Journal of Plant Biology, August 2006, 49(4) : 331-335

SHORT COMMUNICATION

A Role for Evolutionary Predictions in Gene Isolation and Characterization Studies

Todd James Barkman*
Department of Biological Sciences, Western Michigan University, Kalamazoo, MI 49008, USA

Studies of the evolutionary history of gene families can provide important insight into the processes leading to functional conservation or divergence among orthologs and paralogs. While the pattern of relationships among family members may be difficult to predict prior to an analysis, there are several expectations that should be met that can provide evolutionary verifications for gene isolation and characterization studies. In this report, it is shown that a recently described jasmonic acid caboxyl methyltransferase (JMT) sequence from *Capsicum annuum* L. (hot pepper) fails to satisfy several evolutionary predictions. Specifically, the hot pepper sequence is more closely related to some *Arabidopsis thaliana* (L.) Heynh. strains than are other strains of *A. thaliana* or its close relative *Brassica campestris* L. In addition, the level of divergence of hot pepper and *A. thaliana* JMT is more than ten times lower than expected from the levels of divergence of their mitochondrial and plastid sequences. Finally, attempts to PCR amplify JMT from hot pepper DNA were unsuccessful. In light of these empirical results, the identity of the putative hot pepper JMT sequence is called into question and highlights the need for phylogenetic perspectives even for functional studies.

Keywords: jasmonic acid carboxyl methyltransferase, multigene family evolution, phylogenetic analysis

The evolutionary history of multigene families can be quite complex due to the pervasive processes of family member birth and death, recombination, and gene conversion (Sanderson and Doyle, 1992; Hillis, 1994; Clegg et al., 1997; Zhang 2003). Yet, although orthology/paralogy relationships may be hard to predict at the onset of a study, there are clear predictions that can be expected for any gene family under investigation. First, orthologous sequences isolated from different strains or populations of a single species should be monophyletic to the exclusion of orthologs from other more distantly related species, and especially genera, although see Igic and Kohn (2001) for a rare exception. This expectation becomes more likely given increasing amounts of time since divergence from the common ancestor of any two species. Second, if orthologs are available for more than two species then it is expected that the relationships among them should reflect the phylogenetic history of the organisms. That is, if two species are known to belong to the same lineage (or plant family), then a third unrelated species should appear more distantly related based on phylogenetic analyses of any orthologous sequence isolated from its genome. Third, for any two plant species, there are clear predictions related to the number of substitutions per synonymous site (K_S) estimated among mitochondrial, plastid, and nuclear sequences. Specifically, mitochondrial K_S estimates are predicted to be at least three times lower than those of the plastid, which in turn are predicted to be at least two times lower than nuclear K_S estimates (Wolfe et al., 1987), excepting a few unusual cases (Adams et al., 2002; Cho et al., 2004).

As an example of the use of the preceding evolutionary predictions, jasmonic acid carboxyl methyltransferases (JMT) sequences were analyzed below. JMT is involved in the production of methyl jasmonate which is a volatile signal released in response to insect herbivory (Cheong and Choi, 2003). JMT is a member of the SABATH gene family of methyltransferases that includes genes responsible for the production of various secondary chemicals including methyl jasmonate, methyl salicylate and methyl benzoate (D'Auria et al., 2003). The first fully characterized JMT sequence was obtained from *Arabidopsis thaliana* (L.) Heynh. (Seo et al., 2001). Recently, the isolation of JMT from *Capsicum annuum* L. (hot pepper) using RT-PCR was reported (Song et al., 2005). Song et al. (2005) compared amino acid alignments of the *C. annuum* sequence they isolated to *A. thaliana* JMT and a *Brassica campestris* L. putative JMT sequence (Song et al., 2000). It was reported that the *A. thaliana* and *C. annuum* sequences share 92% deduced amino acid identity, which is surprising given that these two species are only very distantly related (The APG, 2003). Here evidence is provided that reveals that the *C. annuum* JMT sequence fails to satisfy several evolutionary predictions thereby casting doubt on the validity of the sequence reported by Song et al. (2005).

