

A molecular re-examination of phylogenetic relationships among *Juniperus*, *Cupressus*, and the *Hesperocyparis-Callitropsis-Xanthocyparis* clades of Cupressaceae

Randall G. Terry

Biology Department, Lamar University, P. O. Box 10037, Beaumont, TX, 77710, USA
rgterry@lamar.edu

and

Robert P. Adams

Baylor University, Biology Department, One Bear Place, 97388, Waco, TX 76798, USA

ABSTRACT

Previous molecular phylogenetic studies have recovered conflicting hypotheses of relationship among *Juniperus* (J), *Cupressus* (C), and *Hesperocyparis-Callitropsis-Xanthocyparis* (HCX). Conflict between nuclear genes, chloroplast genes, and nuclear and chloroplast data have all been realized in recovering all possible topologies among the three clades. In this study, we use 2.2 kb of aligned sequence from two nuclear loci, and 11.4 kb of sequence from 11 chloroplast regions, in re-examining relationships among J-C-HCX. Unlike previous studies, we find unambiguous support for relationships in the nuclear data, whether the genes are analyzed individually or in combination. In contrast, character conflict between different chloroplast partitions, or even between characters from a single region, results in nearly equally well-supported but conflicting hypotheses of relationship. Statistical tests of likelihood values indicate the chloroplast data always fails to distinguish between two of three competing sister group relationships, and in one instance cannot differentiate between any of the three possible J-C-HCX topologies. Results presented here suggest a complex evolutionary history in which molecular processes in addition to possible ancient hybridization have obscured J-C-HCX relationships. Published on-line www.phytologia.org *Phytologia* 97(1): 67-75 (Jan 2, 2015). ISSN 030319430.

KEY WORDS: chloroplast DNA, nrDNA, *Juniperus*, *Hesperocyparis-Callitropsis-Xanthocyparis* (HCX), parsimony, Bayesian analysis

Cupressaceae is the third largest gymnosperm family with over 130 species in about 33 genera (Farjon, 2005; Farjon et al., 2002; Adams et al., 2009). The family is well represented in both the northern and southern hemispheres, with members occupying all habitable continents and occurring in a variety of habitats (Farjon, 2005; Adams, 2014.) Rarity and high degrees of endemism are disproportionately represented in the family, with 18 genera being monotypic and 27 having five or fewer species (Farjon, 2005). Among the more diverse genera in the family are *Juniperus* (67 species, 34 varieties), many species of which are adapted to semi-arid habitats in the northern Hemisphere, *Cupressus*, a genus of 12 species (Little, 2005) geographically centered in Asia (Mao et al., 2010) and generally known as the “Old World cypresses” (OWC), and *Hesperocyparis*, a recently recognized genus of 17 species (Adams et al., 2009; Adams et al., 2014; Wolf, 1948) from the western United States, Mexico, and central America (i. e., the New World cypresses or NWC).

A spate of phylogenetic studies published over the last decade have resulted in new perspectives on the phylogeny of *Juniperus*, *Cupressus*, *Hesperocyparis* and related taxa (Little et al., 2004; Little, 2005; Little, 2006; Adams et al., 2009; Yang et al., 2012; Terry et al., 2012). The recovery and taxonomic recognition of *Hesperocyparis* as distinct from *Cupressus* (Adams et al., 2009), strong support for inclusion of *Callitropsis* and *Xanthocyparis* in a lineage with *Hesperocyparis* (i. e., the HCX lineage of Terry et al., 2012; Little et al., 2004; Little, 2006; Adams et al., 2009), and studies elucidating species

relationships (Terry et al., 2012) and the recognition of new species (Adams et al., 2014) within *Hesperocyparis* and *Juniperus* (Mao et al., 2010), collectively represent our improved understanding of evolutionary and taxonomic relationships in the group.

