

Phylogenetic relationships of *Prunus* (Rosaceae) distributed in North East India: insights from the ribosomal internal transcribed spacer sequence (ITS)

Jennifer N. Mekrini

Institute of Bioresources and Sustainable Development

Biseshwori Thongam (✉ b_thongam07@yahoo.com)

Institute of Bioresources and Sustainable Development <https://orcid.org/0000-0003-2675-8971>

Research Article

Keywords: Internal transcribed spacers, phylogeny, *Prunus*, ML, Rosaceae, Bayesian

Posted Date: February 7th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-2519946/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Prunus L. is a large genus including trees and shrubs such as apricot, almond, peaches, plum and nectarines belonging to family Rosaceae. The phylogenetic analysis using ITS sequence was conducted to understand the evolutionary relationship between the *Prunus* sp. of North East India.

Methods and results

The present study focuses on the 32 accessions, comprising of six voucher specimen of *Prunus cerasoides*, one voucher specimen of *Prunus campanulata*, one voucher specimen of *Prunus napaulensis*, one voucher specimen of *Prunus domestica* and other twenty two voucher specimen of *Prunus* comprising 5 subgenera according to Rehder (1940) and one voucher specimen of *Pyrus communis* (outgroup) archived from GenBank. The molecular marker (ITS) was utilized in the present study to understand the phylogenetic relationship. Various molecular markers are utilized in assessing the phylogenetic relationship and genetic diversity of *Prunus*. The reconstruction using the ITS region is to evaluate the evolutionary relationship of our species collected from North East region and to reflect on the relatedness of the taxa based on the ML and Bayesian method. The result of phylogenetic analyses based on ITS sequence clustered two major groups comprising of the *Cerasus-Padus-Laurocerasus* complex and the *Prunus-Amygdalus* complex.

Conclusion

The ITS sequences of the *Prunus* species from the present investigation would be helpful in the future study as it provides access to the sequences from the region of North East India as there is no phylogenetic study conducted till date.

Introduction

Prunus L. is a genus in Rosaceae comprising of trees and shrubs such as almond, apricot, cherry, peaches, plum and nectarines [1]. Morphologically Rosaceae have pentamerous flowers with valvate sepals, five conspicuous clawed petals and numerous stamens inserted on a receptacular cup [2]. About 430 species of *Prunus* accepted worldwide and are known to be mainly distributed in North Temperate region. In India, a total of 36 species were represented [3–4]. Among 36 taxa, 18 taxa are significantly used for cultivation and commercial purposes in terms of its edible fruits and kernels and another 18 taxa for its wild and economic importance. It is commercially cultivated in Northern hilly states of India like Himachal Pradesh, Jammu and Kashmir and Uttaranchal. [5–6]. A total of 19 *Prunus* are available widely in the Himalayan region. In Northeast India, around 13 species of *Prunus* are recognized [7]. *P.napaulensis*, *P.candulata* and *P.cerasoides* are reported in Khasi Hills Meghalaya [8].

Prunus is characterized by having simple, alternate, usually lanceolate, in *Amygdalus mira*, *Prunus simonii*, *Cerasus cerasoides*, *Cerasus glandulosa* [9]. The flowers are usually white to pink, sometimes red, with five petals, five sepals and diverse stamens. The flowers appear singly in *Prunus domestica*, *Prunus spinosa*, *Amygdalus dasycarpa*, in umbels of two to 6 in *Cerasus campanulata*, *Cerasus conradine*, *Cerasus pseudocerasus* and sometimes racemes in *Prunus napaulensis*, *P.grayana* and *P.obtusata* [9]. The fruits of the genus *Prunus* are easy to recognise with fleshy relatively hard-coated seed, known as the drupe [10]. *Prunus* are known to have adapted low winter temperatures and summer drought [11]. Previous molecular studies were attempted using different markers like ITS, rbcL, matK. Some studies attempted to construct the evolutionary relationships of the genus using this marker [12–13]. The

utilization of nucleotide sequences of RNA, DNA and amino acids of a protein should be incorporated for molecular phylogenetic study and achieved using new techniques [14]. The ITS sequences have high evolution rate and are useful for phylogenetic analysis among related species or among populations within a species. The ITS marker was proposed as a barcode for flowering plants [15]. The internal transcribed spacer (ITS) region of ribosomal DNA (rDNA), contain the ITS1 spacer, 5.8S rRNA gene, and ITS2 spacer, has been a useful molecule marker in determining taxonomic relationships of many plant families. [16]. ITS region is known for its uses in phylogenetic studies of the many taxa of flowering plants, including groups within the family Rosaceae [16–17].

