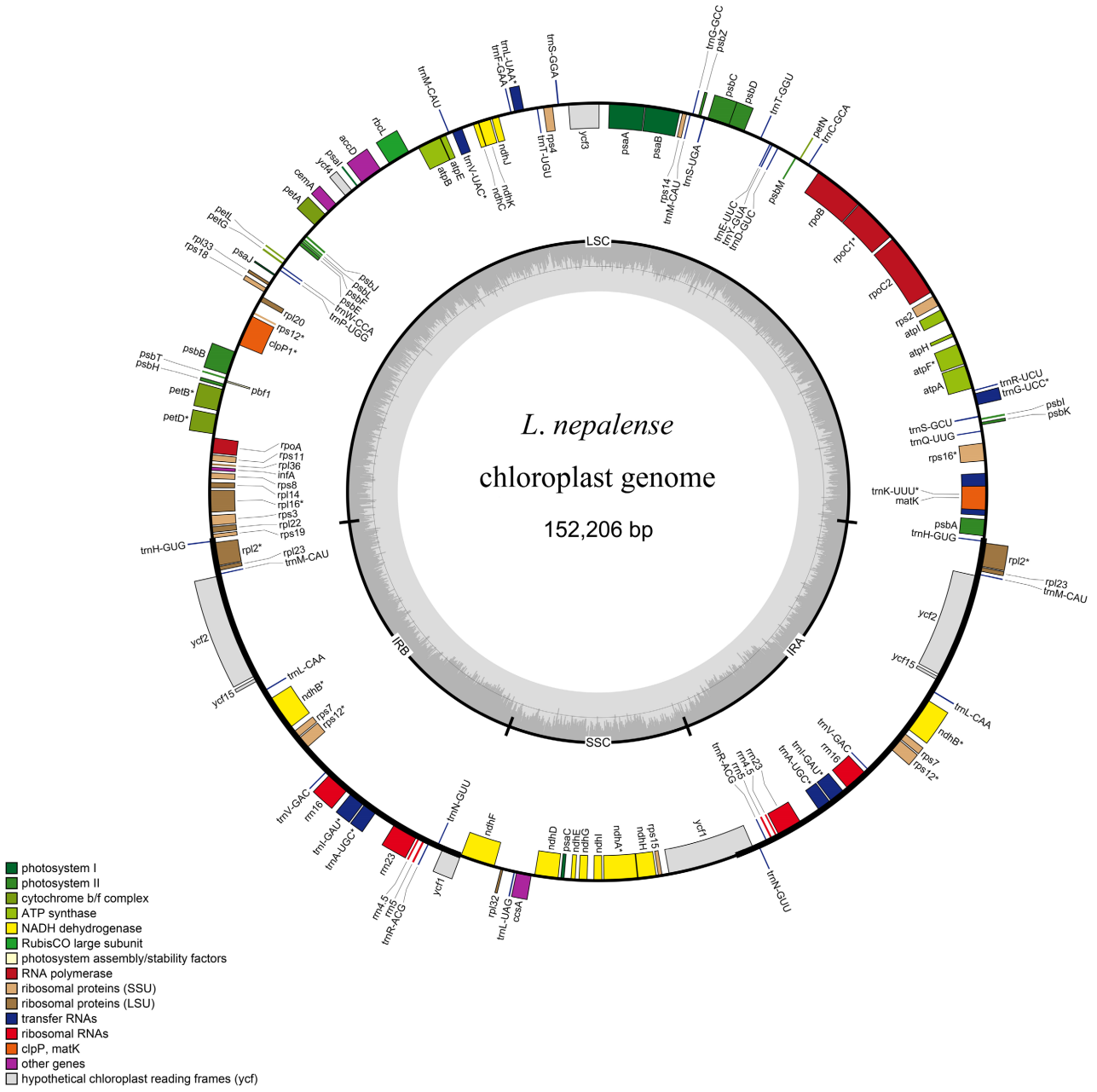

Figure 1. The circular map of the cp genome of *L. nepalense* D. Don.