MATERIALS AND METHODS

Phylogenetic Analyses

A wide diversity of functionally characterized SABATH gene family members were chosen for phylogenetic analysis including JMT, benzoic acid/salicylic acid methyltransferase (BSMT), benzoic acid methyltransferase (BAMT), indole acetic acid methyltransferase (IAMT), and salicylic acid methyltransferase (SAMT) (D'Auria et al., 2003). BLAST was used to obtain multiple putative JMT sequences from *A. thaliana* from GenBank. GenBank accession numbers for all sequences analyzed are shown next to each species in Figure 1. Prior to phylogenetic analysis, DNA sequences were aligned using Clustal X (Thompson et al., 1997) and were

*Corresponding author; fax +1-269-387-5609
e-mail todd.barkman@wmich.edu

Abbreviations: JMT, jasmonic acid carboxyl methyltransferase; Ks, number of substitutions per synonymous site; mt, mitochondrial; pt, plastid

subsequently adjusted by eye. PAUP*4.0 (Swofford, 2003) was used for phylogenetic estimates using neighbor joining and parsimony. Neighbor-joining tree estimates were performed using the optimal nucleotide substitution model estimated using Modeltest 3.06 (Posada and Crandall, 2001). Unweighted parsimony analyses were performed with 100 random addition sequences during heuristic searches. To evaluate node-specific support with both methods, bootstrap proportions (Felsenstein, 1985) were estimated with 1000 replicates. DNAsp 3.53 (Rozas and Rozas, 1999) was used to estimate the number of substitutions per synonymous site (K_S).

DNA Isolation and PCR

To isolate genomic JMT sequences, DNA was extracted from fresh leaves of *C. annuum* L. and *A. thaliana* (L.) Heynh. After standard DNA extraction procedures (Doyle and Doyle, 1987), polymerase chain reaction (PCR) was performed to amplify JMT using three pairs of gene specific primers: JMT 1: Forward, 5′-GGTCCGAACAGTCTCTTGTCCATC-3′ and Reverse 5′-AACACGATTCTCCTCCTCCACG-3′, JMT 2: Forward, 5′-CGGAGAAGGAAGACAGGACAATAAC-3′ and Reverse 5′-GGAACGATAAAACCATTCGGC-3′, JMT 3: Forward, 5′-GCTCCGAAGAGTTGAAAATGGTG-3′ and Reverse 5′-AACAATAGCGTATCGTGGCGAG-3′. These primer pairs were designed to amplify three separate regions of JMT that correspond to portions of exons 2 to 4 of *C. annuum* (DQ222856), respectively. Mitochondrial *coxIII* (Forward: 5′-CCGTAGGAGGTGTGATGT-3′ and Reverse: 5′-CTCCCCACCAATAGATAGAG-3′), *trnL* (using primers E and F; Taberlet et al., 1991), and SAMT (using F6 and R3; Martins and Barkman, 2005) were used as positive controls to ensure that the extracted DNAs were amplifiable. Each 50 μl reaction contained 0.1 μl Platinum Taq DNA polymerase (Invitrogen, USA), 1 μl DNA (ca. 0.5 μg), 1 μl of each primer (10 μM), 2 μl of nucleotide mix (10 mM), 5 μl 10X reaction buffer, 2.5 μl $MgCl_2$ (50 mM) and 37.4 μl distilled water. Cycling conditions included an initial denaturation for 3 min at 94°C followed by 35 cycles of denaturation for 1 min at 94°C, annealing for 1 min at 50°C and extension for 30 sec at 72°C. The annealing temperature chosen for these primer pairs is 2 to 6°C below the optimal temperature predicted for each pair of primers against *C. annuum* JMT. The cycling ended with a final extension at 72°C for 30 min. DNA sequencing followed standard DNA sequencing protocols (Martins and Barkman, 2005).