In NE India, the early studies focuses mostly on the species distribution and chemical constituent of *Prunus*, therefore detail studies of the genus is required since it will be helpful in identification, resolve its phylogenetic relationship and helps in conservation. Since morphological data can be limited in data in some species, molecular data can be used to construct the phylogenetic tree to obtain a more clear and robust phylogenetic tree [14]. Hence preliminary investigation was undertaken regarding the extent of biodiversity of some *Prunus* species found in North East India to infer the phylogenetic relationship and establish relatedness among the taxa. To gain insights on the *Prunus* of NE India, this study was designed:

1. To construct phylogenetic reconstruction by employing sequences of ITS of ribosomal nuclear region to help resolve phylogenetic relationship of *Prunus* with an emphasis on North East India.
2. To understand the relationship between the *Prunus* species collected from North East India and species already deposited in GenBank reflecting relatedness among the taxa and the evolutionary relationship of the genus.

Materials And Methods

Sample collection and Identification

Extensive botanical explorations in different districts of Manipur namely Senapati, Ukhrul, Tamenglong, Bishnupur and Chandel district in the year 2019–2021 were carried out during the flowering and fruiting season (from October–December) for collection of *Prunus* germplasm. Several species were also collected from NE Indian states like Nagaland, Meghalaya, Arunachal Pradesh, Sikkim and extended to other parts of India namely Tamil Nadu. Most of the samples were obtained from wild populations. Samples were then stored at -80°C until the extraction of DNA. Living plant material especially seeds of *Prunus* species were collected from the natural habitat and was grown in pots in the greenhouse at Institute of Bioresources and Sustainable Development (IBSD) Talyelpat, Manipur. The voucher specimen were deposited in the herbaria of Institute of Bioresources and Sustainable Development (IBSD) for identification of species using conventional phenotypic method and the herbarium specimen were consulted at Botanical Survey of India, Kolkata. The GPS co-ordinates of the sample collected from different parts of North East India are depicted in table 1.

DNA Isolation, Amplification and Sequencing

The data consist of nine accessions, representing six voucher specimens of *Prunus cerasoides* collected from 6 different locations, one of *Prunus campanulata*, one of *Prunus domestica* and one *Prunus napaulensis*. The species were collected based on the different locations. Isolation of total genomic DNA was carried out using Qiagen DNA isolation Mini kit and the extracted DNA were quantified using Nanodrop ND-1000 spectrophotometer (Thermoscientific) for purity of the sample. Universal primers ITS1 (5'TCCGTAGGTGAACCTGCGG-3') and ITS4 (5'TCCTCCGCTTATTGATATGC-3') respectively designed were used for PCR amplification [18]. PCR was carried out in a Thermal cycler (Biorad). Each PCR was performed in a total volume of 20 µl containing 1µl of genomic DNA, 2µl of

10x PCR buffer, 0.25 µl of 1 mM dNTPs, 1 µl each of 5 µM ITS 1 and ITS 2, 0.2 µl of 5 units/µl Taq polymerase (Sigma-Aldrich) and 14.55 µl of Nuclease Free Water (Qiagen). The PCR process consisted of an initial denaturation at (94°C) for 5 min, followed by denaturation at (95°C) for 30 sec, annealing at (58°C) for 1 min, and extension at (72°C) for 2 min for a total of 35 cycles and a final extension at (72°C) for 7 min. The samples were held at 4°C. Amplified PCR products were separated via electrophoresis on 1.2% (W/V) agarose gels with 1x TBE buffer at 60V for 65 min, stained with GelRed™ Nucleic Acid Stain and visualized before cleaning under UV light using UV transilluminator (Bio-Rad) and Bio-Rad Molecular Imager GelDoc™ Model-Universal Hood (Bio-Rad Lab Inc. USA). Gene ruler 1 kb DNA Ladder (Promega) was loaded to identify the size of the amplicons. The amplified products were purified using the Qiagen gel/PCR purification kit according to the instruction (Valencia California USA) and the purified products were then sent for Sanger DNA sequencing to Bionivid Technology Private Limited Bangalore. The same primers used for amplification were also used in the process of sequencing.

Nucleotide Sequence Data

The nucleotide sequence of the ITS region for all the nine individuals have been submitted to GenBank. The nucleotide sequence data obtained in this study (including downloaded and newly sequenced) appear in the NCBI (National Center for Biotechnology information) Genbank and the accessions number are listed in Table 2.

