Molecular Phylogenetics and Evolution 51 (2009) 269–280

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

# Allopolyploidy in Fragariinae (Rosaceae): Comparing four DNA sequence regions, with comments on classification

Magnus Lundberg [a,*], Mats Töpel [b], Bente Eriksen [b], Johan A.A. Nylander [a], Torsten Eriksson [a,c]

[a] Department of Botany, Stockholm University, SE-10691, Stockholm, Sweden
[b] Department of Environmental Sciences, Gothenburg University, Box 461, SE-40530, Göteborg, Sweden
[c] Bergius Foundation, Royal Swedish Academy of Sciences, SE-10405, Stockholm, Sweden

ARTICLE INFO

*Article history:*
Received 23 June 2008
Revised 25 February 2009
Accepted 26 February 2009
Available online 5 March 2009

*Keywords:*
Allopolyploidy
Fragariinae
Incongruence
Molecular phylogeny
Bayesian convergence diagnostics

ABSTRACT

Potential events of allopolyploidy may be indicated by incongruences between separate phylogenies based on plastid and nuclear gene sequences. We sequenced two plastid regions and two nuclear ribosomal regions for 34 ingroup taxa in Fragariinae (Rosaceae), and six outgroup taxa. We found five well supported incongruences that might indicate allopolyploidy events. The incongruences involved *Aphanes arvensis*, *Potentilla miyabei*, *Potentilla cuneata*, *Fragaria vesca*/*moschata,* and the *Drymocallis* clade. We evaluated the strength of conflict and conclude that allopolyploidy may be hypothesised in the four first cases. Phylogenies were estimated using Bayesian inference and analyses were evaluated using convergence diagnostics. Taxonomic implications are discussed for genera such as *Alchemilla*, *Sibbaldianthe*, *Chamaerhodos*, *Drymocallis* and *Fragaria*, and for the monospecific *Sibbaldiopsis* and *Potaninia* that are nested inside other genera. Two orphan *Potentilla* species, *P. miyabei* and *P. cuneata* are placed in Fragariinae. However, due to unresolved topological incongruences they are not reclassified in any genus.



## 1. Introduction

The importance and detection of hybridisation and its role in species formation have been, and continue to be, major foci in evolutionary plant research, notably because hybridisation followed by polyploidisation is suggested to be an important mode of speciation among vascular plants. Following hybridisation, alternations at genomic and gene levels occur in the offspring with potential advantages of the polyploid species compared to its diploid progenitors. This may be caused by changes of the newly formed hybrids by reproductive isolation, or by changes in biochemical, physiological and developmental flexibility; changes that are expressed, for example in plant size, flowering time and reproductive output, but also in new combinations of characters (Levin, 1983; Schranz and Osborn, 2000; Song et al., 1995). Speciation through polyploidy was discovered in the beginning of the last century by Winkler (1916) and Winge (1917). To distinguish between the two main types of polyploidy commonly recognised, Kihara and Ono (1926) introduced the terms allopolyploidy (duplication of chromosomes in interspecific hybrids) and autopolyploidy (duplication of chromosomes within a species).

Many estimates of the frequency of polyploidy among angiosperms have been presented. Early estimates, based on chromosome numbers in related species or arbitrary levels above which species were thought to be polyploid, ranged between 30% and 52% (Müntzing, 1936; Darlington, 1937; Stebbins, 1950; Grant, 1963, 1981). More recent studies based on guard cell sizes suggest that more than 70% of all extant angiosperms are of polyploid origin (Masterson, 1994), and polyploidy has been suggested to be the cause of at least 2–4% of all recent speciation events (Otto and Whitton, 2000), thus suggested to be an important mode of speciation (e.g. Mallet, 2007). Such estimates, in particular the older ones, may be rather inaccurate. As an example of the problems in making such estimates, the model plant *Arabidopsis thaliana* is defined as a functional diploid with a relatively small genome (The Arabidopsis Genome Initiative, 2000). However, results from whole genome sequencing indicates that there have been two or more rounds of genome duplication events in the evolution of that species (Vision et al., 2000; Blanc et al., 2003; Bowers et al., 2003; Simillion et al., 2002). Recent genetic and genomic studies suggest that most or perhaps all angiosperms have undergone one or several rounds of polyploidisation followed by extensive diploidisation (Wolfe, 2001; Eckhardt, 2001), the evolutionary process where the genomic content of a polyploid species degenerates into a diploid state again.

Before molecular methods were developed, polyploid speciation was detected by chromosome counts and crossing experiments (Grant, 1981). Drawbacks of this method include that it assumes recent polyploidy events and that parental species are extant. More recently, several analyses using molecular data have been developed to detect and reconstruct hybrid speciation, e.g.

* Corresponding author. Fax: +46 (0)8 165525.
*E-mail address:* Magnus.Lundberg@botan.su.se (M. Lundberg).

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.02.020

splits decomposition (Bandelt and Dress, 1992; Huson, 1998) and linkage disequilibrium (e.g. Rieseberg et al., 2003). A third, and in our view the currently most powerful way to reconstruct allopolyploid speciation is to use phylogenies based on nuclear single or low copy DNA markers (e.g. Sang and Zhang, 1999; Smedmark et al., 2003; Mason-Gamer, 2001; Popp et al., 2005). In the ideal case, one copy of the gene would be present in the phylogeny for each ploidy level. Such a complex but bifurcating tree can then be turned into a more realistic estimation of the phylogeny: a network picturing both reticulate and bifurcating branches. This method has several advantages: (1) The entire phylogeny of the group is analysed and allopolyploidy can be detected even if ancestors are not extant. (2) Parentage is traceable because allopolyploids have one gene copy from each parent, at least initially. (3) It may be possible to trace several rounds of allopolyploidy events. (4) It is possible to distinguish allopolyploidy (where gene copies from single organisms are distantly related) from autopolyploidy (where gene copies are sisters). There is now a software package available that aims to turn a bifurcating tree with gene copies into a reticulate tree ("Padre"; Huber et al., 2006).