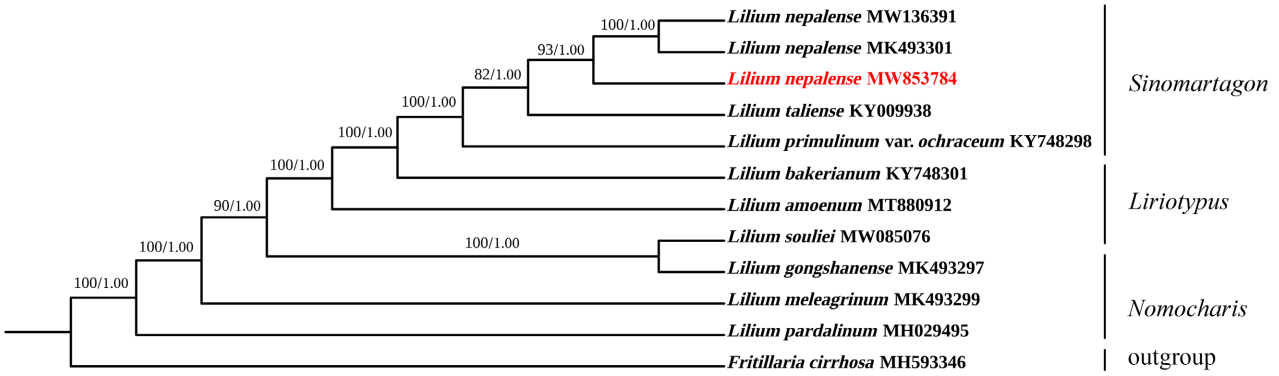

Figure 2. Phylogenetic tree of *L. nepalense* D. Don with other species in section *Sinomartagon*, section *Liriotypus* and *Nomocharis* Clade in the genus *Lilium* constructed on the basis of the whole chloroplast genome. The bootstrap percentage values and Bayesian posterior probability were listed at each node.

and a small single copy (SSC) of 17,563 bp, which were separated by a pair of inverted repeats (IRa/IRb) of 26,399 bp (Table 1). The average GC content of whole chloroplast genome sequence was 37%, and the GC content in LSC, SSC, IRs regions were 34.8%, 30.6%, and 42.5%, respectively (Table 1). There were 135 genes found from *L. nepalense* plastid genome sequence, which included 89 protein-coding genes, 38 tRNAs, and 8 rRNAs. Among those genes, 20 duplicate genes (8 protein coding genes, 8 tRNA genes, 4 rRNA genes) and 23 introns contained genes were also detected (Figure 1, Table 1). The complete chloroplast genome of *L. nepalense* was submitted to the genomics into the GenBank with accession number of MW853784. The data obtained in this study are slightly different from those of Wu *et al.* [11], except that GC content, rRNA gene number and tRNA gene (Table 1).

3.2. Phylogenetic Analysis

The phylogenetic relationships among 10 species, along with sequenced cp genomes were further investigated using IQ-TREE and MrBayes through the likelihood and Bayesian methods with a bootstrap of 1000 replicates to assess the reliability. The phylogenetic results showed that *L. nepalense* sampled from different habitats (dry valley and wet forest, respectively) was more closely related to each other, this suggests that genetic variations between different populations of the same species were smaller than in different species. Besides, all of the three *L. nepalense* samples showed more close relationship with *L. taliense*, belonging to the section *Sinomartagon*, than to any other *Lilium* species with a strong bootstrap value (Figure 2), indicating that our results were in accordance with that of traditional taxonomy (Comber, 1949). Which revealed that *L. nepalense* should be included in section *Sinomartagon* but then section *Liriotypus* (Figure 2).

4. Conclusions

Chloroplasts are the sites of photosynthesis in plants and the cp genome usually be used for systematic analysis of their highly conservative genome structure but

significantly different gene sequence. We studied the characteristics of the cp genome of *L. nepalense*, an economic plant with great value as horticultural crops and medicine. Our tree result indicated that *L. nepalense* should be placed in the section *Sinomartagon* but the section *Liriotypus* with highly bootstrap value. This study provides a basis for solving a long-standing puzzle for the status of *L. nepalense* in *Lilium*.

Habitat type may play an important role in the genetic divergence of species, which has been proved by studies on *Phragmites australis [19]*, *Blanus* sp. [20], *Dendrobates pumilio* [21] and so on. *L. nepalense* presents an excellent opportunity for this kind of research. Unlike other *Lilium* species, *L. nepalense* has a wide range and is distributed across different habitats. Our study showed that there were some differences in the genetic structure between different habitats (dry valley and wet forest, respectively).

This characterized cp genome will also provide essential data for the development of molecular markers and the further population genetics research of *L. nepalense*.

Data Availability Statement

The genome sequence data that support the findings of this study are openly available in GenBank under the accession no. MW853784 (https://www.ncbi.nlm.nih.gov/nuccore/MW853784.1/).

Authors' Contribution

Genshen Yin conceived the idea and wrote the draft of the paper. Shilong Wang performed the data analyses. Shuangshuang Zhang and Wenlei Cheng performed the experiment. Minghua Dong helped perform the analysis with constructive discussions. All the authors contributed to the final version of the manuscript.

Acknowledgements

We thank Mingzhong Mo (Forestry and Grassland Bureau of Honghe Hani and Yi Autonomous Prefecture, Yunnan, China) for his assistance with field sampling. This work was supported by the National Natural Science Foundation of China [Grant No.31860107] and the Basic Research Joint Special Fund project of Local undergraduate universities in Yunnan Province [Grant No.202001BA070001-163].

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Woodcock, H.B.D. and Stearn, W.T. (1950) Lilies of the World: Their Cultivation and Classification. Country Life Limited, London.

[2] Commission, C.P. (2020) Pharmacopoeia of People's Republic of China. China Medical Science Press, Beijing, 138.

[3] Comber, H.F. (1949) A New Classification of the Genus *Lilium*. In: Chittenden, F.J., ed., *Royal Horticultural Society Lily Year Book*, RHS, London, 85-105.