Sequence Analysis

Basic Local Alignment Search Tool (BLAST) was used to find the most homological sequences in the GenBank databases. Determination of the conserved sites (CS), variable sites (VS), Parsimony informative sites (PIS), Singleton sites (SS) and transition and transversion ratio and transition and transversion pair were generated by performing multiple sequence alignment using Mega X [19]. The sequences were aligned with other *Prunus* sequences deposited in GenBank whose phylogenetic relationships were previously reported [20]. The sequences were then visually inspected to verify alignments; after discarding unaligned sequences, the sequences are realigned in Mega X using MUSCLE algorithm to verify the alignment along with the other amplicon.

Construction of phylogenetic tree

The phylogenetic tree was re-rooted using ITS Sequence of *Pyrus communis* (NCBI accession No JQ392467.1) as outgroup. The nuclear ITS sequences were submitted to NCBI GenBank. Best fit substitution models were calculated and selected based on lowest BIC scores (Bayesian Information Criterion) using MEGA software and the best fit model of the ITS nucleotide sequence was found to be K2 + G. The heuristic search for the maximum likelihood analysis was performed with the subsequent options: Nearest-Neighbour-Interchange (NNI) [21]. The dataset for phylogenetic analysis includes 31 ITS sequences representing five subgenus of the *Prunus*. Among these, 22 ITS sequence were retrieved from GenBank and 9 were newly sequenced. The phylogenetic position was analyzed by maximum likelihood and Bayesian analyses of the nuclear ITS DNA sequences. The maximum likelihood tree is presented in Fig. 1 with the posterior probability (PP) indices of Bayesian analysis and also the bootstrap values (BT) of the maximum likelihood method. The phylogenetic tree was constructed using Mega X software. Before phylogenetic analysis, sequence characteristics of all genome regions were calculated using MEGA X [19]. The aligned sequences were analyzed by maximum likelihood (ML) method using Mega X and Bayesian method using MrBayes 3.2.7 [22]. Markov chain Monte Carlo (MCMC) analysis was run for 1,000,000 generations with one cold and three heated chains, starting from random trees and sampling one out of every 200 generations, until the

average standard deviations of split frequencies fell below 0.01, indicating the convergence of the iterations. The first 25% of trees were discarded as burn-in and the remaining tree were used to calculate the posterior probabilities (PP).

Result

PCR amplification of ITS and Sequence analysis

PCR and sequencing with primer ITS1 and ITS 4 successfully amplified the nuclear ITS region from all the accessions of *Prunus* investigated in the study. The length of ITS1, 5.8S and ITS2 regions of the taxa study ranged from, 226–256 bases for ITS1, 232–305 bases for ITS2 and 5.8S was a constant 102 bases (Table 3). The sequences have been submitted in the GenBank under the accession number MT811814, MT811815, MT811816, MT811817, MT811818, MT811819, MT811820, MT811821 and MT828305. The mean average Guanine + Cytosine content of the ITS1, 5.8S and ITS2 was 59.24%, 57.8% and 57.81% respectively. The highest number of nucleotide for ITS sequence was observed in *P.cerasoides* Ukhul (633 bases) whereas the lowest was recorded in *P.cerasoides* Nagaland (566 bases). The maximum GC content (63.8%) and the lowest AT content (36.2%) was observed in *P.napaulensis*. Nevertheless, the lowest GC content (53.20%) and maximum AT content (46.8%) were recorded in *P.cerasoides* Arunachal Pradesh. Similar GC content was also observed in the sequence ITS region as reported by Lee & Wen [20] supporting the observed pattern of GC variation. Alignment of the entire regions required 10 indels, 4 in ITS1, 1 in 5.8s and 5 in ITS2. SNPs were present in the ITS region with 2 in ITS1 and 10 in ITS2. Further the Tajima neutrality test was used to estimate nucleotide diversity value (π) [23]. A total of 189 segregating sites (S) and 484 maximum number of positions (N) demonstrate a nucleotide diversity N (π) of 0.055852 was observed. The estimated Transition/ Transversion bias (R) is 1.67.

