Molecular Phylogenetics and Evolution xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand

Simon Joly [a,*], Peter B. Heenan [b], Peter J. Lockhart [a,c]

[a] *Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, New Zealand*
[b] *Allan Herbarium, Landcare Research, Lincoln, New Zealand*
[c] *Institute of Molecular Biosciences, Massey University, New Zealand*

ARTICLE INFO

Article history:
Received 6 September 2008
Revised 17 February 2009
Accepted 23 February 2009
Available online xxxx

Keywords:
Adaptive radiation
Allopolyploidy
Divergence time estimates
Hybridization
Pachycladon

ABSTRACT

The Southern Alps in New Zealand contain many herbaceous plant groups that have radiated during the Plicoene–Pleistocene. The species in these genera tend to be polyploid relative to their overseas close relatives, an observation of much interest given that hybridization and allopolyploidy have recently been suggested as a possible stimulus for adaptive radiation. We were interested to determine whether or not allopollyploidy was a feature of *Pachycladon*, a genus which is hypothesised to have adaptively diversified onto different geological substrates in the mountains of the South Island of New Zealand. Phylogenetic analyses of five single-copy nuclear genes show that *Pachycladon* species have two copies of each gene representing two highly diverged evolutionary lineages from the Brassicaceae. Molecular clock analyses of all loci suggest that the two genome copies in *Pachycladon* diverged 8 million years ago, and that the allopolyploid origin of the genus occurred during the Pleistocene between 1.6 and 0.8 million years ago. This hybridization event at the origin of the *Pachycladon* radiation is perhaps the most extreme example yet reported of successful hybridization between distantly related parents.

1. Introduction

Hybridization is an important feature of plant and animal evolution (Anderson, 1949; Arnold, 1997; Barton, 2001; Grant, 1981; Rieseberg, 1997; Stebbins, 1959). It is increasingly accepted as a process for generating biotic diversity (Arnold, 1997; Barrier et al., 1999; Ferguson and Sang, 2001; Rieseberg et al., 2003) and for promoting rapid adaptation (Ellstrand and Schierenbeck, 2000; Grant and Grant, 1996; Lewontin and Birch, 1966). It has also been suggested that hybridization might have acted as a trigger for the adaptive radiation of some genera as both processes require similar ecological requirements such as the presence of novel habitats (Seehausen, 2004). There are currently few well characterised examples that support the latter role for hybridization in evolution apart from the demonstration that an allopolyploid event preceded the adaptive radiation of the Hawaiian silversword alliance (Barrier et al., 1999). Allopolyploidy, which results in the presence of more than two complete sets of chromosomes from distinct species in one individual, is indeed one means of stabilizing hybrid advantage and is a prevalent evolutionary process of vascular plants (Otto and Whitton, 2000; Soltis and Soltis, 2000). Hybridization and allopolyploidy have been suspected as being important for the evolution of certain floras such as that of New Zealand (Hair, 1966), which has produced many Late tertiary species radiations (Winkworth et al., 2005).

Pachycladon (Brassicaceae) in New Zealand is one such example of a polyploid genus in which hybridization might have been important. It consists of eight small perennial species that are endemic to the South Island of New Zealand, and an additional species that is endemic to Tasmania, Australia. *Pachycladon* is said to have diversified within the last 1–3.5 million years (Heenan et al., 2002), with adaptation to different geological substrates suggested as being an important abiotic driver (Heenan and Mitchell, 2003). It harbours considerable inter-specific morphological variation to an extent that species of the genus were previously thought to belong to three different genera that were associated to different tribes or sub-tribes (Mitchell and Heenan, 2000). Some *Pachycladon* species are restricted in altitude and to different soil types (schist or greywacke) and variation in habitat is correlated with differences in reproductive strategy (monocarpic or polycarpic), inflorescence position (lateral or terminal), siliques with seeds uniseriate or biseriate, and seed morphology (with or without wings) (Heenan and Mitchell, 2003). Ongoing work using microarray and transcriptome sequencing is currently investigating potential biotic and abiotic drivers of the *Pachycladon* radiation (e.g., Collins et al., 2008; Voelckel et al., 2008).

The suggestion that *Pachycladon* ($2n = 20$) is of polyploid origin comes from FISH chromosome painting experiments (Lysak, unpublished data) and the presence of duplicated nuclear genes (Collins et al., 2008; McBreen and Heenan, 2006). However, the

* Corresponding author. Address: Biology Department, McGill University, 1205 Docteur Penfield, Montreal, Que., Canada H3A 1B1.
E-mail address: simon.joly@mail.mcgill.ca (S. Joly).

1055-7903/$ - see front matter Crown Copyright © 2009 Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.02.015

polyploid nature of the genus has yet to be clearly demonstrated and it is still unknown whether hybridization was involved in the origin of the genus. Here, we conduct phylogenetic analyses of five single-copy nuclear genes and report that *Pachycladon* is an allopolyploid with very distantly related parents.

2. Methods

Five single-copy nuclear genes were investigated that were almost all located in distant regions of the *Arabidopsis thaliana* genome (TAIR version 7; Fig. 1) and on different blocks of the reconstructed ancestral Brassicaceae karyotype (Schranz et al., 2006): chalcone synthase (*CHS*, Ancestral Karyotype Block R [AKB-R]), cinnamyl alcohol dehydrogenase 5 (*CAD5*, AKB-U), malate synthase (*MS*, AKB-R), phosphoribulokinase (*PRK*, AKB-B), and a nodulation gene from the MtN21 gene family (*MtN21*, AKB-J). The chloroplast *rbcL* gene was also analysed to reconstruct the maternal history of the genus.

Pachycladon exilis and *Pachycladon fastigiata*, which represents the oldest dichotomy in the *Pachycladon* phylogeny (Heenan and Mitchell, 2003), were sequenced for all markers but additional species were sequenced for *CHS*, *MS*, and *PRK*. We attempted to sequence nine Brassicaceae species that cover several lineages in the family according to previous studies (Bailey et al., 2006; Heenan et al., 2002; Koch et al., 2007), but other sequences available in GenBank were also included in phylogenetic analyses (Supplementary material Appendix 1).

2.1. Molecular methods

DNA was extracted using the plant DNeasy kit (QiaGen) and the genes were amplified using standard PCR conditions (see Joly et al., 2006). Published primers were used to amplify *CHS* (Koch et al., 2000), *MS* (Lewis and Doyle, 2001), and *PRK* (Lewis and Doyle, 2002). *CAD5* was amplified using primers CAD5-P2F (5′-CTGCCACACCGATCTTCATCAAACT-3′) and CAD5-P4R (5′-TAGTATRCCTCCTCTTAGGCCTGG-3′) and an annealing temperature of 50 °C, whereas *MtN21* was amplified using primers MtN21-1F (5′-GCCATTACGTTCTTGTTGCATACCG-3′) and MtN21-5R (5′-TGTCAAAACCGATGTTCAAAGCAG-3′) and an annealing temperature of 54 °C. All individuals of *Pachycladon* possessed two divergent copies of each gene, which allowed primers specific to each copy to be developed (Supplementary material Appendix 2). Most *Pachycladon* individuals were amplified and sequenced using these specific primers.