RESULTS AND DISCUSSION

Figure 1 shows the results of a neighbor-joining phylogenetic analysis of SABATH gene family members (minimum evolution score = 4.81). The parsimony consensus tree (15 equal length trees, 2839 steps) was very similar to the NJ tree particularly with respect to the placement of the JMT sequences and is therefore not shown. Based on the results shown in Figure 1, it appears that several putative JMT accessions from *A. thaliana* available on GenBank are not monophyletic (Fig. 1). Surprisingly, it appears that the *C. annuum* sequence (Song et al., 2005) is more closely related to some sequences of *A. thaliana* than are other sequences from the same species. Not only is the placement of *C. annuum* with *A. thaliana* strongly supported by high bootstrap values (85/100), fairly high levels of support place it embedded within the species lineage as well (BP = 93 in the parsimony tree). This is an unexpected result because it is predicted that these *A. thaliana* orthologues should be monophyletic to the exclusion of any other species, including *C. annuum*. Second, phylogenetic analyses of the reported *C. annuum* sequence clearly support its relationship to *A. thaliana* JMT with high statistical support to the exclusion of the orthologous sequence from *Brassica*. Because *A. thaliana* and *Brassica* are closely related and classified in the same family (Brassicaceae) whereas *C. annuum* (Solanaceae) is only distantly related to either of them (The APG, 2003), this is a highly unexpected result. Third, the K_S obtained from comparing the mitochondrial encoded *coxIII* of *C. annuum* (DQ855008) to *A. thaliana* (NC000932) was 0.09 while the K_S obtained from comparing the plastid encoded *rbcL* of *C. annuum* (U08610) to *A. thaliana* (NC001284) was 0.38. Thus, according to expectations, the level of K_S was at least three times lower for *coxIII* as compared to *rbcL*. The estimated K_S from a comparison of the nuclear encoded JMT sequences of *C. annuum* and *A. thaliana* is only 0.03. This value is in sharp contrast to what was predicted. According to relative intergenomic synonymous subtitution expections, K_S for the two nuclear sequences should have been approximately 0.54 to 0.76 (an order of magnitude higher).

Based on these molecular evolutionary analyses it seems likely that the *C. annuum* JMT sequence is contamination by either *A. thaliana* or a close relative such as *A. lyrata*. To determine if the previously reported JMT sequence from *C. annuum* may have been a contaminant, PCR was attempted in order to isolate JMT from *C. annuum* using three pairs of primers designed from the published sequence (Song et al., 2005). Because these primers perfectly matched the putative *C. annuum* sequence, it was expected that they would result in amplification of JMT. However, these three primer pairs only amplified JMT from *A. thaliana* (Fig. 2A). DNA sequencing revealed that JMT 1 (DQ855005), 2 (DQ855006) and 3 (DQ855007) from *A. thaliana* (Fig. 2A) were all identical to a sequence from *A. thaliana* strain Col-0 (BX813439) available on GenBank. Surprisingly, JMT 1 and 3 from *A. thaliana* (Fig. 2A) were also identical to *C. annuum* (DQ222856). However, JMT 2 from *A. thaliana* differed from the *C. annuum* sequence at four nucleotide positions. To verify that the lack of JMT PCR amplification from *C. annuum* was not due to poor DNA quality, we used several positive control primers designed to amplify mt, pt and nuclear sequences. As shown in Figure 2B, *coxIII*, and *trnL* amplified from both *A. thaliana* and *C. annuum* indicating that the hot pepper DNA is amplifiable. Finally, SAMT was amplified from hot pepper (DQ855009) using primers that correspond to exon 3 of that gene. The obtained product (Fig. 2B) matched *C. annuum* SAMT from GenBank (AY741493) in this region thereby confirming the identity of this PCR product. The amplification of SAMT from hot pepper indicates that other SABATH gene family members could be amplified from the extracted *C. annuum* DNA (Fig. 2B). These SAMT primers were designed to be specific to Solanaceae so the lack of amplification from *A. thaliana* is not unexpected (Fig. 2B). The PCR products obtained using these "positive control" primers indicates that the *C. annuum* DNA is amplifiable and the lack of amplification of *C. annuum* JMT calls into question the origin of the reported sequence of Song et al. (2005).