Despite these advances, a number of outstanding questions remain. Among these are relationships among certain genera of Cupressaceae (Gadek et al., 2000; Yang et al., 2012), including those among *Juniperus* (J), *Cupressus* (C), and HCX. One of the first studies to address relationships among *Juniperus* and Old and New world representatives of *Cupressus* was that of Gadek et al. (2000), which used molecular and morphological data in addressing relationships among the major lineages of Cupressaceae. Parsimony analysis of cpDNA sequences recovered a clade containing distinct Old and New World *Cupressus* as sister to *Juniperus*, i. e., J (OWC, NWC) (Gadek et al., 2000). Two subsequent studies used cpDNA sequences to corroborate the findings of Gadek et al. (2000) in recovering a J(C,HCX) topology (Adams et al., 2009; Yang et al., 2012), while relationships among the three lineages were unresolved in a third study that used cpDNA (see Little, 2006). Three studies have used DNA sequences from a total of five nuclear loci in addressing relationships among J, C, and HCX. Two general patterns emerge from these studies: nrITS sequences always yield a C(J,HCX) topology (Adams et al., 2009; Little, 2006), and the other data sets, either alone or in various combinations, yield a HCX(J,C) topology (Little, 2006; Adams et al., 2009; Yang et al., 2012). Collectively, these findings indicate conflict between nrITS and other nuclear data sets (ABI3, 4CL, Needly, and Leafy) in resolving relationships among J-C-HCX, but in no instance are phylogenies derived from nuclear data congruent with those based on cpDNA, a finding some authors attribute to ancient hybridization (Yang et al., 2012).

In this study, we re-examine relationships among J-C-HCX using nearly 13.7 kb of aligned DNA sequence from both the chloroplast and nuclear genomes. Results from separate analyses of the cytoplasmic and nuclear data as well as combined analyses are used to re-assess relationships among J-C-HCX.

MATERIALS AND METHODS

Specimens used in this study with voucher information and GenBank accession numbers are provided in Table 1. For all specimens, one gram (fresh weight) of foliage was placed in 20g of activated silica in the field, and subsequently stored at -20°C in the lab.

DNA extraction, PCR amplifications, and preparation of sequencing templates are according to Terry et al. (2012). Briefly, total genomic DNA was extracted from 0.020 g of silica dried leaf tissue using a DNeasy Plant Mini Kit according to the manufacturer's instructions (Qiagen, Valencia, CA, USA). The psbD-trnT intergenic spacer and the trnC-trnD intergenic region containing spacer sequence and a portion of the psbM coding region were amplified and sequenced for two species of *Calocedrus*, four species of *Juniperus*, and three species of *Cupressus* (Table 1). All other sequences were previously published (Gadek et al., 2000; Little et al., 2004; Little, 2006; Terry et al., 2012) and are available in GenBank. Thermal cycling protocols for all amplifications were as follows: 94°C for 5 min, followed by 30 cycles of 94°C for 1 min, 2 min at the optimized annealing temperature, and 72°C for 2 min, followed by 72°C for 7 min. Annealing temperatures were 47.5°C for psbD-trnT and 50°C for trnC-trnD. Primer sequences and other amplification details are given in Terry et al. (2012). PCR products were purified by agarose gel electrophoresis according to Terry et al. (2012) and sequenced at McLab Inc. (San Francisco, CA).

Combining data from this study with chloroplast and nuclear sequences from GenBank produced 13.7 kbp of aligned sequence from 9 noncoding chloroplast regions (8 intergenic spacers and one intron), 2 chloroplast genes (rbcL and psbB), and 2 nuclear genes (nrITS and NEEDLY intron 2). For sequences published here, raw sequence from forward and reverse strands was assembled and aligned using Clustal

Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) or MAFFT. Computer generated alignments were manually refined using Seq-AL v.2.0a9 (Rambaut 2002). Both parsimony and Bayesian analyses were performed on each of three data sets: chloroplast data only, nuclear data only, and combined chloroplast and nuclear data. Parsimony analyses were conducted using PAUP*v.4.0b10 (Swofford 2002), with the heuristic search option in effect, simple stepwise addition of taxa, and TBR branch swapping, saving multiple trees. Branch support was assessed by conducting 1000 replicates of bootstrapping with the settings described above. Bayesian analyses were conducted using MrBayes 3.2.1 (Ronquist and Huelsenbeck, 2003) according to Terry et al. (2012). Best-fit evolutionary models were estimated for individual gene regions using the Akaike information criterion (AIC) implemented in jModelTest v.0.0.1 (Posada 2008; Guindon and Gascuel 2003) using the default settings for likelihood calculations and the uncorrected AIC. Bayesian analyses were fully partitioned by gene region, with two independent runs of four Metropolis coupled chains each. Chains were generated from different random trees and run for 1 million generations, sampling every 1,000th generation. In each run, three chains were heated using a temperature of 0.2 with one swap between chains every generation. The burnin fraction was enforced to 0.2 using the “relburnin” command, resulting in the first 200 of 1,000 trees being discarded, and the remaining trees pooled to construct the posterior distribution of the phylogeny. A 50 % majority-rule consensus tree was produced using the “contype = halfcompat” command. Convergence and mixing were assessed by examining plots of likelihood against chain generation over the course of the run and by monitoring the standard deviation of split frequencies among runs in MrBayes.

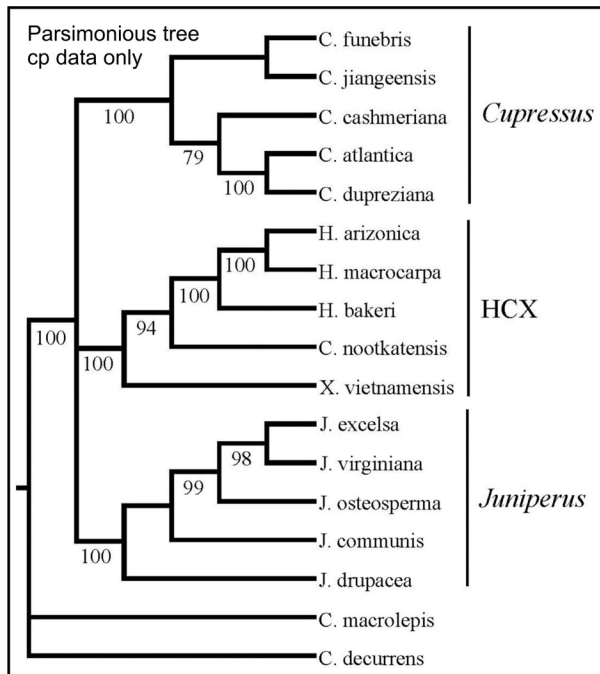
We statistically compared log likelihood values in assessing the relative support of the nuclear and chloroplast data for each the three possible J-C-HCX topologies. Three tests were performed; a 1-sided Kishino-Hasegawa (KH; Goldman et al., 2000), the Shimodaira-Hasegawa (SH; Shimodaira and Hasegawa, 1999), and the expected likelihood weights (ELW; Strimmer and Rambaut 2002). Each test was performed on each of three user defined trees, the two trees from parsimony analysis of the cpDNA only [(HCX(J,C) and J(C,HCX)], and the single tree from the combined parsimony analysis (C(J,HCX). Maximum likelihood analyses and statistical tests of fit were performed in Tree Puzzle 5.2 (Schmidt et al., 2002). Default settings were used in all tests except a gamma distribution with four rate categories was used in estimating rate heterogeneity.

RESULTS AND DISCUSSION

Parsimony analysis of the chloroplast data only produced two shortest-length tree of 759 steps (CI=0.90, RI=0.91; Fig. 1). Bootstrapping of these data produced strong support for most branches, but relationships among J, C, and HCX were unresolved in the 50% majority rule tree (Fig. 1). Of the two most parsimonious trees recovered, one was consistent with previous reports (Gadek et al., 2000; Adams et al., 2009; Yang et al., 2012) in recovering J(C,HCX), while the other recovered HCX(J,C). Bayesian analysis of the chloroplast data also recovered a HCX(J,C) topology, but while nearly all branches had posterior probabilities (pp) of 1.0, the J-C clade was weakly supported (pp= 0.62; Fig. 2) .