Molecular Phylogeny Analysis

We have analyzed the dataset of the ITS region of 32 taxa which was generated from the multiple alignment for both the maximum likelihood (ML) and Bayesian inference. The accession number JQ392467.1 *Pyrus communis* archived from the GenBank was selected as outgroup for the present investigation. Bootstrap values below 50% for branches were excluded. Our present study includes the five subgenera for phylogenetic inference provided by Rehder namely *Amygdalus*, *Prunus*, *Padus*, *Cerasus* and *Laurocerasus* [16]. In the ML phylogenetic tree inferred from the ITS sequences, two major groups can be recognized comprising of subgenera *Cerasus*-*Padus*-*Laurocerasus* complex and *Amygdalus*-*Prunus*-complex clustering as group. Our phylogenetic analysis showed that the subgenera *Cerasus* and *Prunus* each form a separate clade. In the ML tree, our samples were closely clustered with the other individuals of the subgenera *Cerasus*. In the ML tree the subgenus *Amygdalus*, *Laurocerasus* and *Padus* are found basal to the subgenus of *Cerasus*. The subgenus *Cerasus* was found to be nested above the subgenera clade *Amygdalus*-*Prunus* complex. In the ML tree, the clade comprising of *Prunus caroliniana*, *P.grayana* and *P.napaulensis* was found to be basal to subgenera *Prunus*. Most species of subgenera *Cerasus* form a group. The accession numbers of the species included in the alignment are presented in the tree.

Discussion

Our analyses shows that with greater range of primer universality, universal primers could amplify the ITS region from plant, fungus and animal DNA samples [24]. The nuclear ribosomal ITS region has been reported to display a high degree of divergence between other species but highly conserved within species due to which they are treated

the most preferred genetic marker for species level identification. ITS shows the highest discriminating power in terms of infrageneric levels [25]. Our present study investigates on the molecular data variation of 32 nucleotide sequence of the ITS region which represents 5 subgenera of *Prunus* and *Pyrus communis* as outgroup of the family Rosaceae. Based on our analysis the length of trimmed nucleotide sequence of the complete ITS region including ITS1, 5.8S and ITS 2 was found to be in the range of 566–633 bases. The previous studies reported that their complete ITS region of *Prunus* was in the range of 591–615 bases [20]. Thus showing that there is not much variation in the base length. Our present study investigates that subgenus *Cerasus* and *Prunus* each form a clade showing consistent with the findings from previous analysis [20]. The phylogenetic analyses strongly support that the subgenera *Laurocerasus* and *Padus* intermingled with each other and together form a paraphyletic clade. The two species sampled for the subgenus *Amygdalus* clade clustered into a group and this report is similar with the earlier studies [20]. Based on some study the subgenera *Amygdalus* was reported as monophyletic group, but the variation in our study might be limited number of taxa of the subgenera *Amygdalus* but the sampled included in the study clustered together [26]. The *Cerasus-Padus-Laurocerasus* complex in the maximum likelihood ITS tree was supported by bootstrap value of 69%. Within the *Cerasus-Padus-Laurocerasus* complex, taxa of subgen. *Cerasus* formed a large cluster including the seven individuals collected for the study. The subgenus *Cerasus* differs from the subgenera *Padus* and *Laurocerasus* morphologically by having axillary or umbels shape inflorescence and the later having racemose, many flowered. Previous findings also reports that *P.tomentosa* of the subgenera *Cerasus* is group within the taxa of subgenus *Prunus* and *Amygdalus*, our present study also supports that *P.tomentosa* of the subgenera *Cerasus* is nested within the subgenera *Amygdalus* and *Prunus*[24]. *P.tomentosa* has a solitary or 2 in fascicle inflorescence like taxa of subgenera *Prunus* and *Amygdalus*. Unlike other taxa of subgen. *Cerasus*, *P.tomentosa* and other species in section *Microcerasus* do have solitary or 2 in a fascicle inflorescence similar with those in the subgenera *Prunus* and *Amygdalus* which maybe the reason why they are intermixed with those subgenera. The subgenera *Cerasus* and the section *microcerasus* needs to be studied more for the phylogenetic taxa position by adding more species in future.

Based on our study, the separation of subgenera *Padus* and *Laurocerasus* is not supported. The ITS phylogenetic tree analysis clearly indicates that species of both subgenera are intermingled with each other. These two taxa share common character morphologically by having racemose inflorescence [9]. Previous studies in relevance to Rosaceae using molecular markers are reported in *Prunus* using ITS and Trnf [27]. Earlier studies also show support in utilization of marker like ITS, S6pdh and trnI-trnf spacer in phylogenetic study [26]. It concludes that ITS alone obtained result similar than with combined datasets, it is most likely because ITS had twice the number of informative sites. Hence the study discusses that molecular marker like ITS can be utilized for phylogenetic study in species level. Here we present a phylogeny reconstruction for 31 ITS sequence data of *Prunus* using maximum likelihood and Bayesian inference method. Our goal is to produce phylogenetic evidence for classification that reflects the evolutionary history and its relationship between the genus and subgenus of *Prunus* from North East India. The study is carried out by comparing with species archived within the GenBank as there is very limited study in the North East region in respect to molecular data.