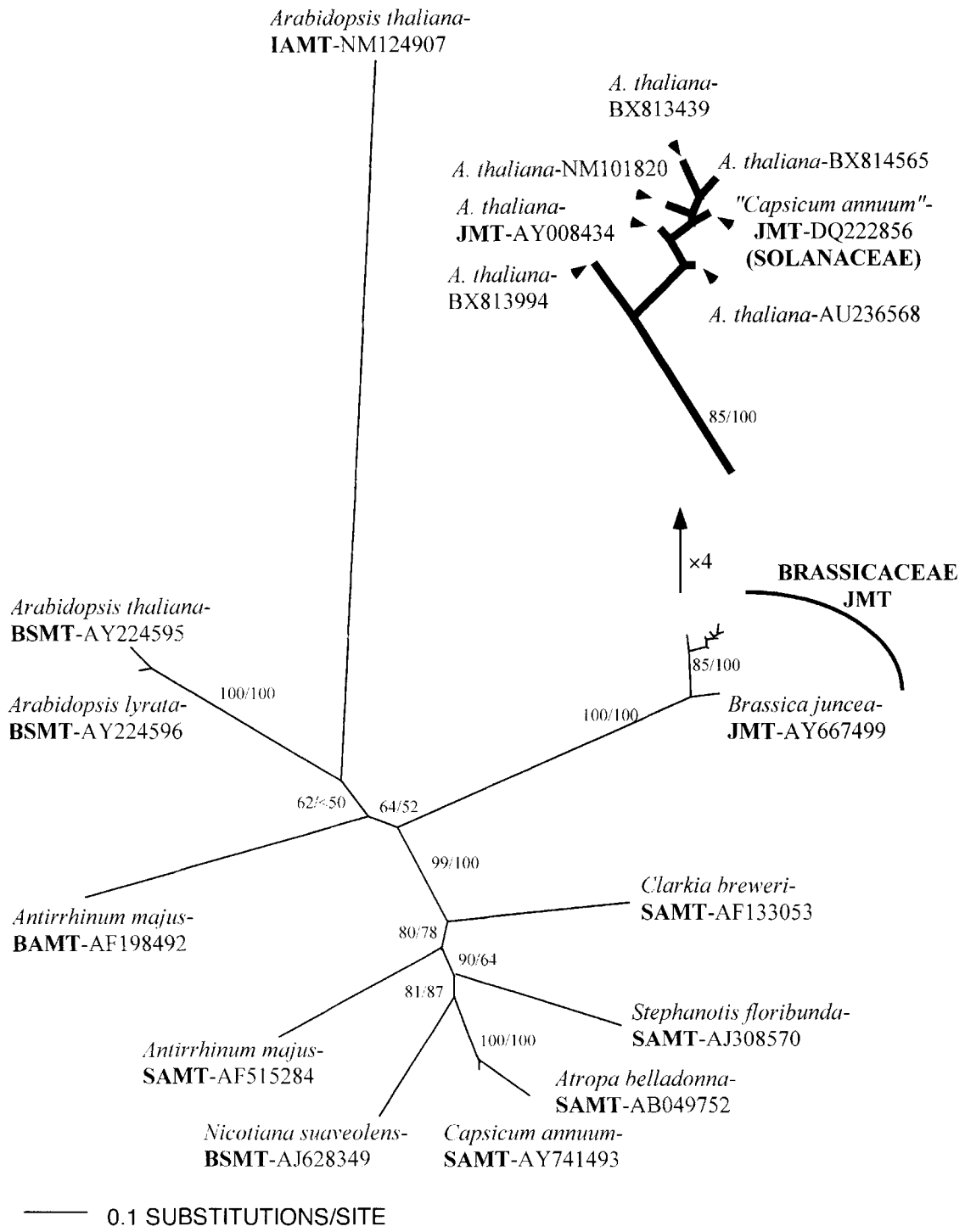

Figure 1. Neighbor-joining tree showing phylogenetic relationships among SABATH gene family members. Levels of node-specific support > 50 are shown for neighbor-joining/parsimony, respectively. GenBank accession numbers are listed next to each sequence analyzed. JMT, jasmonic acid methyltransferase; SAMT, salicylic acid methyltransferase; BSMT, benzoic acid/salicylic acid methyltransferase; BAMT, benzoic acid methyltransferase; IAMT, indole acetic acid methyltransferase. The phylogenetic position of *C. annuum* JMT nested within the *A. thaliana* lineage is highly unexpected. Branch lengths are proportional to amount of estimated nucleotide change.