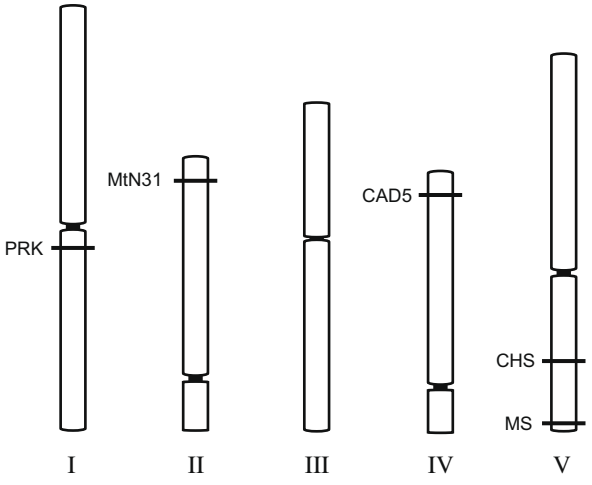

Fig. 1. Position of the nuclear markers used in this study on the *Arabidopsis thaliana* chromosomes.

The PCR products were purified with polyethylene glycol (Joly et al., 2006). Sequencing reactions used the BigDye terminator chemistry (version 3.3; Applied Biosystems) and were run on a 3730 automated sequencer (Applied Biosystems). Polymorphic PCR products were cloned with a TOPO-TA kit (Invitrogen). Colonies were screened by PCR (T3–T7 primers) and five positive clones were typically sequenced from the PCR product using T3 and T7 primers. Sequences were edited in Sequencher (version 4.7; GeneCodes) and aligned in ClustalX (Thompson et al., 1997). Because introns were too variable to be aligned with confidence across the Brassicaceae, they were removed from all genes prior to the phylogenetic analyses.

For the *rbc*L gene, we identified available Brassicaceae sequences by blasting a *P. fastigiata* sequence (GenBank Accession No. EF015666) to the nr database of GenBank using megablast. The best 100 hits were kept and then aligned in ClustalX.

2.2. Phylogenetic analyses

Every marker was tested for evidence of recombination using the Φ statistic (Bruen et al., 2006) implemented in PhiPack (Bruen, 2005), with 10,000 permutations and using a window size of 100. Maximum likelihood (ML) phylogenetic analyses were performed on the complete matrices for each gene (i.e., containing all clone sequences). The tree search was undertaken with PhyML version 2.4.4 (Guindon and Gascuel, 2003) using the best fitting substitution model according to a Akaike Information Criterion (AIC) evaluated on a neighbour joining tree (ModelTest version 3.7, Posada and Crandall, 1998). This preliminary analysis was used to filter sequences likely to contain *Taq* polymerase induced errors. Only one sequence was kept where several clone sequences were assumed to come from a single gene copy: sequences with more than 99% sequence identity were considered to be the result of polymerase errors. The clone sequences left in the final matrices were the ones with the most basal position on the ML tree and/or with the shortest branch, consistent with it containing a minimal number of *Taq* induced errors. Sequences that contained uncommon amino acid substitutions, compared to the other clones and to the whole sequence alignment, were also considered to result from *Taq*-induced errors and were removed.

The filtered matrices (TreeBase study accession number S2261) were subjected to ML and Bayesian phylogenetic analyses. The ML tree search was performed in PhyML using the best model selected according to the AIC criterion in ModelTest. Branch support was estimated using 1000 bootstrap replicates. Bayesian phylogenetic analyses were performed in BEAST version 1.4.7 (Drummond and Rambaut, 2007) using the model available in BEAST that received the best AIC score in ModelTest. Each matrix was analysed with an uncorrelated lognormal clock (Drummond et al., 2006) fixing the mean substitution rate to one, using a Yule prior for the tree, and using other default parameters. Two independent runs were performed for each dataset to ensure convergence of the chains. Convergence to the stationary distribution for all parameters and the collection of estimated sample sizes above 1×10^4 was confirmed using Tracer (Rambaut and Drummond, 2005). The chains were run for 1×10^7 generations and sampled every 1000 generations, discarding the first million generations as burnin (except for the *CHS* and *rbc*L genes that had a "run length/burnin" of 2×10^7/1.2×10^7 and 5×10^7/6×10^6 generations, respectively). The best root positions estimated in BEAST were used to root ML trees.

2.3. Divergence times estimates

The topologies of the ML trees obtained with the filtered matrices were used to test if the genes studied were evolving according to a molecular clock hypothesis with a likelihood ratio test (Felsen-

Please cite this article in press as: Joly, S., et al. A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. Mol. Phylogenet. Evol. (2009), doi:10.1016/j.ympev.2009.02.015

stein, 1981). ML scores of the data for the two alternative hypotheses with the appropriate evolutionary model were obtained from PAUP* version 4.10b (Swofford, 2002). A sequential Bonferroni correction was applied (Rice, 1989) to account for the multiple statistical tests performed.

Estimating divergence times in the Brassicaceae is a difficult task because of the rarity of available fossils to calibrate substitution rates. One source of information is the presence of *Rorippa* pollen deposits in geological samples from the Pliocene (2.5–5 Mya) (Mai, 1995). But because estimating divergence times on a phylogeny using methods that allows variation of rates among branches using few calibration points is challenging (Knapp et al., 2005; Soltis et al., 2002; Yang and Yoder, 2003) and because some methods are inaccurate when calibration points are located toward the tips of the phylogeny (Pérez-Losada et al., 2004) such as with the *Rorippa* pollen, these approaches were not pursued further here. Instead, a molecular clock was assumed and synonymous substitution rates were used for dating specific events on the phylogeny. Because the family is relatively recent (ca. 30 mya), the signal at synonymous sites should not be affected by saturation. To estimate the divergence time between taxa and groups, we used a range of possible rates of synonymous substitution of 1.9–4.1×10^{-8} synonymous substitutions per synonymous site (d_S) per year. This was calculated given that d_S values between *A. thaliana* and *Brassica* vary between 0.4 and 0.6 for most genes (estimated from 256 genes: Blanc et al., 2003) and that the divergence between these species is likely to have occurred between 14.5 and 21 mya according to two independent estimates (Koch et al., 2001; Yang et al., 1999). We feel that the present range of synonymous substitution rates is more informative than the values of 1.5×10^{-8} obtained by Koch et al. (2000) from a single gene (*CHS*) and of 7.3×10^{-9} (Lynch and Conery, 2000) obtained from tRNA genes and calibrated using the split between monocots and dicots. Indeed, calibrating a rate with an old event such as the divergence of the monocots and the dicots may underestimate the synonymous substitution rate due to signal saturation at such distances. Average numbers of synonymous substitution per synonymous sites were estimated independently for every gene using the maximum likelihood method of Goldman and Yang (1994) implemented in PAML (Yang, 2007). Given these numbers, divergence times (T) were then obtained given that the synonymous substitution rate = $d_S/2T$.