Conclusion

Based on our analysis, it can be noted that ITS region can be used in the phylogenetic study. The study was carried out to find out the evolutionary relatedness of *Prunus* from North East region using ITS sequences along with those archived in the GenBank. Our study also found paraphyletic branches within the subgenera *Padus* and *Laurocerasus*. Thus the subgeneric study of the status of *Padus* and *Laurocerasus* is not fully resolved as both the species representing the genera intermingled with each other. Our study also correlates with the study by [20] [27]

that the subgenera *Prunus* and *Amygdalus* are closely related to each other, but more data needs to be added in the future to fully resolve the relationship between the *Prunus* species for classification of the genus to support in the phylogenetic inference. Our ITS tree is overwhelmingly congruent with the tree presented earlier by the [20] as it have the same placement with majority of the taxa. Nuclear region used in this study provide sufficient resolution for *Prunus* species of the North East region by forming a major cluster with its own subgenera *Cerasus*, *Prunus* and *Padus*. This is the first report of 9 ITS sequences generated for phylogenetic study from the North East region. The sequences were submitted for the first time in NCBI database. Thus, the study would provide valuable genetic information for further conservation and evolutionary research and the sequences is available as reference sequences in future. However, extensive sampling of this genus is highly recommended as modern DNA markers and other chloroplast markers can help in infrageneric study of *Prunus*.

Declarations

Acknowledgements

The authors wish to express gratitude to the Director, Institute of Bioresources and Sustainable Development (IBSD), Manipur, India for facilities and support. The authors also thank the Department of Biotechnology (DBT) Govt. of India, New Delhi for financial support. The authors thank the Botanical Survey of India; Kolkata for providing access in consulting the herbarium specimens. The authors also wish to acknowledge Kajio Joel, Loli Sani for their field support.

Author contributions BT and JNM planned and designed the study. JNM performed all the experiments. JNM did data analysis. JNM wrote the manuscript with input from BT; BT checked and improved the manuscript.

Funding The authors are very much grateful to Department of Biotechnology (DBT), Government of India, New Delhi for financial support to carry out this research work.

Data availability The dataset that was analysed during the current study is available in GenBank at <https://www.ncbi.nlm.nih.gov/>, with accession number MT811814, MT811815, MT811816, MT811817, MT811818, MT811819, MT811820, MT811821 and MT828305.

Competing interest The authors declare there is no conflict of interest.

Ethics approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Joseph, N., Anjum, N., & Tripathi, Y. C. (2018). *Prunus cerasoides* D. Don: A review on its ethnomedicinal uses, phytochemistry and pharmacology. *International Journal of Pharmaceutical Sciences Review and Research*, 48(1), 62–69.
2. Wang, J. R., Wang, X., Su, N., Li, Q. J., Zhang, X. H., Ma, Y. P., Zhao, L., Ginefra Toni, J. F., & De Craene, L. R. (2020). Floral morphology and morphogenesis in *Sanguisorba* (Rosaceae): Flower diversification despite petal reduction and spatial constraints. *Botanical Journal of the Linnean Society*, 193(1), 47–63. <https://doi.org/10.1093/botlinnean/boaa009>