In contrast to the chloroplast data, analyses of the nuclear data alone, or of combined nuclear and chloroplast data, consistently produced strong support for a clade containing J and HCX (Figs. 3-5). In addition, parsimony analysis of the nuclear data alone recovered four shortest length trees, all of which contained C(J,HCX), and strong support for a J-HCX sister group relationship (Fig 3). Similarly, parsimony analysis of the combined data produced a single tree of 1522 steps (CI=0.87, RI=0.90) in which a well-supported J-HCX clade was recovered (Fig. 5). Bayesian analysis of nuclear data alone or of combined nuclear and chloroplast data always recovered a C(J,HCX) topology with strong support (pp=1.0) for the J-HCX clade (Figs. 4 and 5).

Maximum likelihood analysis using Tree Puzzle found the HCX(J,C) and J(C,HCX) topologies explained the cpDNA nearly equally well, while the C(J,HCX) topology produced the least likely



explanation of the data. All tests found no significant difference between the HCX(J,C) and J(C,HCX) topologies, and one of three tests (SH) found no difference among the three possible J-C-HCX alternatives (Table 2).

Here, we re-examine relationships between *Juniperus*, *Cupressus*, and HCX with 13.7 kb of aligned DNA sequence. Our data include 2329 bp of aligned sequence from two nuclear genes (nrITS and Needly), and 11402 bp of sequence from 11 chloroplast regions (Table 1). We consistently recover a C(J,HCX) topology from the nrITS and Needly data sets, analyzed either alone or in combination. This finding is supported by results from only one previous study (i.e. Adams et al., 2009, which used combined nuclear and chloroplast data), and is in conflict with the HCX(J,C) topology recovered from analyses of several other nuclear loci

Figure 1. 50% majority-rule consensus of two most parsimonious trees generated from analysis of chloroplast data only. Length = 759 steps, CI = 0.90, RI = 0.91. Numbers below branches are bootstrap values, and are not given for values less than 50%.

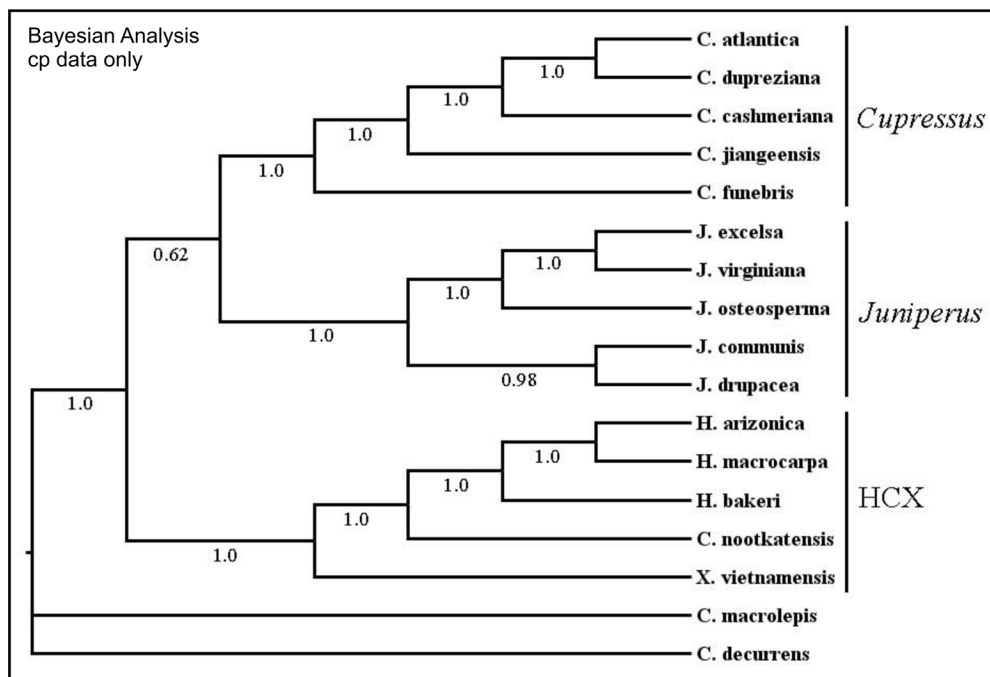
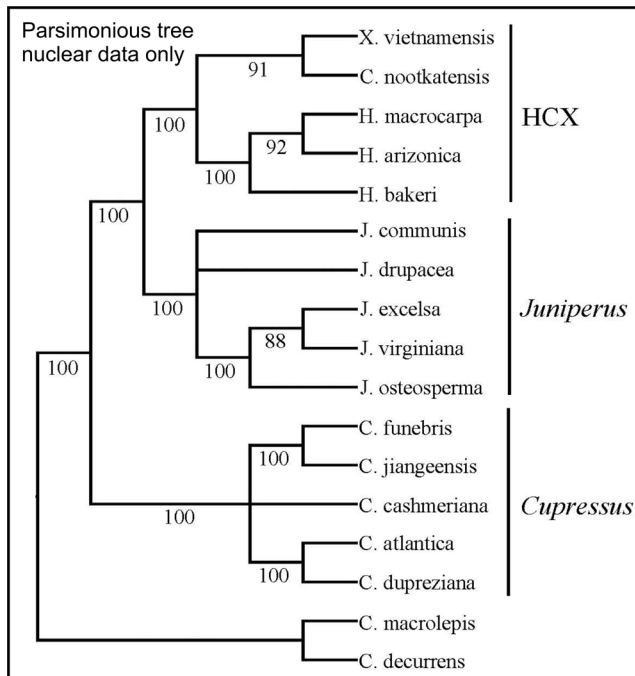