3. Results

The sequences generated for this study were deposited in GenBank under the Accession Nos. FJ645063–FJ645256 (see Supplementary material Appendix 1 for a list of all Accession Nos. used in this study). Table 1 lists some characteristics for each gene investigated in this study. There was no evidence of pseudogene sequences in any dataset after inspection of the translated amino acid sequences. There was also no evidence of recombination in the datasets analysed ($p > 0.05$). ML and Bayesian trees were always highly congruent and the only disagreements occurred at nodes that received very low support (bootstrap < 50% and posterior probability < 0.5).

3.1. Two genome copies

CHS phylogeny has been a focus of previous studies (Koch et al., 2000, 2001) and thus was the gene with the most thorough sampling across the family. The *CHS* phylogeny clearly illustrates the presence of two distant copies in the genome of *Pachycladon* (Fig. 2) that are associated with the two main evolutionary lineages in the Brassicaceae (Bailey et al., 2006; Koch et al., 2007), which we will henceforth call the *Arabidopsis* and the *Brassica* lineages (see Fig. 2). Phylogenies of the other nuclear genes revealed that *Pachycladon* possesses two copies of each (Fig. 3). The phylogenetic positions of the two genome copies were highly similar across gene trees (Figs. 2 and 3). Together, these observations support an allopolyploid origin for the genus. One of the genome copies present in *Pachycladon* was consistently in a derived position in the *Arabidopsis* lineage and it will henceforth be called the A genome copy (for *Arabidopsis*). This genome copy appears to be most closely associated with genera *Crucihimalaya*, *Transberingia*, and *Boechera*, although other genera such as *Capsella* and *Olimarabidopsis* are close to this copy for some of the genes. The phylogenetic position of the other genome copy, the B copy, was always close to the split of the *Arabidopsis* and *Brassica* lineages, although none of the species sampled had high genetic affinities with this copy.

3.2. Reconstructing the maternal evolutionary history of Pachycladon

The chloroplast *rbcL* gene was chosen to reconstruct the maternal history of *Pachycladon* because it could be aligned unambiguously within the Brassicaceae. We used a BLAST approach to sample as many Brassicaceae species as possible: 76 out of the first hundred results were Brassicaceae species and 24 of these represented identical sequences within species that were removed for the phylogenetic analysis. Although the phylogenetic analysis was performed on all sequences, we only present the portion of the tree containing Brassicaceae species (Fig. 4), which formed a monophyletic group with high support (99.6% bootstrap and 1.0 posterior probability). The *rbcL* tree shows that the chloroplast of *Pachycladon* branches at the base of the *Arabidopsis* lineage and that it is relatively distant from the genera *Transberingia* and *Crucihimalaya*, which is very similar to the B genome copy identified above with the nuclear genes. Given that chloroplasts are generally maternally inherited in angiosperms and also in Brassicaceae (e.g., Johannessen et al., 2005), this suggests that the B genome was the maternal parent of the allopolyploid event.

Table 1
Gene characteristics and likelihood ratio tests.

Gene	Length[a]	All sequences		Filtered matrix		ln$L$		$\delta = -2\log(L_0/L_1)$	df	p
		Taxa	Evolutionary model	Taxa	Evolutionary model	Molecular clock (L_0)	No clock (L_1)			
CAD5	376	19	TrN+Γ	14	GTR+G	−1250.62382	−1241.01793	19.21178	12	0.0835442
CHS	1167	64	SYM+I+Γ	49	SYM+I+G	−8466.34482	−8396.3419	140.0058	47	3.53×10^{-11}***
MS	515	51	SYM+Γ	22	SYM+I+G	−1760.644	−1744.92766	31.43268	20	0.0497314
MtN21	572	42	TrN+Γ	13	HKY+G	−2022.1267	−1996.81908	50.61524	11	0.0004851**
PRK	677	51	TrN+I+Γ	23	GTR+I+G	−2543.12135	−2520.1437	45.9553	21	0.0012954**
rbcL	1318	77	K81uf+I+Γ	NA	NA	NA		NA	NA	NA

Significance levels after sequential Bonferroni correction: ** $p < 0.01$, *** $p < 0.001$.
[a] Introns removed.

Please cite this article in press as: Joly, S., et al. A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. Mol. Phylogenet. Evol. (2009), doi:10.1016/j.ympev.2009.02.015

