PHYLOGENETIC ANALYSES OF THE 13 MAPLE (*ACER*) SPECIES BASED ON CHLOROPLAST GENES PSBA AND RBCL

A Thesis
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Master of Professional Studies

by
Linkai Mei
August 2021

Abstract This study uses two chloroplast genes (psbA and rbcL) and morphological data collected online and from herbarium specimens to generate phylogenetic trees for 13 species of maple (*Acer*). Alignment of sequences and phylogenetic analyses were performed and visualized with the programs MUSCLE, RAXML, MrBayes, MESQUITE, and ITOL. ML and Bayesian analyses were used to generate phylogenetic trees, and these were compared with results from a previous study by Li. Results indicate that *Acer griseum*, *Acer caesium*, *Acer hyrcanum*, and *Acer platanoides* form a monophyletic terminal clade, and this is the most robust result of these analyses. *Acer grandidentatum* and *Acer saccharum* are monophyletic in the ML analysis, but unresolved in the Bayesian analyses, which makes sense because *Acer grandidentatum* is a subspecies of *Acer saccharum*.

INTRODUCTION

Acer is one of the most important genera of trees in the broad-leaved deciduous forests of the Northern Hemisphere, with about 130 species in eastern Asia, 10 species in North America, and 12 species in Europe and western Asia. (Li, et al. 2019) They are primarily deciduous shrubs and trees of temperate forest regions of the Northern Hemisphere, with a few species extending into the subtropics. (Ackerly, et al. 1998) Maple trees have been used as sources of commercial products, for example, maple syrup from *Acer saccharum* and construction timber from *A. saccharum*, *A. rubrum*, and *A. pseudoplatanus*. (Li, et al. 2019)

Although the genus *Acer* is easily recognizable based on leaf phyllotaxy and the distinctive samara, it is very polymorphic. The leaves vary from simple to pinnate and palmate. However, more important than the leaves for the classification of the taxa, is the large variability of inflorescences, flowers, fruits, and the number of bud scales. This enormous variability of maples certainly has attracted the attention of botanists and there are still many gaps and unanswered questions. (De Jong, et.al 2004) As in previous articles, different phylogenetic trees are obtained with different methods.

For over a century there has been no further discussion about defining the genus *Acer*. The distinctive winged fruits (samaras), in combination with the opposite leaves, make maples easily recognizable and a well-defined genus. Researchers have been studying maples for centuries, but in many cases, the authors had no flowers or fruits at their disposal, and as a consequence, the importance of leaf characteristics has been overestimated, e.g., the absence or presence of hair on the underside of the leaves. Hence, techniques such as DNA analysis can be helpful for further studying and defining the *Acer* species and the corresponding subspecies. (De Jong, et al. 2004) A previous study on the phylogeny of *Acer* was based on DNA sequences of over 500 nuclear loci generated using the anchored hybrid enrichment method (Li, et al. 2019). This paper adds another original perspective by using two chloroplast genes, which ensures that these results are independent of the previous results (Li, et al., 2019) based only on nuclear markers.

This study uses two genes downloaded from GenBank with additional morphological data that were collected online and in the LH Bailey Herbarium to generate phylogenetic trees.

MATERIALS AND METHODS

Based on the availabilities in GenBank and a previous study (supplemental figure 1). *Acer Platanoides*, *Acer Griseum*, and *Acer Rubrum* were used as outgroups. Morphological data were gathered from the internet and the LH Bailey Herbarium.

Molecular sequences - The genes psbA and rbcL were downloaded in FASTA format from GenBank for as many as 13 *Acer* species (see appendices for accession numbers). These were assembled into 2 FASTA files, each file has the same gene with different species.

Alignment - The program MUSCLE was used to align sequences downloaded from GenBank, using default settings. This produced aligned sequences in FASTA format, which were then used for subsequent analyses.

ML analysis - MESQUITE (http://www.mesquiteproject.org/) and an online converter (http://sequenceconversion.bugaco.com/converter/biology/sequences/fasta_to_phylip.php) to convert the aligned FASTA files to PHYLIP files. Analyze the PHYLIP files with RAXML (Stamatakis, et al. 2014) (http://herbarium.bh.cornell.edu/courses/plbio4400_2021/PROGRAMS/raxml.exe (Windows), http://www.sfu.ca/biology2/staff/dc/raxml/ (MAC)). Example command as follows: raxml -m GTRCAT -s /datasets/ruhfel_2014_phy/accD.phy -n accD. Load the "best tree" files into MESQUITE or the ITOL website (https://itol.embl.de/upload.cgi) and output trees (as graphic trees) to PDF files, using new Simple newick/phylip as the interpreter.