Figure 2. 50% majority-rule consensus tree generated from Bayesian analysis of chloroplast data only. Numbers below branches are posterior probabilities.



(Little, 2006; Adams et al., 2009; Yang et al., 2012). Nevertheless, we find strong support for a J-HCX sister group relationship (bootstrap=100, pp=1.0) in all analyses including nuclear data (Figs. 3-5), and note that is no instance is nuclear data unable to statistically distinguish C(J,HCX) from either of the other two J-C-HCX alternatives (data not shown). Moreover, character analysis identified 42 synapomorphies for the J-HCX clade in the combined analysis, 35 from the nuclear genes and nearly equally divided between nrITS and Needly (combined CI of 0.88), and 7 from the chloroplast data (CI=0.79).

In contrast to the nuclear data, chloroplast sequences do not provide strong support for any particular hypothesis of J-C-HCX relationship. Perhaps this is best exemplified by results in which the chloroplast data never

Figure 3. 50% majority-rule consensus of four most parsimonious trees generated from analysis of nuclear data only. Length = 754 steps, CI = 0.86, RI = 0.89. Numbers below branches are bootstrap values.

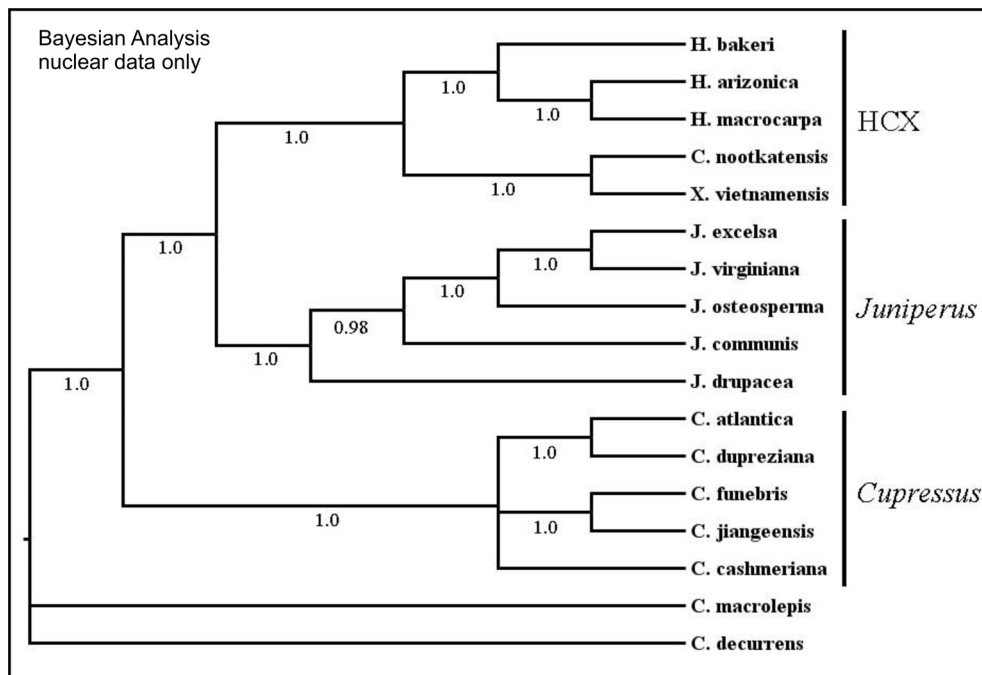
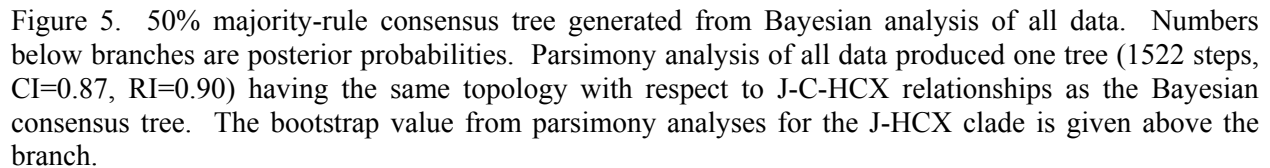


Figure 4. 50% majority-rule consensus tree generated from Bayesian analysis of nuclear data only. Numbers below branches are posterior probabilities.



Some authors have suggested *Cupressus* originated through hybridization between *Juniperus* and the common ancestor of HCX, an assertion based on conflict between different nuclear loci (Needly and Leafy vs. nrITS), similarity in *Cupressus* Needly and Leafy sequences to those of both *Juniperus* and HCX, and conflict between topologies derived from cpDNA (matK) and nuclear sequences (Yang et al., 2012). Results present here are different from those of previous studies in that we find little or no conflict among different nuclear partitions in recovering a well-supported C(J,HCX). In addition, we find little support for J-C-HCX relationships in the cpDNA data, although the cpDNA data never supports a J-HCX clade, and C(J,HCX) is excluded from the other two alternatives in two of three statistical comparisons of topology (Table 2). Collectively, these findings suggest that if conflict in topologies supported by nuclear and cpDNA data is attributable to ancient hybridization (Yang et al., 2012), then other processes producing ambiguity in the chloroplast data, or conflict between different nuclear genes (Yang et al., 2012), have also been important in the evolutionary history of the group.

We thank Jordan Britten for lab assistance. This research was funded by the Department of Biology and College of Arts and Sciences, Lamar University.