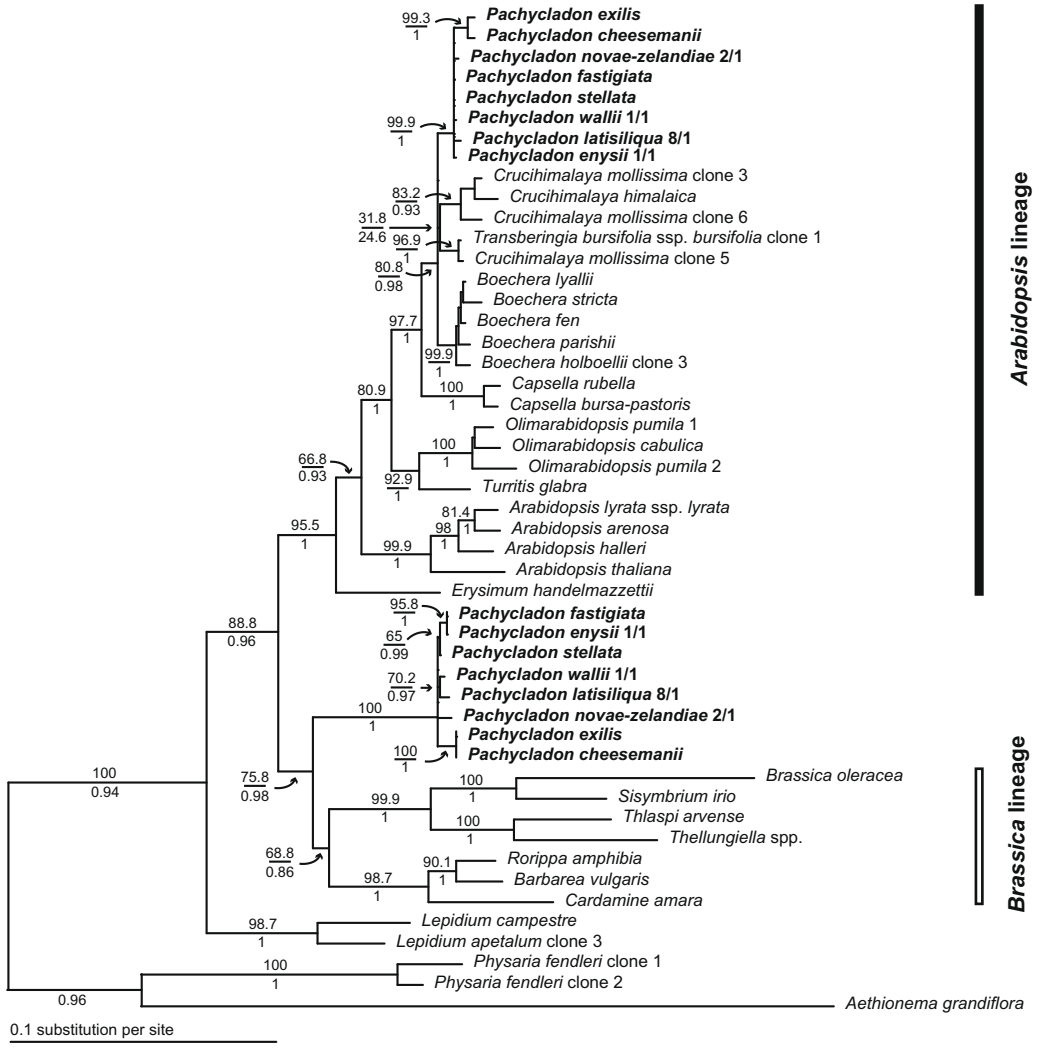

Fig. 2. Maximum likelihood phylogeny of the chalcone synthase (*CHS*) gene. Bootstrap support and posterior clade probabilities are indicated above and below the branches, respectively. The phylogeny was rooted at the position that received the highest posterior probability in the Bayesian analysis.

3.3. Divergence time estimates

The hypothesis of a molecular clock was rejected for *CHS*, *MtN21*, and *PRK* (Table 1). However, when only species of interest were considered, the molecular clock hypothesis was only rejected for the *PRK* gene at the 5% level ($p < 0.01$ after sequential Bonferroni correction). This suggests that although the genes *CHS* and *MtN21* do not evolve neutrally in all Brassicaceae species sampled, they do evolve at the same rate in lineages of interest. The divergence times reported below are the average divergence times across all markers. The gene *PRK* was also included because its estimate was not more extreme then those of other markers (Table 2). The average divergence time between *A. thaliana* and *Arabidopsis lyrata* was 4.92 ± 1.18 mya (Table 2), which is consistent with previous estimates (Koch et al., 2000). The analysis suggests that the two genome copies of *Pachycladon* started to diverge ca. 8.18 ± 4.37 mya. Because the divergence of *P. fastigiata* and *P. exilis* represents the oldest dichotomy in the *Pachycladon* phylogeny (Heenan and Mitchell, 2003), the distance between these species suggests that the genus started to diverge ca. 0.8 mya. To obtain a maximum age estimate for the hybridization event, we looked at the average distance between the A genome in *Pachycladon* species and species that branched at the node immediately below this copy in the phylogeny because they all share a most recent common ancestor with the node that is immediately below *Pachycladon* in the phylogeny. Consequently, all these species must have diverged over the same period of time since their most recent common ancestor. This distance will always constitute an upper bound for the time since the hybridization event because it is possible that the true parent has not been sampled, in which case it would have diverged from *Pachycladon* more recently than that node. The B genome copy was also considered in a calculation for the maximum age of the hybridization event, but because the distance from *Pachycladon* to the closest sampled relatives for the B genome were always greater than that for the A copy, the latter always gave better (i.e., smaller) estimates. These calculations indicated that the maximum age estimate for the hybridization event is 1.61 mya.

4. Discussion

4.1. The origin of the genus Pachycladon

Phylogenetic analysis of molecular markers is a method of choice to identify polyploid events and to determine if a polyploid is of auto- or allopolyploid origin (Straub et al., 2006). The first phylogenetic studies that investigated the position of the genus *Pachycladon* within the Brassicaceae used the internal transcribed spacer of the nuclear ribosomal 18S-5.8S-26S gene family (nrITS)

Please cite this article in press as: Joly, S., et al. A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. Mol. Phylogenet. Evol. (2009), doi:10.1016/j.ympev.2009.02.015