Bayesian analyses - MrBayes (Ronquist, et al. 2001) was used for Bayesian phylogenetic analyses. MESQUITE was used to convert the aligned FASTA files to NEXUS files. For these analyses the following parameters were used: ngen = 100000 and the GTRGAMMA model was used within the mrbayes command block. Using the command: mb - i filename.nex.

'Total evidence' - MESQUITE was used to combine different molecular data sets with partition information as follows:

```
DNA, gene1 = starting position – ending position

DNA, gene2 = starting position – ending position

raxml -m GTRCAT -p 12345 -q gene1_gene2_part.txt -s gene1_gene2.phy -n gene1_gene2.
```

Tree output was converted to graphic trees as PDF files using MESQUITE.

Quick bootstrap values - raxml -f a -m GTRGAMMA -p 12345 -T2 -x 12345 -#100 -s gene1_gene2.phy -n gene1_gene2. The resulting Raxml_bipartitions.gene1_gene2 can be read into the Mesquite, treat the node values as bootstrap values to display the bootstrap values.

Posterior probabilities for MrBayes - gene1_gene2_mb.nex.con.tre is read into ITOL, in the control panel -> advanced -> bootstraps / metadata -> Display -> data source -> prob+-sd -> prob(percent). Export it as PDF file.

Mapping morphology onto the tree - MESQUITE was used.

RESULTS

Molecular trees: Neither ML nor Bayesian analyses resolved any major subgroups within the selected *Acer* species used for this analysis (Fig.1,Fig. 2, and Fig. 3). The results of Bayesian and ML were also different in details of relationships. A monophyletic group that was resolved with both methods was a clade that includes four species: *Acer griseum*, *A. caesium*, *A. hyrcanum* and *A. platanoides*, and within this clade, the same arrangement of species was found.

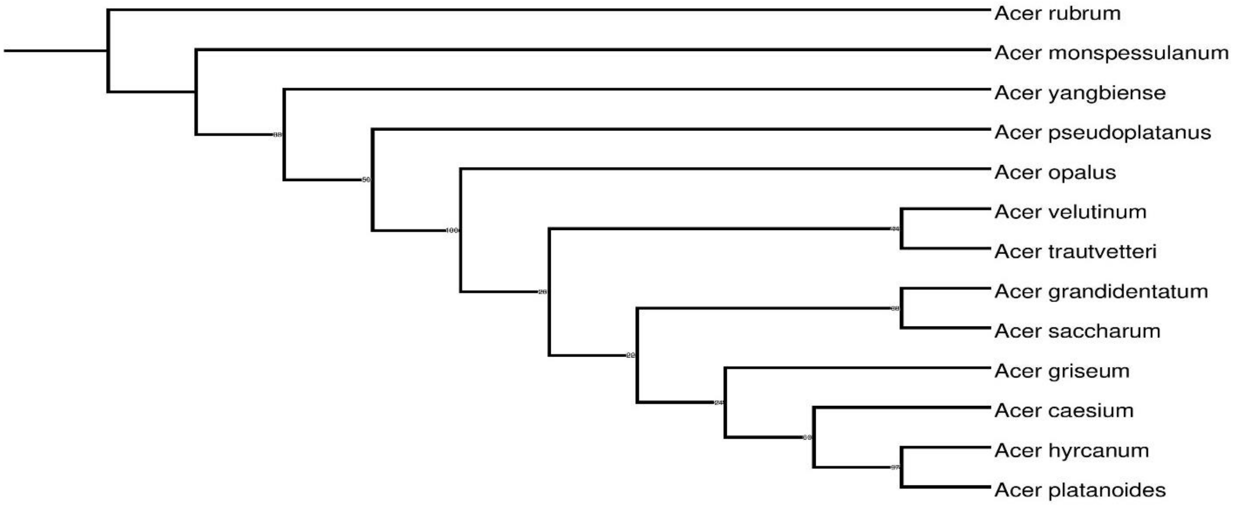


Figure 1: (RAXML) Maximum likelihood Bootstrap mapped on the best tree (psbA_rbcL)