LITERATURE CITED

- Adams, R. P. 2014. Junipers of the world: The genus *Juniperus*, 4th edition. Trafford Publ., Victoria, BC.
- Adams, R. P., J. A. Bartel and R. A. Price. 2009. A new genus, *Hesperocyparis*, for the cypresses of the western hemisphere (Cupressaceae). Phytologia 91:160–185.
- Farjon, A. 2005. A monograph of Cupressaceae and Sciadopitys. Royal Bot. Gard., Kew, London.
- Farjon, A., N. T. Hiep, D. K. Harder, P. K. Loc and L. Averyanov. 2002. A new genus and species in the Cupressaceae (Coniferales) from northern Vietnam, *Xanthocyparis vietnamensis*. Novon 12:179–189.
- Gadek, P. A., D. L. Alpers, M. H. Heslewood and C. J. Quinn. 2000. Relationships within Cupressaceae *sensu lato*: a combined morphological and molecular approach. Am. J. Bot. 87:1044–1057.
- Goldman, N., J. P. Anderson and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49:652–670.
- Guindon, S. and O. Gascuel. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Syst. Biol. 52:696–704.
- Little, D. P. 2006. Evolution and circumscription of the true cypresses (Cupressaceae: *Cupressus*). Syst. Bot. 31:461–480.
- Little, D. P. 2005. Evolution and circumscription of the true cypresses (Cupressaceae: *Cupressus* and *Callitropsis*): A combined molecular approach. PhD diss. Cornell University, Ithaca, NY.
- Little, D. P., A. E. Schwarzbach, R. P. Adams and C. F. Hsieh. 2004. The circumscription and phylogenetic relationships of *Callitropsis* and the newly described genus *Xanthocyparis* (Cupressaceae). Am. J. Bot. 91:1871–1880.
- Mao, K., G. Hao, J. Liu, R. P. Adams and R. I. Milne. 2010. Diversification and biogeography of *Juniperus* (Cupressaceae): variable diversification rates and multiple intercontinental dispersals. New Phytol. 188:254–272.
- Posada, D. 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25:1253–1256.
- Rambaut, A. 2002. Se-Al: Sequence Alignment Editor. Available from <http://tree.bio.ed.ac.uk/people/arambaut/>, University of Oxford, Oxford.
- Ronquist, F. and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.
- Schmidt, H. A., K. Strimmer, M. Vingron and A. von Haeseler. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18:502–504.
- Shimodaira, H. and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16:1114–1116.
- Strimmer, K. and A. Rambaut. 2002. Inferring confidence sets of possibly misspecified gene trees. P. R. Soc. London B 269:137–142.
- Swofford, D. L. 2002. PAUP*: Phylogenetic analysis using parsimony(* and other methods), version 4b10. Sinauer Associates, Sunderland, Mass.
- Terry, R. G., J. A. Bartel and R. P. Adams. 2012. Phylogenetic relationships among the New World cypresses (*Hesperocyparis*; Cupressaceae): evidence from noncoding chloroplast DNA sequences. Plant Syst. Evol. 298:1987–2000.
- Wolf, C. B. 1948. The New World cypresses. Aliso 1:1–250.
- Yang, Z. Y., J. H. Ran and X. Q. Wang. 2012. Three genome-based phylogeny of Cupressaceae *s.l.*: further evidence for the evolution of gymnosperms and Southern Hemisphere biogeography. Mol. Phylogenet. Evol. 64:452–470.