Considered together, the molecular evolutionary and PCR results indicate that the reported isolation of JMT from *C. annuum* (Song et al., 2005) may instead be due to contamination. Although Song et al. (2005) were able to detect bands using Southern blots that hybridized to the JMT probe, it is possible that they represent other *C. annuum* SABATH gene family members (Fig. 3; Song et al., 2005), especially given that *A. thaliana* JMT and *C. annuum* SAMT are ca. 60% identical at the nucleotide level. Likewise, the significance of Southern blots of *C. annuum* RT-PCR products obtained after wounding is difficult to interpret (Fig. 4, 5, and 6; Song et al., 2005). Had it been shown that methyl jasmonate is produced in *C. annuum* leaves in response to wounding, the expression results would correlate with secondary chemical production. However,

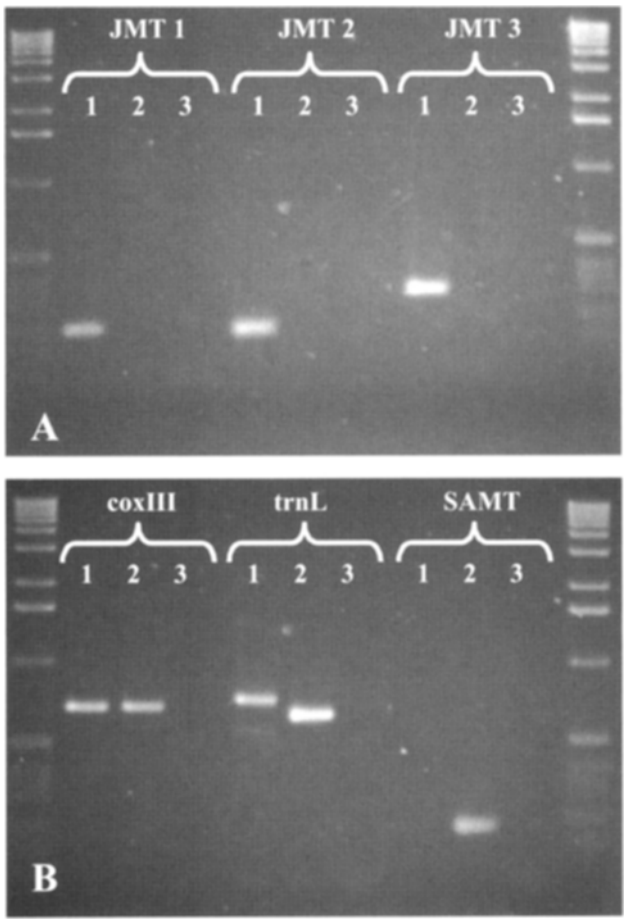

Figure 2. A. Agarose gel showing PCR products obtained from three non-overlapping primer pairs designed from *C. annuum* JMT. Although the three primer pairs were designed from the published *C. annuum* JMT sequence, amplification only was obtained from *A. thaliana*. **B.** Agarose gel showing PCR products obtained from mitochondrial *coxIII*, plastid *trnL*, and nuclear SAMT. Amplification and sequencing of *coxIII*, *trnL*, and *SAMT* from *C. annuum* demonstrates that the DNA used is correctly identified and amplifiable. 1, *A. thaliana*; 2, *C. annuum*; 3, negative control (no DNA). A 1 kb ladder is shown on the outermost lanes of each gel.

because other volatiles like methyl salicylate produced by other SABATH gene family members are known to be released from hot pepper insect wounded leaves (van Den Boom et al., 2004), the blot results do not reject the possibility that SAMT, rather than JMT, is upregulated in the *C. annuum* leaves (Song et al., 2005).