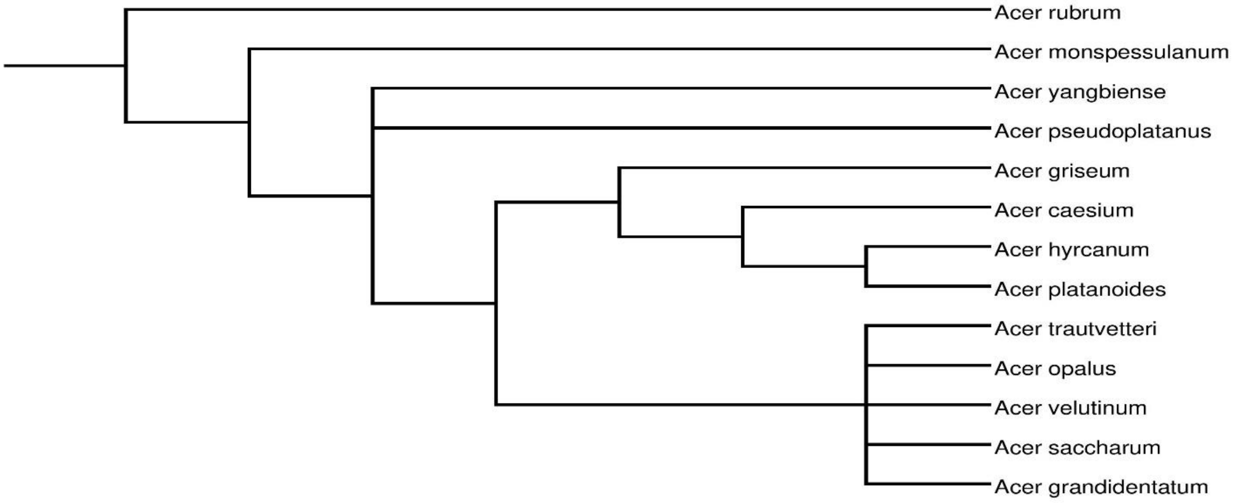


Figure 2: (MrBayes) Bayesian analysis (psbA_rbcL)

Figure 3: (MrBayes) Bayesian analysis with posterior probability (psbA_rbcL)

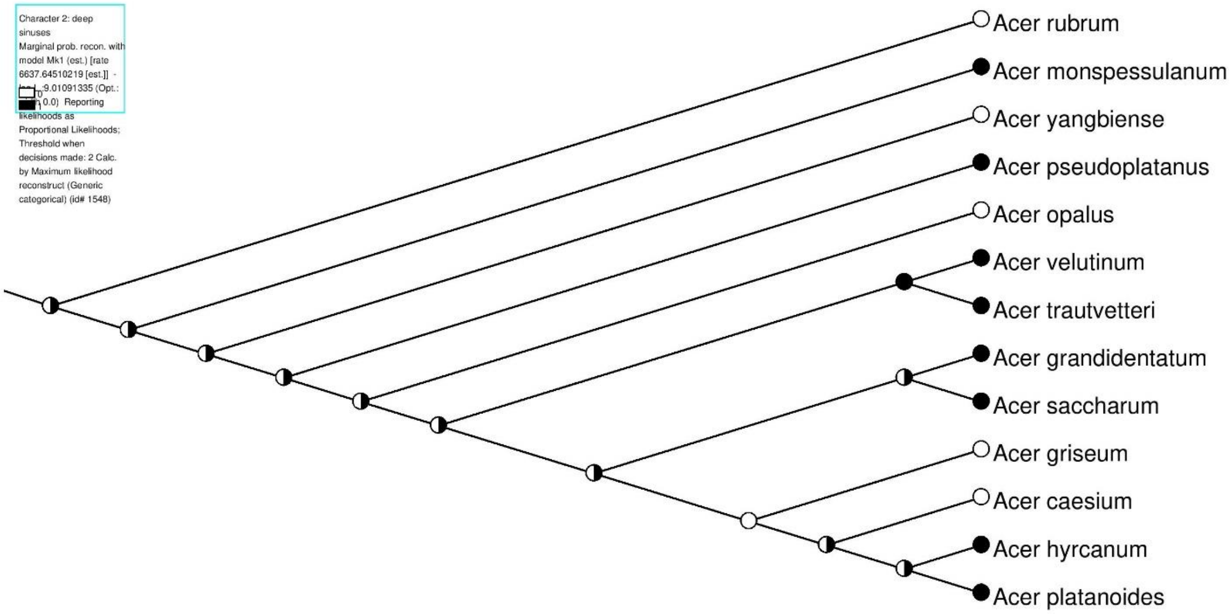


Figure 4: (MESQUITE) Morphological character (with deep sinuses: 1; without deep sinuses: 0) mapped on Maximum likelihood Bootstrap (psbA_rbcL) with MK model