Taxon	Voucher	Locus												
		mtLS	Needly	rbcd	trnK-matK	petB-petD	psbB	trnT-trnD	trnL-trnF	trnS-trnG	trnV-intron	psbD-trnT	trnC-trnD	rps4-trnS
C. macrolepis	Adams 10652, (KR 7315), Vietnam	AY380855.1	AY988269.1	HM024270.1	HM023982.1	HM024078.1	HM024175.1	HM024449.1	AY988171.1	HM024618.1	HM023887.1	*KP177874	*KP177865	HM024355.1
C. decurrens	Adams 10297, Oregon, USA	AY380854.1	AY988268.1	HM024269.1	HM023981.1	HM024077.1	HM024174.1	HM024448.1	AY988170.1	HM024617.1	HM023886.1	*KP177873	*KP177864	HM024354.1
C. atlantica	Adams 8429, Morocco	AY988367.1	AY988280.1	HM024275.1	HM023987.1	HM024083.1	HM024180.1	HM008344.1	AY988182.1	JO740558.1	HM023892.1	JO740555.1	JO740511.1	HM024360.1
C. dupreziana	Adams 8432, (ex Hillier Gardens), Alger	AY988375.1	AY988290.1	AY988243.1	AY988342.1	NA	NA	NA	AY988191.1	JO740559.1	NA	JO740536.1	JO740512.1	NA
C. finchris	Adams 8139, (KR ex Hubei), China	AY988377.1	AY988292.1	AY988245.1	HM023991.1	HM024087.1	HM024184.1	HM008329.1	AY988194.1	HM008346.1	HM023896.1	*KP177879	*KP177871	HM024364.1
C. jiangensis	Adams 9300, (ex Wang 026A), Tibet	AY988382.1	AY988298.1	AY988249.1	HM023993.1	HM024089.1	HM024186.1	NA	AY988199.1	HM008348.1	HM023898.1	*KP177880	*KP177872	HM024366.1
C. cashmeriana	Adams 8125, (ex Hillier), England	AY988372.1	AY988286.1	AY988240.1	HM023988.1	HM024084.1	HM024181.1	NA	AY988187.1	NA	HM023893.1	*KP177881	*KP177866	HM024361.1
C. nootkatensis	Adams 9086, Washington, USA	KJ849660.1	AY988304.1	HM024268.1	HM023980.1	HM024076.1	HM024173.1	HM024531.1	AY988207.1	JO740538.1	HM023885.1	JO740514.1	JO740490	HM024353.1
X. vietnamensis	Adams 10142, Vietnam	AY380877.1	AY988329.1	HM024352.1	HM024074.1	HM024170.1	HM024267.1	HM008343.1	AY988229.1	JO740539.1	HM023979.1	JO740515.1	JO740491	HM024447.1
H. bakeri	Adams 9362, California, USA	AY988369.1	AY988283.1	AY988237.1	HM023999.1	HM024095.1	HM024192.1	HM008340.1	AY988184.1	JO740540.1	HM023904.1	JO740516.1	JO740492	HM024372.1