Because considerable nucleotide diversity may exist among different strains of plant species, a phylogenetic perspective is useful to assess the significance of observed sequence differences and the implications of estimated relationships (Baum et al., 2005). Not only can the phylogenetic position of a newly isolated sequence be assessed in terms of expectations related to its function, the number of differences between sequences can also be interpreted in terms of intergenomic expectations (Wolfe et al., 1987). The incorporation of an evolutionary perspective into gene isolation and characterization studies will be critical to provide insight into complex gene family evolution especially as more studies are focused on non-model organisms.

ACKNOWLEDGEMENTS

Don Kane, Bruce Bejeck, Wendy Ransom-Hodgkins, Talline Martins, and two anonymous reviewers are thanked for providing feedback on earlier versions of this manuscript. This study was supported by NSF grant DEB 03 44496.

Received March 28, 2006; accepted June 30, 2006.

LITERATURE CITED

Adams KL, Qiu Y-L, Stoutemyer M, Palmer JD (2002) Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc Natl Acad Sci USA 99: 9905-9912

Baum DA, Smith SD, Donovan SS (2005) The tree-thinking challenge. Science 310: 979-980

Cheong JJ, Choi YD (2003) Methyl jasmonate as a vital substance. Trends Genet 19: 409-413

Cho Y, Mower JP, Qiu Y-L, Palmer JD (2004) Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. Proc Natl Acad Sci USA 101: 17741-17746

Clegg MT, Cummings MP, Durbin ML (1997) The evolution of plant nuclear genes. Proc Natl Acad Sci USA 94: 7791-7798

D'Auria JC, Chen F, Pichersky E (2003) The SABATH family of methyltransferases in *Arabidopsis thaliana* and other plant species. Rec Adv Phytochem 37: 253-283

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bulletin 19: 11-15

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783-791

Hillis DM (1994) Homology in molecular biology, *In* BK Hall, ed, Homology: The Hierarchical Basis of Comparative Biology. Academic Press, San Diego, pp 339-368

Igic B, Kohn J (2001) Evolutionary relationships among self-incompatibility RNases. Proc Natl Acad Sci USA 98: 13167-13171

Martins TR, Barkman TJ (2005) Reconstruction of Solanaceae phylogeny using the nuclear gene SAMT. Syst Bot 30: 435-447

Posada D, Crandall KA (2001) Selecting the best-fit model of nucleotide substitution. Syst Biol 50: 580-601

Rozas J, Rozas R (1999) DnaSP version 3: An integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15: 174-175

Sanderson M, Doyle JJ (1992) Reconstruction of organismal and gene phylogenies from data on multigene families: Concerted evolution, homoplasy, and confidence. Syst Biol 41: 4-17

Seo HS, Song JT, Cheong JJ, Lee YH, Lee YW, Hwang I, Lee JS (2001) Jasmonic acid carboxyl methyltransferase: A key enzyme for jasmonate-regulated plant responses. Proc Natl Acad Sci USA 98: 4788-4793

Song JT, Seo HS, Song SL, Lee JS, Choi YD (2000) NTR1 encodes a floral nectary specific gene in *Brassica campestris* L. ssp. pekinensis. Plant Mol Biol 42: 647-655

Song MS, Kim DG, Lee SH (2005) Isolation and characterization of a jasmonic acid carboxyl methyltransferase gene from hot pepper (*Capsicum annuum* L.). J Plant Biol 48: 292-297

Swofford DL (2003) PAUP*: Phylogenetic analysis using parsimony (*and other methods) Version 4. Sinauer Associates, Sunderland

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17: 1105-1109

The APG (2003) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants:

APG II. Bot J Linn Soc 141: 399-436
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 24: 4876-4882
van Den Boom CEM, van Beek TA, Posthumus MA, de Groot A, Dicke M (2004) Qualitative and quantitative variation among volatile profiles induced by *Tetranychus urticae* feeding on plants from various families. J Chem Ecol 30: 69-89
Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84: 9054-9058
Zhang J (2003) Evolution by gene duplication-an update. TREE 18: 292-298