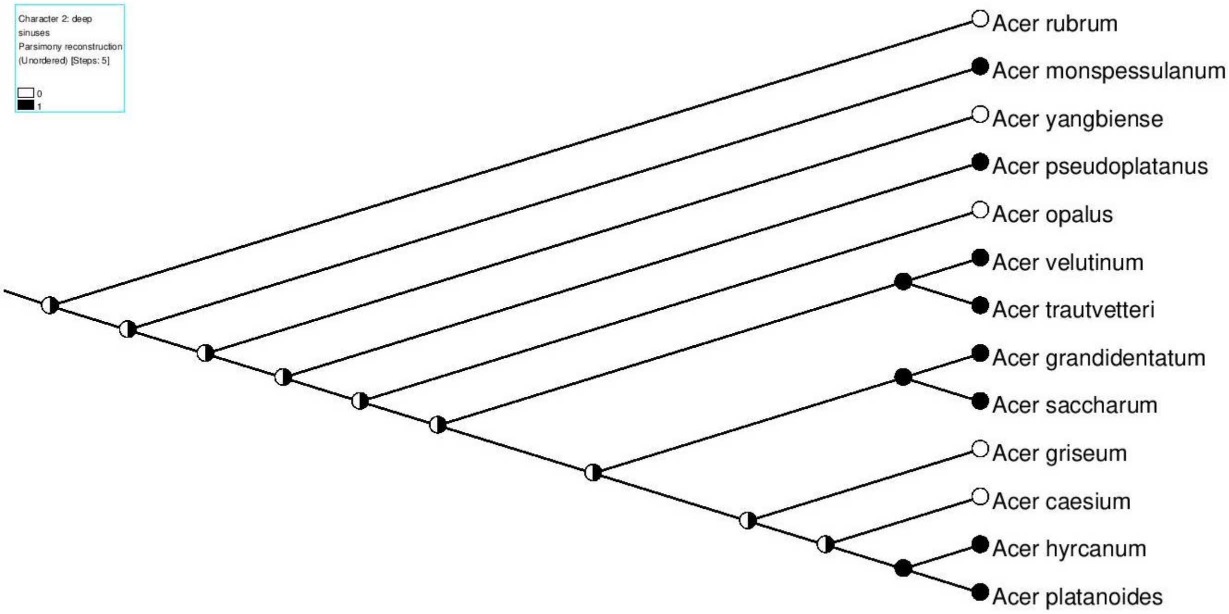


Figure 5: (MESQUITE) Morphological character (with deep sinuses: 1; without deep sinuses: 0) mapped on Maximum likelihood Bootstrap (psbA_rbcL) with parsimony