H. arizonica	Adams 9378, Arizona, USA	U77962.1	AY988278.1	AY988233.1	HM023998.1	HM024094.1	HM024191.1	HM024456.1	AY988181.1	JO740552.1	HM023903.1	JO740529.1	JO740505	HM024371.1
H. macrocarpa	Adams 11460, California, USA	KJ849658.1	AY988301.1	HM024284.1	HM024005.1	HM024101.1	HM024198.1	HM024462.1	AY988204.1	HM024631.1	HM023910.1	JO740520.1	JO740496	HM024378.1
J. communis	Adams 7846, Sweden	AY988396.1	AY988314.1	HM024297.1	HM024019.1	HM024115.1	HM024212.1	HM024476.1	GO301207.1	HM024645.1	HM023924.1	*KP177875	*KP177867	HM024392.1
J. excelsa	HM001195.1	NA	HM024303.1	HM024025.1	HM024121.1	HM024218.1	HM024218.1	HM024482.1	HM024605.1	HM024565.1	HM023930.1	*KP177877	*KP177869	HM024398.1
J. virginiana	Terry 139, Texas, USA	EU277699.1	NA	HM024343.1	HM024065.1	HM024161.1	HM024258.1	HM024522.1	HM024601.1	HM024691.1	HM023970.1	*KP177878	*KP177870	HM024438.1
J. osteosperma	Terry 058, Utah, USA	AY988400.1	AY988320.1	HM024318.1	HM024040.1	HM024136.1	HM024233.1	HM024497.1	HM024580.1	GU223325.1	HM023945.1	JO740519.1	JO740495.1	HM024413.1
J. drupacea	Adams 8795, Greece	AY380872.1	AY988317.1	HM024301.1	HM024023.1	HM024119.1	HM024216.1	HM024480.1	JF950995.1	HM024649.1	HM023928.1	*KP177876	*KP177868	HM024396.1

Table 1. Voucher and GenBank accession data for taxa used in this study. Accession numbers with an asterisk (*) are for sequences published in this study.

Hypothesis	logL	Difference	Kishino-Hasegawa	Shimodaira-Hasegawa	Likelihood Weight
C(J,HCX)	-18669.9	18.2	0.05 (-)	0.06 (+)	0.03 (-)
J(C,HCX)	-18653.1	1.4	0.30 (+)	0.61 (+)	0.32 (+)
HCX(J,C)	-18651.7	Best	1.00 (+)	1.00 (+)	0.65 (+)

Table 2. Results from maximum likelihood analysis testing the fit of the three possible J-C-HCX topologies to the chloroplast data. p-values are given under the test name and + indicates inclusion in the confidence set.