3. Ghora, C., & Panigrahi, G. (1984). Fascicles of flora of India. Rosaceae: Genus Prunus. Botanical Survey of India, Calcutta, India, 18, 1–44.
4. Santapau, H., & Henry, A. N. (1973). Dictionary of the flowering plants in India.
5. Sharma, J. K. (2000a). Behmi-a wild fruit from Himalayan cold desert region. *Journal of American Pomological Society*, 54(1), 27–28.
6. Sharma, J. K. (2000b). Morphological studies on apricot and its wild relatives. *Journal of Hill Research*, 13(1), 5–10.
7. Pandey, A., Nayar, E. R., Venkateswaran, K., & Bhandari, D. C. (2008). Genetic resources of Prunus (Rosaceae) in India. *Genetic Resources and Crop Evolution*, 55(1), 91–104. <https://doi.org/10.1007/s10722-007-9217-6>
8. Rana, J. C., Dutta, M., & Rath, R. S. (2012). Plant genetic resources of the Indian Himalayan region - An overview. *Indian Journal of Genetics and Plant Breeding*, 72(2), 115–129.
9. Lingdi L., Cuizhi G., Chaoluan L., Alexander C., Bartholomew B., Brach A.R., Boufford D.E., Ikeda H., Ohba H. & Robertson K.R. 2003: Rosaceae. — *Flora of China* 9: 46-434.
10. Brian, M. (2003). The wealth of India, raw materials. *Council for Scientific and Industrial Research*, 7, 69–70.
11. Dirlewanger, E., Quero-García, J., Le Dantec, L., Lambert, P., Ruiz, D., Dondini, L., Illa, E., Quilot-Turion, B., Audergon, J. M., Tartarini, S., Letourmy, P., & Arús, P. (2012). Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three Prunus species: Peach, apricot and sweet cherry. *Heredity*, 109(5), 280–292. <https://doi.org/10.1038/hdy.2012.38>
12. Badenes, M. L., & Parfitt, D. E. (1995). Phylogenetic relationships of cultivated Prunus species from an analysis of chloroplast DNA variation. *Theoretical and Applied Genetics*, 90(7–8), 1035–1041. <https://doi.org/10.1007/BF00222918>
13. Mowrey, B. D., & Werner, D. J. (1990). Phylogenetic relationships among species of Prunus as inferred by isozyme markers. *Theoretical and Applied Genetics*, 80(1), 129–133. <https://doi.org/10.1007/BF00224026>
14. Amit Roy, S. R. (2014). Molecular Markers in Phylogenetic Studies-A Review. *Journal of Phylogenetics & Evolutionary Biology*, 02(02). <https://doi.org/10.4172/2329-9002.1000131>
15. Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., & Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, 102(23), 8369–8374. <https://doi.org/10.1073/pnas.0503123102>
16. Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., & Donoghue, M. J. (1995). The ITS Region of Nuclear Ribosomal DNA: A Valuable Source of Evidence on Angiosperm Phylogeny. *Annals of the Missouri Botanical Garden*, 82(2), 247. <https://doi.org/10.2307/2399880>
17. Alice, L. A., & Campbell, C. S. (2017). Phylogeny of Rubus (Rosaceae) Based on Nuclear Ribosomal DNA Internal Transcribed Spacer Region Sequences Author (s): Lawrence A. Alice and Christopher S. Campbell Published by: Botanical Society of America, Inc. Stable URL: <http://www.jstor.org>. 86(1), 81–97.
18. White, T. J., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and Direct Sequencing of Fungal Ribosomal Rna Genes for Phylogenetics. *PCR Protocols*, January, 315–322. <https://doi.org/10.1016/b978-0-12-372180-8.50042-1>
19. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018a). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549. <https://doi.org/10.1093/molbev/msy096>

20. Lee, S., & Wen, J. (2001). A phylogenetic analysis of *Prunus* and the Amygdaloideae (Rosaceae) using ITS sequences of nuclear ribosomal DNA. *American Journal of Botany*, 88(1), 150–160. <https://doi.org/10.2307/2657135>
21. Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 17(6), 368–376. <https://doi.org/10.1007/BF01734359>
22. Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17(8), 754–755. <https://doi.org/10.1093/bioinformatics/17.8.754>
23. Tajima, F. (1993). Statistical analysis of DNA polymorphism. In *Japanese Journal of Genetics* (Vol. 68, Issue 6, pp. 567–595). <https://doi.org/10.1266/jjg.68.567>
24. Sun, Y. L., Park, W. G., Oh, H. K., & Hong, S. K. (2014). Genetic diversity and hybridization of *Pulsatilla tongkangensis* based on the nrDNA ITS region sequence. *Biologia (Poland)*, 69(1), 24–31. <https://doi.org/10.2478/s11756-013-0284-1>
25. Li, D. Z., Gao, L. M., Li, H. T., Wang, H., Ge, X. J., Liu, J. Q., Chen, Z. D., Zhou, S. L., Chen, S. L., Yang, J. B., Fu, C. X., Zeng, C. X., Yan, H. F., Zhu, Y. J., Sun, Y. S., Chen, S. Y., Zhao, L., Wang, K., Yang, T., & Duan, G. W. (2011). Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proceedings of the National Academy of Sciences of the United States of America*, 108(49), 19641–19646. <https://doi.org/10.1073/pnas.1104551108>
26. Bortiri, E., Sang-Hun, O. H., Jiang, U. O., Baggett, S., Granger, A., Weeks, C., Buckingham, M., Potter, D., & Parfitt, D. E. (2001). Phylogeny and systematics of *Prunus* (Rosaceae) as determined by sequence analysis of ITS and the chloroplast trnL-trnF spacer DNA. *Systematic Botany*, 26(4), 797–807. <https://doi.org/10.1093/biomet/53.3-4.325>
27. Wen, J., Berggren, S. T., Lee, C.-H., Ickert-Bond, S., Yi, T.-S., Yoo, K.-O., Xie, L., Shaw, J., & Potter, D. (2008). Phylogenetic inferences in *Prunus* (Rosaceae) using chloroplast ndhF and nuclear ribosomal ITS sequences. *Journal of Systematics and Evolution*, 46(3), 322–332. <https://doi.org/10.3724/SP.J.1002.2008.08050>

Tables

Table.1. GPS co-ordinates of the *Prunus* species used in the present study.