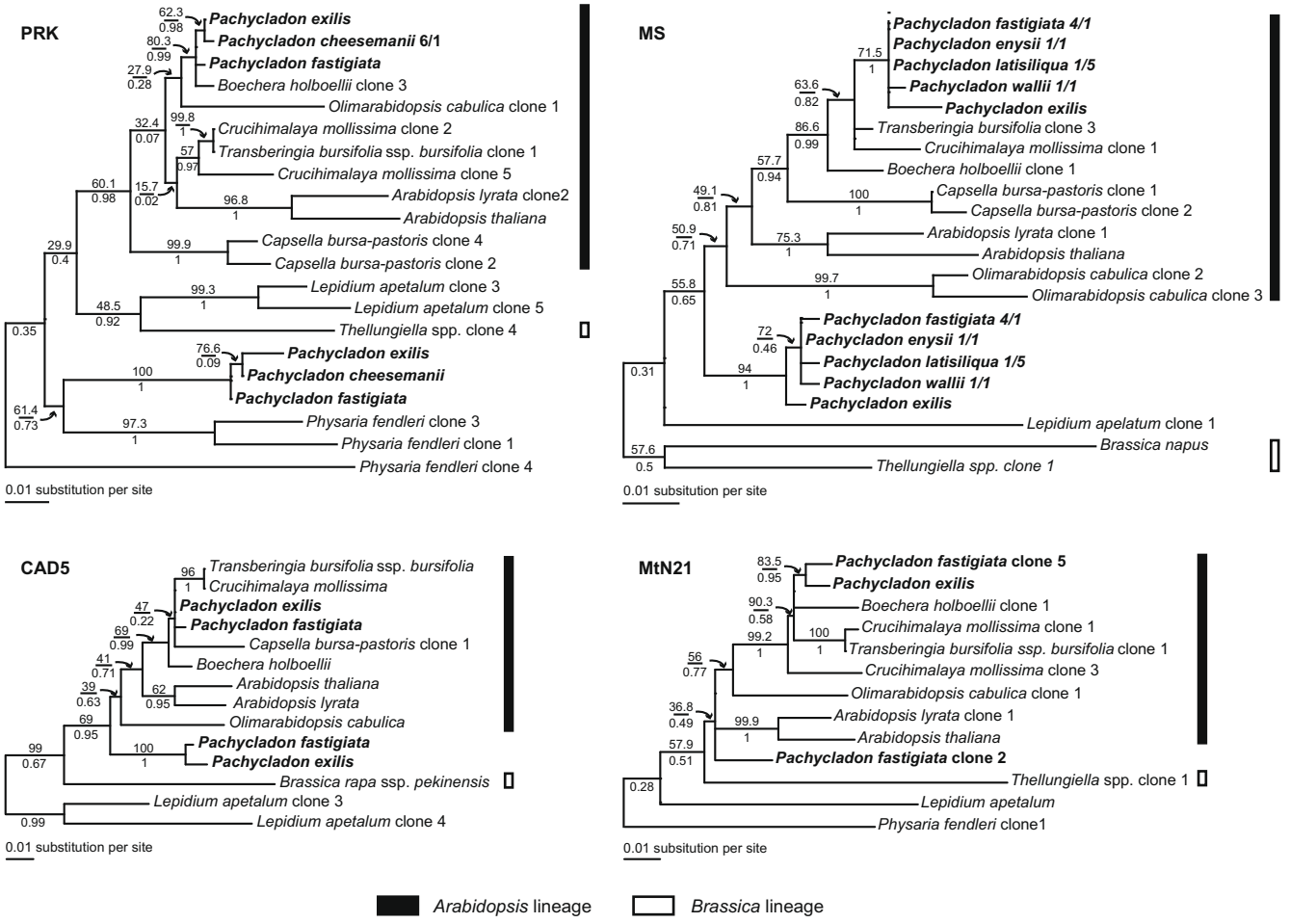

Fig. 3. Maximum likelihood phylogenies for the genes phosphoribulokinase (*PRK*), malate synthase (*MS*), cinnamyl alcohol dehydrogenase 5 (*CAD5*), and *MtN21* nodulation gene. Bootstrap support and posterior clade probabilities are indicated above and below the branches, respectively. The phylogenies were rooted at the positions that received the highest posterior probabilities in the Bayesian analyses.

(Heenan et al., 2002; Mitchell and Heenan, 2000). This marker suggested that *Pachycladon* was closely related to the Eurasian genera *Crucihimalaya* and *Transberingia* (Heenan et al., 2002). There was no evidence of multiple nrITS copies in the *Pachycladon* genome that might have suggested an allopolyploid origin for the genus. However, the 18S-5.8S-26S ribosomal gene family is known to frequently evolve in a concerted manner (Álvarez and Wendel, 2003) and thus it is possible that the nrITS of a single parental species was fixed prior to the species radiation. Since then, although other markers have been sequenced for *Pachycladon* (McBreen and Heenan, 2006), none were analysed within the context of the Brassicaceae family. Thus it has remained unclear whether *Pachycladon* is really polyploid and if so if it is an autopolyploid or an allopolyploid.

The present phylogenetic analyses of five nuclear genes have shown that all of those genes were present in two copies in the genome of *Pachycladon*, which represents strong evidence for a polyploid origin. The hypothesis of independent gene duplication events is unlikely given that the genome copies were in similar positions in the different gene trees and that the rest of the phylogenies did not seem affected by paralogy. Similarly, the possibility that these markers were duplicated during a single event that involved only a fraction of the genome (e.g., a chromosome arm) is improbable given that the five genes investigated are distantly located in genomes of Brassicaceae species. The polyploid event likely occurred before the genus diversified as two copies of all genes have been found in all *Pachycladon* individuals investigated except for *P. fastigiata* with the *MtN21* gene, which is likely a consequence of the few number of clones sequenced (5). These findings thus support previous evidence for the polyploid nature of the genus *Pachycladon*. In addition, the observation that the gene copies group with distantly related species and that the position of these copies were very similar among genes support an allopolyploid origin for *Pachycladon*.

The allopolyploid event that precedes the radiation of *Pachycladon* is the most recent of several whole genome duplication events that have occurred in its evolutionary history. Given that the model plant *A. thaliana* has experienced three whole genome duplication events—the latest of which occurred near the emergence of the Brassicaceae family (Blanc et al., 2003; De Bodt et al., 2005), *Pachycladon* must have experienced four whole genome duplication events in its past. This is a minimal estimate since older genome duplications could have occurred, although it is unlikely that genome duplication events older than the oldest one detected in *A. thaliana* (ca. 101–168 mya: De Bodt et al., 2005) would leave a detectable signature in actual genomes. Such successive rounds of polyploidy followed by genome diploidization (Finnegan, 2002) demonstrate the evolutionary significance of polyploidy in flowering plants.

Identifying parental species for the allopolyploid event that led to *Pachycladon* is difficult due to the variation in phylogenetic position of the *Pachycladon* genome copies among gene trees. This variation may be the result of stochastic error in the phylogenetic analyses, although other factors such as incomplete lineage sort-

Please cite this article in press as: Joly, S., et al. A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. Mol. Phylogenet. Evol. (2009), doi:10.1016/j.ympev.2009.02.015

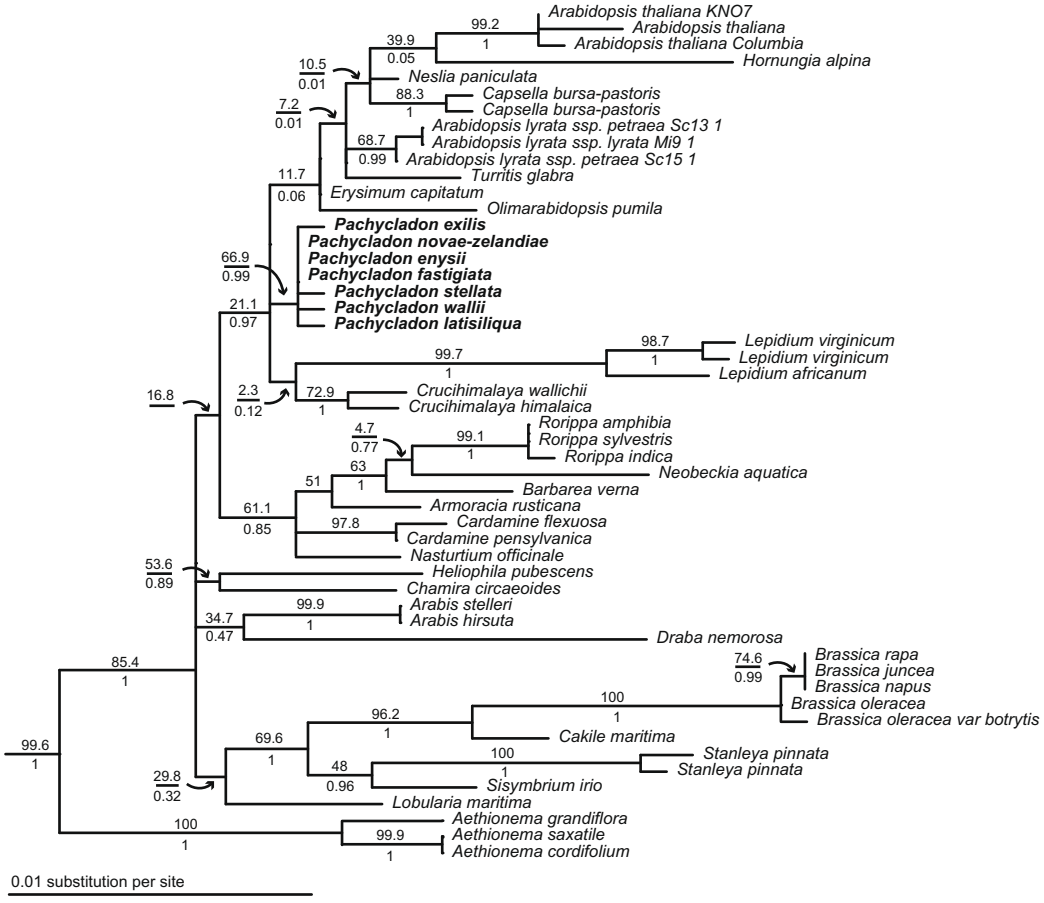

Fig. 4. Portion of the chloroplast (maternal) phylogeny showing all Brassicaceae species. The phylogeny was obtained by maximum likelihood analysis of the large subunit of the riblose-1,5-bisphosphate carboxylase/oxygenase (*rbcL*) gene. Bootstrap support and posterior clade probabilities are indicated above and below the branches, respectively.