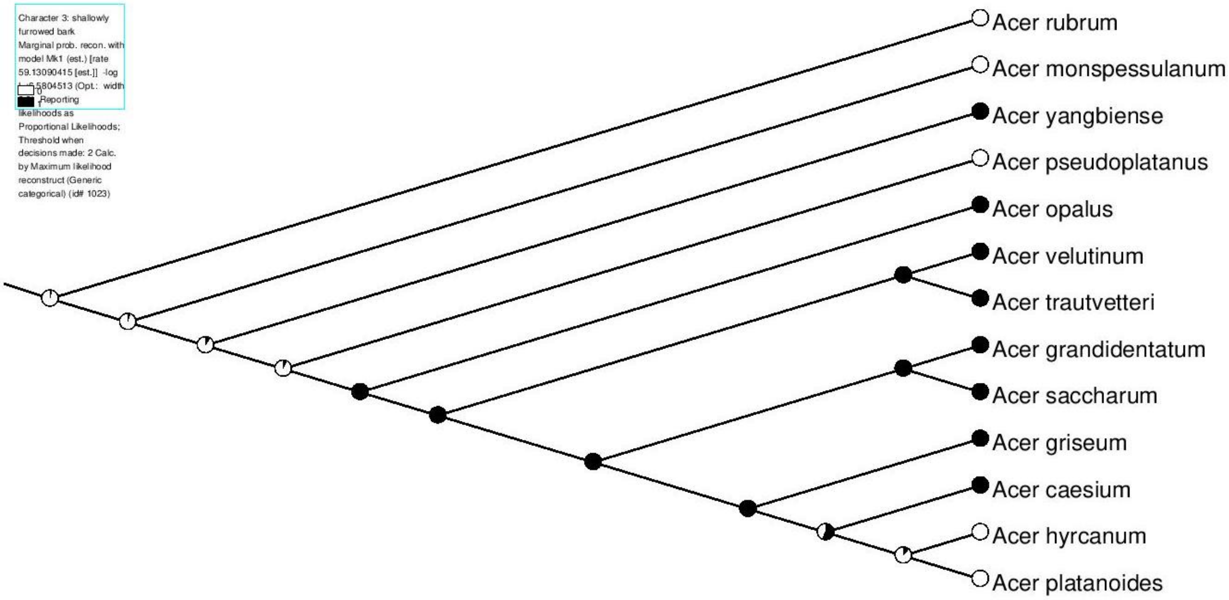


Figure 6: (MESQUITE) Morphological character (with shallowly furrowed bark: 1; without shallowly furrowed bark: 0) mapped on Maximum likelihood Bootstrap (psbA_rbcL) with MK model

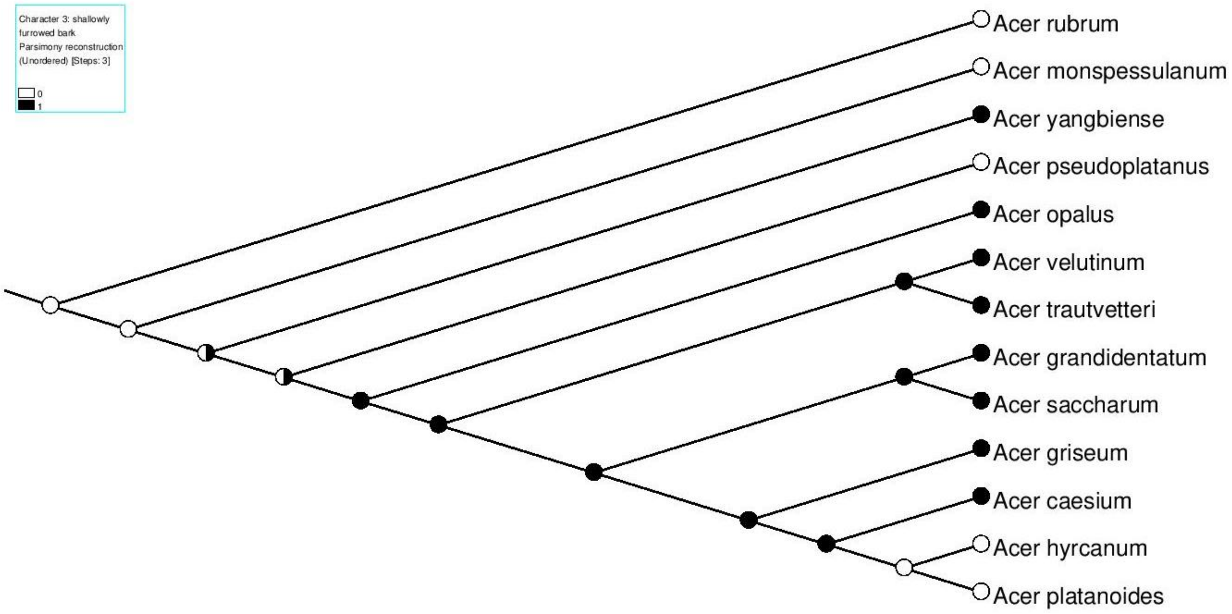


Figure 7: (MESQUITE) Morphological character (with shallowly furrowed bark: 1; without shallowly furrowed bark: 0) mapped on Maximum likelihood Bootstrap (psbA_rbcL) with parsimony

ML analysis: *Acer rubrum*, *Acer monspessulanum*, *Acer yangbiense*, and *Acer pseudoplatanus* form a basal grade relative to a clade of nine species including *Acer saccharum* in both ML and Bayesian analyses. *Acer hyrcanum* and *Acer platanoides* are sister taxa with an ML bootstrap value of 97.