Species name	Location	Latitude	Longitude
1. <i>Prunus domestica</i>	Chandel, Manipur	24° 19' 34.32" N	94° 0' 2.16" E
2. <i>Prunus cerasoides</i>	NH-2 b/w khuzama and viswema (Nagaland)	25° 32' 82.4" N	94° 07' 65.7" E
3. <i>Prunus napaulensis</i>	Laitlyngkot, Meghalaya	25° 26' 53.83" N	91° 50' 34.13" E
4. <i>Prunus cerasoides</i>	Shirui village, Manipur	25° 07' 46.1" N	94° 0' 2.16" E
5. <i>Prunus cerasoides</i>	Sikkim	25° 53' 30" N	88° 51' 22" E
6. <i>Prunus campanulata</i>	Mao, Manipur	25° 20' 25" N	94° 31' 17" E
7. <i>Prunus cerasoides</i>	Arunachal Pradesh, Ziro	27° 35' 41.89" N	93° 50' 18.74" E
8.. <i>Prunus cerasoides</i>	Wards lake, Meghalaya.	25° 34' 48.8" N	91° 53' 26.7" E
9.. <i>Prunus cerasoides</i>	Kodaikanal	10° 13' 41.00" N	77° 28' 12.16" E

Table.2. List of *Prunus* taxa and outgroup sampled for this study and their GenBank accession numbers.

Taxa and site of collection	Accession numbers of ITS sequences in GenBank
<i>Prunus domestica</i> Chandel, Manipur	MT811814
<i>Prunus campanulata</i> Mao, Manipur	MT811815
<i>Prunus napaulensis</i> Meghalaya	MT811816
<i>Prunus cerasoides</i> Meghalaya	MT811817
<i>Prunus cerasoides</i> Chennai	MT811818
<i>Prunus cerasoides</i> Sikkim	MT811819
<i>Prunus cerasoides</i> Nagaland	MT811820
<i>Prunus cerasoides</i> Arunachal Pradesh	MT811821
<i>Prunus cerasoides</i> Ukhrul, Manipur	MT828305
<i>Prunus cyclamina</i>	AF179503.1
<i>Prunus cerasoides</i> var <i>campanulata</i>	AF179503.1
<i>Prunus serrulata</i> var <i>spontanea</i>	AH010149.2
<i>Prunus verecunda</i>	AF179522.1
<i>Prunus maackii</i>	AH010152.2
<i>Prunus fordiana</i>	AF179530.1
<i>Prunus laurocerasus</i>	AH010157.2
<i>Prunus serotina</i>	AH0101542
<i>Prunus ilicifolia</i> subsp. <i>lyonii</i>	AH010155.2
<i>Prunus tenella</i>	AH010142.2
<i>Prunus persica</i>	AF179562.1
<i>Prunus mume</i>	AH009376.2
<i>Prunus tomentosa</i>	AF179500.1
<i>Prunus domestica</i>	AF179485.1
<i>Prunus salicina</i>	AH009373.2
<i>Prunus armericana</i>	AF179489.1
<i>Prunus armericana</i>	AF179488.1
<i>Prunus augustifolia</i>	AF179490.1
<i>Prunus nigra</i>	AH009374.2
<i>Prunus mandshuria</i>	AH009375.2
<i>Prunus caroliniana</i>	AF179540.1

<i>Prunus grayana</i>	AF179531.1
<i>Pyrus communis</i>	JQ392467.1

Table.3. Sizes and Percent G+C content of ITS 1, 5.8s and ITS 2 of *Prunus* taxa.