Table 2
Divergence times for different splits within the Brassicaceae.

Gene	*Pachycladon* genome A	*Pachycladon* genome B	*Pachycladon* A|*Pachycladon* B	*A. thaliana*|*A. lyrata*	*A. thaliana*|*Brassica*	*Pachycladon* A|*A. thaliana*	*Pachycladon* B|*A. thaliana*	*Pachycladon* and its closest species
CAD5	0.55 ± 0.23	0.00 ± 0.00	8.61 ± 3.65	4.17 ± 1.77	17.31 ± 7.33	5.79 ± 2.45	10.49 ± 4.44	0.83 ± 0.35
CHS	0.83 ± 0.35	1.13 ± 0.48	15.06 ± 6.38	6.29 ± 2.67	21.79 ± 9.23	13.02 ± 5.52	18.45 ± 7.82	1.92 ± 0.81
MS	0.65 ± 0.28	0.87 ± 0.37	5.74 ± 2.43	4.03 ± 1.71	16.44 ± 6.97	5.89 ± 2.50	5.88 ± 2.49	2.48 ± 1.05
MtN21	1.10 ± 0.46	****	3.40 ± 1.44	3.99 ± 1.69	****	6.59 ± 2.79	5.56 ± 2.35	1.53 ± 0.65
PRK	0.88 ± 0.37	1.39 ± 0.59	8.11 ± 3.44	6.14 ± 2.60	****	5.75 ± 2.44	11.58 ± 4.91	1.27 ± 0.54
Average	0.80 ± 0.21	0.85 ± 0.60	8.18 ± 4.37	4.92 ± 1.18	18.51 ± 2.87	7.41 ± 3.16	10.39 ± 5.25	1.61 ± 0.63

**** These dates could not be estimated due to lack of information.

ing, gene paralogy or hybridization may also have contributed to the inferred differences. The position of the A genome is consistent with the relationship proposed by the nrITS phylogeny (Heenan et al., 2002) and suggests a close relationship with genera *Crucihimalaya* and *Transberingia*. In contrast to *Pachycladon*, these genera do not have two divergent genome copies and their position is always close to the A genome copy of *Pachycladon*. This indicates that they were not involved in the same polyploid event as *Pachycladon* and thus represent good candidates for one of the parental lines. Yet, further sampling is necessary to be able to comment more precisely on the nature of the A genome parent.

Phylogenetic analysis of the chloroplast *rbcL* gene suggested that the B genome was contributed by the maternal parent in the allopolyploid event. Yet, it was impossible to find a potential parent species for this copy with the actual sampling. According to a recent supermatrix analysis of the family (Bailey et al., 2006), all major tribes that could potentially contain the parent of the B genome of *Pachycladon* are represented in at least one of the genes investigated except for the tribe Smelowskieae. So unless this latter tribe is close to the B genome, the parent of this copy is probably a species whose phylogenetic relationship within the family has yet to be investigated and several genera are in this situation (Al-Shehbaz et al., 2006). The parent of the B genome could also have gone extinct, but it seems unlikely that no closely related species remains today.

Because parental species must have been in contact in the past to hybridize, it will be interesting to identify more accurately the parental species of *Pachycladon* in the future as it will help to understand how the genus evolved and arrived in New Zealand. For example, the A genome is related to *Transberingia* and *Crucihimalaya*, two genera presently found mostly in arctic regions of North American and Russia and in alpine regions of Eurasia

Please cite this article in press as: Joly, S., et al. A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. Mol. Phylogenet. Evol. (2009), doi:10.1016/j.ympev.2009.02.015

(Al-Shehbaz et al., 1999; Price et al., 2001). If the hybridization event that leads to the allopolyploid occurred in Eurasia or North America, it would imply long-distance trans-oceanic dispersal to New Zealand (Heenan and Mitchell, 2003). Yet, it is also possible that the true ancestor of *Pachycladon* was closer to New Zealand. It would appear unlikely that the hybridization event occurred in New Zealand as no Brassicaceae in New Zealand have been found to be closely related to *Pachycladon* (Mitchell and Heenan, 2000). Nevertheless, limited taxon sampling and different tree topologies in published studies make this conclusion tentative. Irrespective of whether or not transoceanic dispersal predated or postdated formation of the allopolyploid, it is interesting to note that only 0.8 my separates our maximum estimate for formation of the allopolyploid (1.61 mya) and the time of diversification of the genus in New Zealand (Heenan and Mitchell, 2003). Such a short period of time between the hybridization event and the species radiation may explain why the genome of *Pachycladon* species shows little signs of diploidization as suggested by the presence of two potentially functional copies of all nuclear genes investigated in all but one *Pachycladon* species.

The hybridization event that led to the evolution of the genus *Pachycladon* is peculiar in terms of the extent of evolutionary distance between the two parental species. Since their divergence, which most likely occurred at the base of the *Arabidopsis* lineage (ca. 8 mya), more than 450 species have evolved, and these have been classified into four different tribes (Boechereae, Camelineae, Descurainieae, Halimolobeae; Al-Shehbaz et al., 2006). Given that the allopolyploidization event which gave rise to *Pachycladon* took place at most 1.6 mya, this means that the lineages represented by the two parental species were ca. 6.4 million years (myr) diverged prior to their hybridization. Although such a long period of independent evolution is striking, it is difficult to compare this estimate of divergence time with similar estimates for other allopolyploid parents. This is because such estimates of age have rarely been calculated. A recent survey (Chapman and Burke, 2007) provides one measure of comparison for the ages of the parent genomes of allopolyploids. It reported the genetic distance at the nrITS locus for parents of several allopolyploids. However, in this study the inferences are limited because the estimates made do not take into account lineage specific rates of molecular evolution or the age of the allopolyploid event. Unfortunately, it is also not possible to measure the nrITS distance between the parent genomes of *Pachycladon* because of the absence of a potential B genome donor and because the nrITS B copies have either been homogenized or lost from the *Pachycladon* genome (see above). Nevertheless, the time of independent evolution of the parents prior to the hybridization event obtained here is greater than that obtained for *Gossypium* (5.2 myr: Senchina et al., 2003), which is one of the allopolyploids with the greatest observed nrITS distance between parent genomes (Chapman and Burke, 2007). Moreover, two of the three allopolyploids reported by these authors to have a larger nrITS distance between parents than *Gossypium—Spartina anglica* (Fortune et al., 2007), and also *Symphyotricum ascendens* when re-estimated using the same sequences and parameters as Chapman and Burke (2007)—are clearly younger than *Pachycladon*. Thus, the successful natural hybridization event at the origin of the *Pachycladon* radiation is certainly one of the most extreme examples yet reported of successful hybridization between distantly related parents in flowering plants.