Bayesian analyses: *Acer griseum*, *Acer caesium*, *Acer hyrcanum*, and *Acer platanoides* form a monophyletic group that is internally unresolved, and they are also monophyletic in the ML analysis. In figure 2: *Acer trautvetteri*, *Acer opalus*, *Acer velutinum*, *Acer saccharum*, and *Acer grandidentatum* are not monophyletic but they form a monophyletic group in Figure 3..

Morphology: Two examples of morphological characters (leaf sinuses and bark texture) are mapped in Figures 4, 5, 6 and 7. These example characters, like others not presented here, are not completely consistent with either the ML or Bayesian trees based upon morphology. Of more interest is the difference between the MK ancestral reconstruction and the parsimony reconstruction. In some cases the MK reconstruction is more ambiguous (as indicated by ancestral states being either state) and in other cases less ambiguous.

DISCUSSION

Molecular trees: Overall, relationships within the genus are ladder-like and not easily resolved into subgroups. In the ML tree, *Acer grandidentatum* and *Acer saccharum* form a monophyletic group, which is expected because *Acer grandidentatum* is usually considered to be a subspecies of *Acer saccharum* (Grimm, et al., 2007). However, this relationship was not recovered in the Bayesian trees. Thus, the results presented here do not support dividing these species of *Acer* into taxonomic subgroups, and the results of Li are equally inconclusive. Differences between the results presented here and the previous results from Li can be attributed to the differences between the genes used in this article vs the genes used by Li as well as a more exhaustive sample of species by Li.

Morphological trees: Morphological data (figures 4 – 7) do not provide robust information for the species when compared with the current and previous molecular results. Differences in ancestral reconstructions between Mk and parsimony are minimal, but might have some effect on interpretation of character evolution. At this point, it appears that the addition of morphological data will not help to resolve relationships within this group of *Acer*.

ACKNOWLEDGEMENTS

The LH Bailey Herbarium, Anna Stalter and Peter Fraissenet for access to materials. Dr. Nixon and Dr. Scanlon for mentoring during the course of this project. Those Lovely maple trees on Cornell campus.

LITERATURE CITED

Ackerly D, Donoghue MJ. 1998. Leaf size, sapling allometry, and Corner's rules: Phylogeny and correlated evolution in maples (*Acer*). The American Naturalist 152: 767–791.

Boulter MC, Benfield JN, Fisher HC, Gee DA, Lhotak M. 1996. The evolution and global migration of the *Aceraceae*. Philosophical Transactions of the Royal Society B: Biological Sciences 351: 589–603.

De Jong PC. 2004. World maple diversity. In: Wiegrefe SJ, Angus H, Otis D, Gregory P eds. International Maple Symposium, Westonbirt Arboretum and Royal Agricultural College in Gloucestershire, England, 2004. Gloucestershire: Westonbirt Arboretum. 2–11.

Grimm GW, Denk T, Hemleben V. 2007. Evolutionary history and systematics of *Acer* section *Acer*—A case study of low-level phylogenetics. Plant Systematics and Evolution 267: 215–253.