Species name (Taxon)/ accession number	Size (bp) ITS1	5.8s	ITS2	A	T	A+T	G	C	G+C	GC%	Total Length
<i>Prunus domestica</i> MT811814	233	102	234	113	118	231	190	176	366	61.3	569
<i>Prunus napaulensis</i> MT811816	233	102	243	109	108	217	194	189	383	63.8	578
<i>Prunus cerasoides</i> MT828305	226	102	305	146	142	288	175	174	349	54.8	633
<i>Prunus campanulata</i> MT811815	232	102	233	130	116	246	170	151	321	56.60	567
<i>Prunus cerasoides</i> MT811821	232	102	232	155	110	265	166	135	301	53.20	566
<i>Prunus cerasoides</i> MT811817	241	102	232	116	121	237	170	168	338	58.80	575
<i>Prunus cerasoides</i> MT811818	232	102	237	112	121	233	169	169	338	59.20	572
<i>Prunus cerasoides</i> MT811821	232	102	232	114	121	235	165	166	331	58.50	566
<i>Prunus cerasoides</i> MT811819	256	102	232	128	118	246	165	179	344	58.30	590

Figures

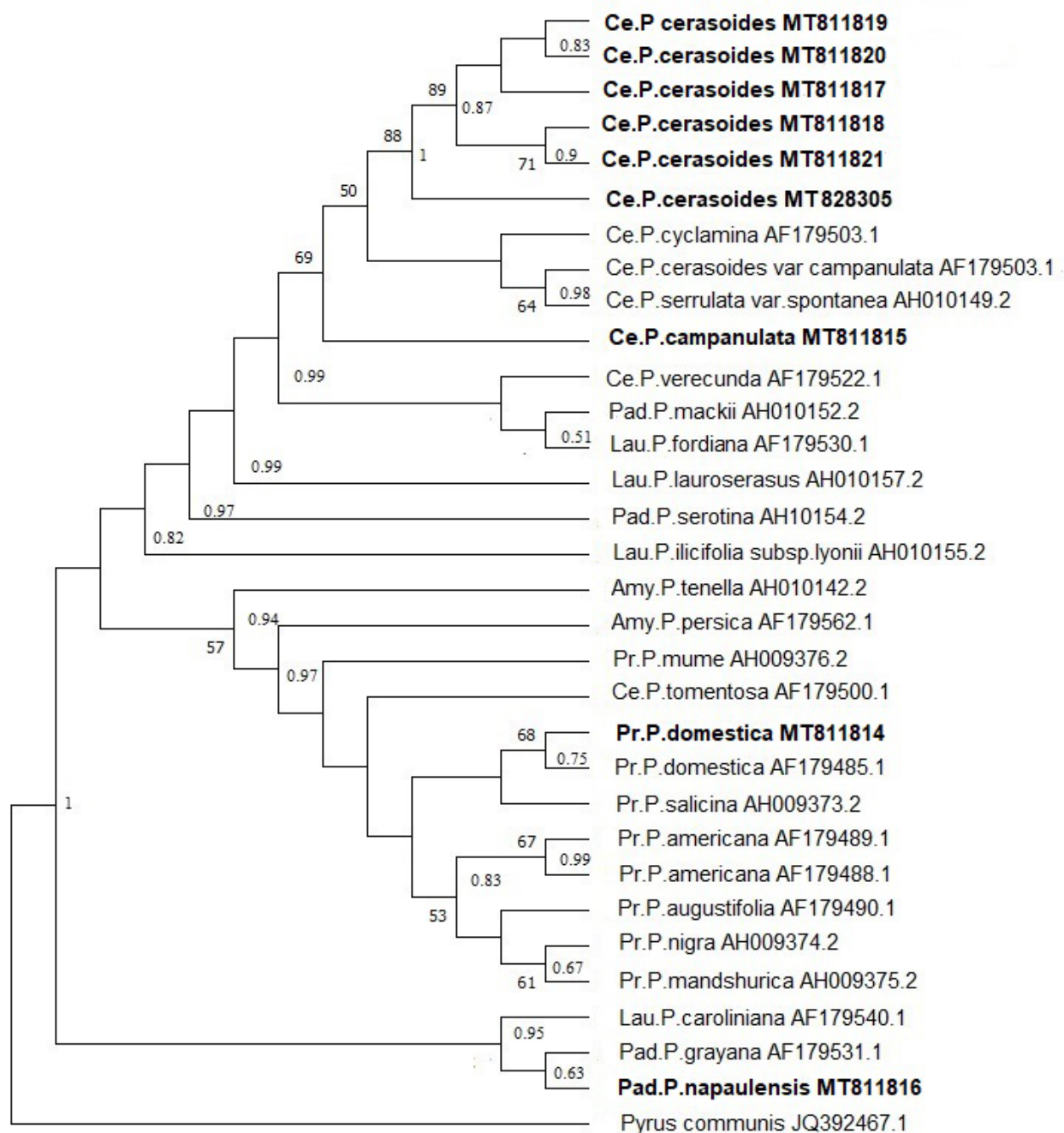


Figure 1

Legend not included with this version