[4] Nishikawa, T., Okazaki, K., Arakawa, K. and Nagamine, T. (2001) Phylogenetic Analysis of Section *Sinomartagon* in Genus *Lilium* Using Sequences of the Internal Transcribed Spacer Region in Nuclear Ribosomal DNA. *Breeding Science*, **51**, 39-46. https://doi.org/10.1270/jsbbs.51.39

[5] Gao, Y.D., Hohenegger, M., Harris, A., Zhou, S.D., He, X.J. and Wan, J. (2011) A New Species in the Genus *Nomocharis* Franchet (Liliaceae): Evidence that Brings the Genus *Nomocharis* into *Lilium*. *Plant Systematics and Evolution*, **298**, 69-85. https://doi.org/10.1007/s00606-011-0524-1

[6] Ikinci, N. (2011) Molecular Phylogeny and Divergence Times Estimates of *Lilium* Section *Liriotypus* (Liliaceae) Based on Plastid and Nuclear Ribosomal ITS DNA Sequence Data. *Turkish Journal of Botany*, **35**, 319-330.

[7] Chen, X.Q., Liang, S.Y., Xu, J.M. and Tamura, M.N. (2000) Flora of China (Vol. 24). Science Press, Beijing, 73-263.

[8] Sugiura, M. (1992) The Chloroplast Genome. *Plant Molecular Biology*, **19**, 149-168. https://doi.org/10.1007/BF00015612

[9] Chen, X.L., Zhou, J.G., Cui, Y.X., Wang, Y., Duan, B.Z. and Yao, H. (2018) Identification of *Ligularia* Herbs Using the Complete Chloroplast Genome as a Super-Barcode. *Frontiers in Pharmacology*, **9**, Article No. 695. https://doi.org/10.3389/fphar.2018.00695

[10] Kim, H.T., Lim, K.B. and Kim, J.S. (2019) New Insights on *Lilium* Phylogeny Based on a Comparative Phylogenomic Study Using Complete Plastome Sequences. *Plants*, **8**, Article No. 547. https://doi.org/10.3390/plants8120547

[11] Wu, H.Z., Bai, W.W., He, S.L., Zhao, Y. and Wu, J.Z. (2021) The Complete Chloroplast Genome of *Lilium nepalense* (Liliaceae). *Mitochondrial DNA Part B*, **6**, 526-527. https://doi.org/10.1080/23802359.2021.1872451

[12] Jin, J.J., Yu, W.B., Yang, J.B., Song, Y., dePamphilis, C.W., Yi, T.S. and Li, D.Z. (2020) GetOrganelle: A Fast and Versatile Toolkit for Accurate *De Novo* Assembly of Organelle Genomes. *Genome Biology*, **21**, Article No. 241. https://doi.org/10.1186/s13059-020-02154-5

[13] Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E.S., Fischer, A., Bock, R. and Greiner, S. (2017) GeSeq—Versatile and Accurate Annotation of Organelle Genomes. *Nucleic Acids Research*, **45**, W6-W11. https://doi.org/10.1093/nar/gkx391

[14] Lowe, T.M. and Eddy, S.R. (1997) tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. *Nucleic Acids Research*, **25**, 955-964. https://doi.org/10.1093/nar/25.5.955

[15] Wheeler, T.J. and Eddy, S.R. (2013) nhmmer: DNA Homology Search with Profile HMMs. *Bioinformatics*, **29**, 2487-2489. https://doi.org/10.1093/bioinformatics/btt403

[16] Greiner, S., Lehwark, P. and Bock, R. (2019) OrganellarGenomeDRAW (OGDRAW) Version 1.3.1: Expanded Toolkit for the Graphical Visualization of Organellar Genomes. *Nucleic Acids Research*, **47**, W59-W64. https://doi.org/10.1093/nar/gkz238

[17] Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. and Lanfear, R. (2020) IQ-TREE 2: New Models and Efficient Methods

for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, **37**, 1530-1534. https://doi.org/10.1093/molbev/msaa015

[18] Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. and Huelsenbeck, J.P. (2012) MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. *Systematic Biology*, **61**, 539-542. https://doi.org/10.1093/sysbio/sys029

[19] Qiu, T., Liu, Z.Y., Yang, Y.F. and Liu, B. (2021) Epigenetic Variation Associated with Responses to Different Habitats in the Context of Genetic Divergence in *Phragmites australis*. *Ecology and Evolution*, **11**, 11874-11889. https://doi.org/10.1002/ece3.7954

[20] Şahin, M.K., Candan, K., Yildirim Caynak, E., Kumlutaş, Y. and Ilgaz, Ç. (2021) Ecological Niche Divergence Contributes Species Differentiation in Worm Lizards (*Blanus* sp.) (Squamata: Amphisbaenia: Blanidae) in Mediterranean part of Anatolian Peninsula and the Levantine Region. *Biologia*, **76**, 525-532. https://doi.org/10.2478/s11756-020-00548-1

[21] Wang, I.J. and Summers, K. (2010) Genetic Structure Is Correlated with Phenotypic Divergence Rather than Geographic Isolation in the Highly Polymorphic Strawberry Poison-Dart Frog. *Molecular Ecology*, **19**, 447-458. https://doi.org/10.1111/j.1365-294X.2009.04465.x