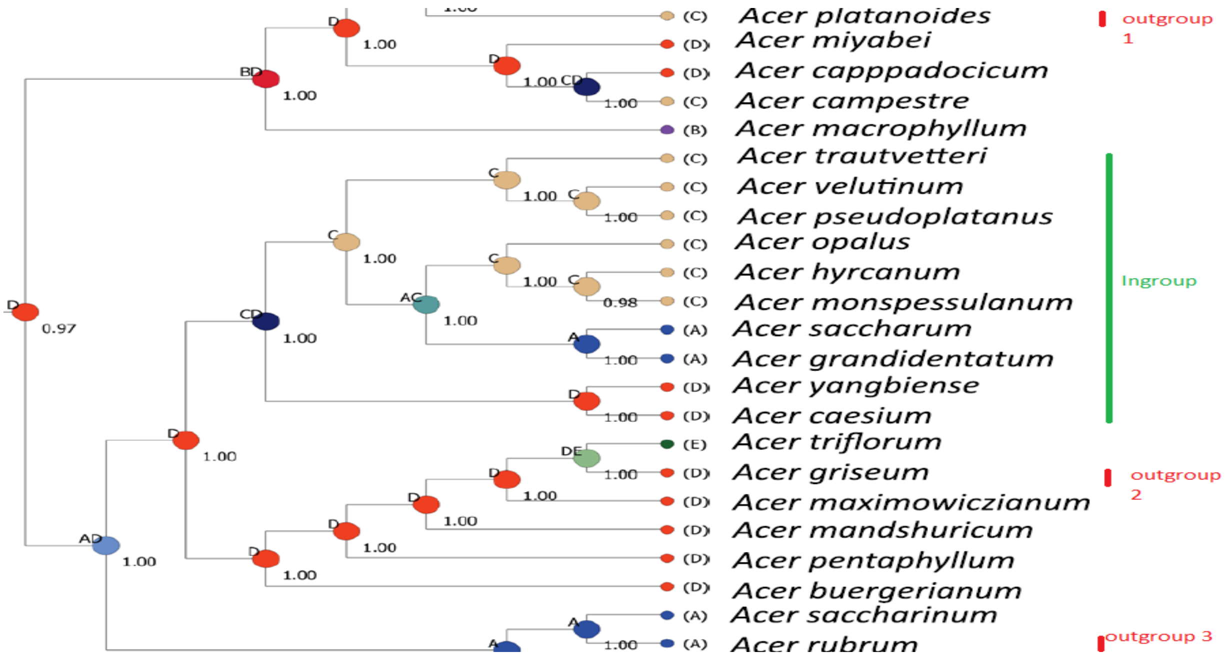
Huelsenbeck, Ronquist, Nielsen & Bollback. 2001. Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology Science 294: 2310-2314.

Li, Jianhua. 2019. Maple phylogeny and biogeography inferred from phylogenomic data. Journal of Systematics and Evolution 57: 594-606.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312–1313.

APPENDICES

Supplemental Figure



Supplemental figure 1: Phylogeny of Acer based on DNA sequences of over 500 nuclear loci (Li, et.al 2019)

GenBank accessions

psbA

>DQ978638.1 Acer trautvetteri PsbA (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His (trnH) gene, partial sequence; chloroplast

>LC335814.1 Acer velutinum var. glabrescens chloroplast DNA, psbA-trnH intergenic spacer region, clone: CLN08

>KU672818.1 Acer pseudoplatanus voucher Aps_5282 psbA-trnH intergenic spacer region, partial sequence; chloroplast

>DQ978614.1 Acer opalus PsbA (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His (trnH) gene, partial sequence; chloroplast

>MT114381.1 Acer hyrcanum psbA-trnH intergenic spacer, partial sequence; chloroplast

>KU672783.1 Acer monspessulanum voucher Amo_6105 psbA-trnH intergenic spacer region, partial sequence; chloroplast

>DQ978630.1 Acer saccharum subsp. saccharum PsbA (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His (trnH) gene, partial sequence; chloroplast

>DQ978627.1 Acer grandidentatum PsbA (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His (trnH) gene, partial sequence; chloroplast

>KU945335.1 Acer yangbiense voucher KM07 psbA-trnH intergenic spacer, partial sequence; chloroplast

>MH116941.1 Acer caesium voucher YLDP002C trnH-psbA intergenic spacer region, partial sequence; chloroplast

>DQ978619.1 Acer platanoides PsbA (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His (trnH) gene, partial sequence; chloroplast

>DQ978600.1 Acer griseum PsbA (psbA) gene, partial cds; psbA-trnH intergenic spacer, complete sequence; and tRNA-His (trnL) gene, partial sequence; chloroplast

>HQ266313.1 Acer rubrum voucher V_658_A_ALheml10 tRNA-His (trnH) gene, partial sequence; trnH-psbA intergenic spacer, complete sequence; and PsbA (psbA) gene, partial cds; chloroplast

rbcL

>DQ978438.1 Acer trautvetteri ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast

>DQ978425.1 Acer pseudoplatanus ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast

>DQ978420.1 Acer opalus ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast

>DQ978432.1 Acer saccharum subsp. saccharum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast

>DQ978397.1 Acer caesium subsp. giraldii ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast

>DQ978424.1 Acer platanoides ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast

>DQ978411.1 Acer griseum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast

>DQ978428.1 Acer rubrum ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) gene, partial cds; chloroplast