Further support for the occurrence of a hybridization event between divergent genomes A and B is provided by the generation of an artificial intergeneric hybrid between *A. thaliana* (maternal parent; $2n = 10$) and *Pachycladon cheesemanii* ($2n = 20$) (Heenan et al., 2008). This intergeneric F1 hybrid has chromosome number of $2n = 15$ that is intermediate between the parents and is both female and male sterile. However, female function was successfully restored by genome duplication induced by colchicine, which demonstrates the importance of chromosome doubling as a mechanism to overcome sterility and facilitating subsequent generation of new hybrids.

4.2. Hybridization: a trigger for adaptive radiations?

Recently, it has been proposed that hybridization can predispose a group to adaptive radiation (Seehausen, 2004). Findings of the present study are consistent with such a view because hybridization clearly preceded the radiation of *Pachycladon*. Our finding is interesting given that there are few documented examples of hybridization events preceding species radiation besides that of the silversword alliance (Barrier et al., 1999). These observations may help to interpret the evolution of flora such as that of New Zealand that has been reported to have a high incidence of polyploidy (Hair, 1966; Murray et al., 2005). In particular, allopolyploidy has been hypothesised to precede the adaptive radiation of the New Zealand alpine *Ranunculus* (Hair, 1966). Our findings provide motivation for investigating other speciose polyploid genera in the New Zealand archipelago.

Although our findings are consistent with a hypothesis that hybridization has been important in the radiation of *Pachycladon*, it remains to be investigated whether or not the inferred hybridization event has facilitated adaptive radiation. It might be possible to test this by looking at the fate of duplicated genes, and evaluating whether or not these gene copies have allowed specialization and adaptation to new environments. This hypothesis is presently being investigated.

4.3. Brassicaceae phylogeny

Although the present study was not intended as an investigation of phylogenetic relationships within the Brassicaceae family, it has nevertheless produced some interesting results, especially since it includes more independently evolving markers than have been analysed in previous analyses. One interesting finding is the basal positions of *Physaria* and *Lepidium* in the family whereas previous multi-gene studies have suggested that these species belong to the *Arabidopsis* lineage (Bailey et al., 2006; Koch et al., 2007). Also, the incongruence observed among our and previously published gene trees (Bailey et al., 2006; Koch et al., 2007) may indicate that the family as a whole has evolved very rapidly and suggests that it will be important to analyse several independently evolving markers to obtain better phylogenetic hypotheses for the Brassicaceae. Finally, the long-distance hybridization event that has preceded the radiation of *Pachycladon* might have significance for understanding diversification within the Brassicaceae family. It suggests that hybridization, even between very distant species, has been an important evolutionary feature of the family that could explain some phenotypic characteristics (Al-Shehbaz et al., 2006).

Acknowledgments

We thank Martin Lysak for plant material and useful feedback on the results and Shannel Courtney and Kerry Ford for assistance with the collection of plant material. S.J. was supported by a NSERC (Canada) post-doctoral fellowship. Financial support for this study comes from a New Zealand Marsden Grant.

Appendix 1 and 2 Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.02.015.

References

Al-Shehbaz, I.A., Beilstein, M.A., Kellogg, E.A., 2006. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Syst. Evol. 259, 89–120.

Al-Shehbaz, I.A., O'Kane', S.L., Price, R.A., 1999. Generic placement of species excluded from Arabidopsis. Novon 9, 296–307.

Álvarez, I., Wendel, J.F., 2003. Ribosomal ITS sequence and plant phylogenetic inference. Mol. Phylogenet. Evol. 29, 417–434.

Anderson, E., 1949. Introgressive Hybridization. John Wiley, New York.

Arnold, M.L., 1997. Natural Hybridization and Evolution. Oxford University Press, New York.

Bailey, D.C., Koch, M.A., Mayer, M., Mummenhoff, K., O'Kane', S.L., Warwick, S.I., Windham, M.D., Al-Shehbaz, I.A., 2006. Toward a global phylogeny of the Brassicaceae. Mol. Biol. Evol. 23, 2142–2160.

Barrier, M., Baldwin, B.G., Robichaux, R.H., Purugganan, M.D., 1999. Interspecific hybrid ancestry of a plant adaptive radiation: allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. Mol. Biol. Evol. 16, 1105–1113.

Barton, N.H., 2001. The role of hybridization in evolution. Mol. Ecol. 10, 551–568.

Blanc, G., Hokamp, K., Wolfe, K.H., 2003. A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. Genome Res. 13, 137–144.

Bruen, T.C., 2005. PhiPack: PHI Test and other Tests of Recombination. McGill University, Montréal, Québec, Canada.

Bruen, T.C., Philippe, H., Bryant, D., 2006. A simple robust statistical test for detecting the presence of recombination. Genetics 172, 2665–2681.

Chapman, M.A., Burke, J.M., 2007. Genetic divergence and hybrid speciation. Evolution 61, 1773–1780.

Collins, L.J., Biggs, P.J., Voelckel, C., Joly, S., 2008. An approach to transcriptome analysis of non-model organism using short-read sequences. Genome Inform. 21, 3–14.

De Bodt, S., Maere, S., Van de Peer, Y., 2005. Genome duplication and the origin of angiosperms. Trends Ecol. Evol. 20, 591–597.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Ellstrand, N.C., Schierenbeck, K.A., 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? Proc. Natl. Acad. Sci. USA 97, 7043–7050.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.

Ferguson, D., Sang, T., 2001. Speciation through homoploid hybridization between allotetraploids in peonies (*Paeonia*). Proc. Natl. Acad. Sci. USA 98, 3915–3919.

Finnegan, E.J., 2002. Epialleles—a source of random variation in times of stress. Curr. Opin. Plant Biol. 5, 101–106.

Fortune, P.M., Schierenbeck, K.A., Ainouche, A.K., Jacquemin, J., Wendel, J.F., Ainouche, M.L., 2007. Evolutionary dynamics of *Waxy* and the origin of hexaploid *Spartina* species (Poaceae). Mol. Phylogenet. Evol. 43, 1040–1055.

Goldman, N., Yang, Z., 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11, 725–736.

Grant, B.R., Grant, P.R., 1996. High survival of Darwin's finch hybrids: effects of beak morphology and diets. Ecology 77, 500–509.

Grant, V., 1981. Plant Speciation, second ed. Columbia University Press, New York.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hair, J.B., 1966. Biosystematics of the New Zealand flora, 1945–1964. NZ J. Bot. 4, 559–595.

Heenan, P.B., Dawson, M.I., Smissen, R.D., Bicknell, R.A., 2008. An artificial intergeneric hybrid derived from sexual hybridization between the distantly related *Arabidopsis thaliana* and *Pachycladon cheesemanii* (Brassicaceae). Bot. J. Linn. Soc. 157, 533–544.

Heenan, P.B., Mitchell, A.D., 2003. Phylogeny, biogeography and adaptive radiation of Pachycladon (Brassicaceae) in the mountains of South Islad, New Zealand. J. Biogeogr. 30, 1737–1749.

Heenan, P.B., Mitchell, A.D., Koch, M.A., 2002. Molecular systematics of the New Zealand *Pachycladon* (Brassicaceae) complex: generic circumscription and relationships to *Arabidopsis* sens. lat. and *Arabis* sens. lat. NZ J. Bot. 40, 543–562.

Johannessen, M., Andersen, B.A., Damgaard, C., Jorgensen, R.B., 2005. Maternal inheritance of Chloroplasts between *Brassica rapa* and F1-hybrids demonstrated by cpDNA markers specific to oilseed rape and *B. Rapa*. Mol. Breeding 16, 271–278.

Joly, S., Starr, J.R., Lewis, W.H., Bruneau, A., 2006. Polyploid and hybrid evolution in roses east of the Rocky Mountains. Am. J. Bot. 93, 412–425.

Knapp, M., Stöckler, K., Havell, D., Delsuc, F., Sebastiani, F., Lockhart, P.J., 2005. Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (Southern Beech). Plos Biol. 3, e14.

Koch, M.A., Dobes, C., Kiefer, C., Schmickl, R., Klimes, L., Lysak, M.A., 2007. Supernetwork identifies multiple events of plastid trnF(GAA) pseudogene evolution in the Brassicaceae. Mol. Biol. Evol. 24, 63–73.

Koch, M.A., Haubold, B., Mitchell-Olds, T., 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). Mol. Biol. Evol. 17, 1483–1498.

Koch, M.A., Haubold, B., Mitchell-Olds, T., 2001. Molecular systematics of the Brassicaceae: evidence from coding plastidic *matK* and nuclear *CHS* sequences. Am. J. Bot. 88, 534–544.

Lewis, C.E., Doyle, J.J., 2001. Phylogenetic utility of the nuclear gene malate synthase in the palm family (Arecaceae). Mol. Phylogenet. Evol. 19, 409–420.

Lewis, C.E., Doyle, J.J., 2002. A phylogenetic analysis of tribe Areceae (Arecaceae) using two low-copy nuclear genes. Plant Syst. Evol. 236, 1–17.

Lewontin, R.C., Birch, L.C., 1966. Hybridization as a source of variation for adaptation to new environments. Evolution 20, 315–336.

Lynch, M., Conery, J.S., 2000. The evolutionary fate and consequences of duplicate genes. Science 290, 1151–1155.

Mai, D.H., 1995. Tertiäre vegetationsgeschichte Europas. Gustzv Fisher, Jena, Stuttgart, New York.

McBreen, K., Heenan, P.B., 2006. Phylogenetic relationships of Pachycladon (Brassicaceae) species based on three nuclear and two chloroplast DNA markers. NZ J. Bot. 44, 377–386.

Mitchell, A.D., Heenan, P.B., 2000. Systematic relationships of New Zealand endemic Brassicaceae inferred from nrDNA ITS sequence data. Syst. Bot. 25, 98–105.

Murray, B.G., De Lange, P.J., Fergusson, A.R., 2005. Nuclear DNA variation, chromosome numbers and polyploidy in the endemic and indigenous grass flora of New Zealand. Ann. Bot. 96, 1293–1305.

Otto, S.P., Whitton, J., 2000. Polyploid incidence and evolution. Annu. Rev. Genet. 34, 401–437.

Pérez-Losada, M., Hoeg, J.T., Crandall, K.A., 2004. Unraveling the evolutionary radiation of the Thoracian barnacles using molecular and morphological evidence: a comparison of several divergence time estimation approaches. Syst. Biol. 53, 244–264.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Price, R.A., Al-Shehbaz, I.A., O'kane, S.L., 2001. *Beringia* (Brassicaceae), a new genus of Arabidopsoid affinities from Russia and North America. Novon 11, 332–336.

Rambaut, A., Drummond, A.J., 2005. Tracer. Available from: <http://tree.bio.ed.ac.uk/software/tracer/>.

Rice, W.R., 1989. Analysing tables of statistical tests. Evolution 43, 223–225.

Rieseberg, L.H., 1997. Hybrid origins of plant species. Annu. Rev. Ecol. Syst. 28, 359–389.

Rieseberg, L.H., Raymond, O., Rosenthal, D.M., Lai, Z., Livingstone, K., Nakazato, T., Durphy, J.L., Schwarzbach, A.E., Donovan, L.A., Lexer, C., 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. Science 301, 1211–1216.

Schranz, M.E., Lysak, M.A., Mitchell-Olds, T., 2006. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends Plant Sci. 11, 535–542.

Seehausen, O., 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19, 198–207.

Senchina, D.S., Alvarez, I., Cronn, R.C., Liu, B., Rong, J., Noyes, R.D., Paterson, A.H., Wing, R.A., Wilkins, T.A., Wendel, J.F., 2003. Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. Mol. Biol. Evol. 20, 633–643.

Soltis, P.S., Soltis, D.E., 2000. The role of genetic and genomic attributes in the success of polyploids. Proc. Natl. Acad. Sci. USA 97, 7051–7057.

Soltis, P.S., Soltis, D.E., Savolainen, V., Crane, P.R., Barraclough, T.G., 2002. Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. Proc. Natl. Acad. Sci. USA 99, 4430–4435.

Stebbins, G.L., 1959. The role of hybridization in evolution. Proc. Am. Philos. Soc. 103, 231–251.

Straub, S.C.K., Pfeil, B.E., Doyle, J.J., 2006. Testing the polyploid past of soybean using a low-copy nuclear gene—is *Glycine* (Fabaceae: Papilionoideae) an auto- or allopolyploid? Mol. Phylogenet. Evol. 39, 580–584.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and other Methods). Sinauer Associates, Sunderland.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

Voelckel, C., Heenan, P.B., Janssen, B., Reichelt, M., Ford, K., Hofmann, R., Lockhart, P.J., 2008. Transcriptional and biochemical signatures of divergence in natural populations of two species of New Zealand alpine *Pachycladon*. Mol. Ecol. 17, 4740–4753.

Winkworth, R.C., Wagstaff, S.J., Glenny, D., Lockhart, P.J., 2005. Evolution of the New Zealand alpine flora: origins, diversification and dispersal. Organ. Divers. Evol. 5, 237–247.

Yang, Y.-W., Lai, K.-N., Tai, P.-Y., Wen-Hsiung, L., 1999. Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. J. Mol. Evol. 48, 597–604.

Yang, Z., 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.

Yang, Z., Yoder, A.D., 2003. Comparison of likelihood and bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. Syst. Biol. 52, 705–716.

Please cite this article in press as: Joly, S., et al. A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. Mol. Phylogenet. Evol. (2009), doi:10.1016/j.ympev.2009.02.015