However, a difficulty of this method is that it involves extensive cloning and sequencing of low copy genes, and in order to be able to obtain sequences from all copies it is often required to work with fresh plant material. For this reason, it is practical to first screen a group for potential allopolyploidy events and then focus on specific parts of the phylogeny, using low copy gene sequences. Screening involves sequencing a large set of species for easily amplified DNA regions. Phylogenies are then estimated for each data set and topologies are compared, contrasting plastid sequence data that are generally maternally inherited with nuclear data that are potentially paternal because of biparental inheritance. This comparative step allows us to detect and evaluate potential conflicts between the separate data sets. Any incongruence found should be well supported and any conflict that can be attributed to problems with analysis rather than data should be dismissed. The next step is to form hypotheses that such well supported incongruences are caused by allopolyploidy and continue with further investigations using low copy gene regions. The initial screening approach has been successful in detecting possible allopolyploid speciation in other studies (e.g. Smedmark and Eriksson, 2002). This simplified approach may underestimate the number of allopolyploid speciation events because some might not be detected. The ribosomal DNA regions, which are used in this study are usually under the influence of concerted evolution, a process that serves to homogenise all copies towards a single repeat type or more rarely to a mix (Wendel et al., 1995; Alice et al., 2001), and this may serve to obscure information about allopolyploidy events. If the paternal copy should become fixed, a topological incongruence would be observed when comparing the nuclear phylogeny with the plastid. If, on the other hand the maternal copy was fixed, the two topologies would be congruent and the hybrid origin will be hidden. In such cases, evidence of polyploidy in itself may help detection. In addition, in case of incomplete homogenisation rare repeat types of ribosomal DNA can be hard to recover by direct sequencing (e.g. Rauscher et al., 2002) and more sensitive methods e.g. cloning may be required.

Other processes that may produce topological incongruent trees include incomplete lineage sorting, independent gene duplications, random loss in multiple genes, horizontal gene transfer, selection, and heterotachy (e.g. Rokas et al., 2003; Richardson and Palmer, 2007; Lopez et al., 2002). An analysis using low copy nuclear sequence data may help to distinguish between such processes. Additionally, during the analysis, poor MCMC simulations or the effect of priors may cause incongruent topologies. Problems in the data and/or misleading settings may produce trees with conflicting topologies and with apparent high clade credibilities (Huelsenbeck et al., 2002; Nylander et al., 2008). It is possible to investigate the performance of the MCMC runs and to reduce the risk of analysis artifacts by using convergence diagnostics (Nylander et al., 2008).

Fragariinae is a subclade of Rosoideae in Rosaceae (Morgan et al., 1994; Eriksson et al., 2003; Potter et al., 2002, 2007). The clade Fragariinae was discovered by Eriksson et al. (1998) based on nuclear ribosomal sequence data and it was further supported by combining nuclear ribosomal data with plastid sequences (Eriksson et al., 2003). Fragariinae is the sister group of *Potentilla* in strict sense. Different segregates of genera in Fragariinae have previously often been classified together with *Potentilla*, but genera such as *Alchemilla* and *Potaninia* have in most cases been placed elsewhere in Rosaceae (Focke, 1894; Hutchinson, 1964; Schulze-Menz, 1964; Kalkman, 1988, 2004). Eriksson et al. (1998, 2003) redefined *Potentilla* and noted that the rather few species that had been classified as *Potentilla* by earlier authors but were found to belong to Fragariinae had already also been classified in other genera. However, a few orphan *Potentilla* species still remain in Fragariinae without having been reclassified.

Fragariinae contain 10 genera and a few orphan *Potentilla* species: *Alchemilla* in the wide sense (at least 350 spp.) including *Aphanes* and *Lachemilla* (Gehrke et al., 2008), *Comarum* (2 spp.), *Dasiphora* (ca. 4 spp.), *Potaninia* (1 sp.), *Sibbaldia* (2–6 spp.), *Sibbaldianthe* (2 spp.), *Sibbaldiopsis* (1 sp.), *Chamaerhodos* (5–8 spp.), *Fragaria* (ca. 20 spp.) and *Drymocallis* (ca. 30 spp.), most of which occur in the northern hemisphere with their highest diversity in temperate regions. Previously used taxon names are presented in Table 1. There are no known distinctive morphological synapomorphies of Fragariinae but it is strongly supported by DNA sequence data. The Fragariinae share a number of non-molecular characters with *Potentilla* such as that a few are woody (*Comarum salesovianum*, *Dasiphora* and *Potaninia*); they have pistils with lateral to basal styles; and a basal chromosome number of $x = 7$ (*Alchemilla* is an exception with $x = 8$). An enlarged receptacle is commonly present in Fragariinae, although lacking in *Alchemilla* and *Potaninia*. Pistils are numerous except in *Potaninia* and some *Alchemilla* and the fruit is an achene or achenetum (Potter et al., 2007). Members of Fragariinae (*Fragaria*, *Alchemilla*) and other closely related genera (*Rosa*, *Rubus*, *Potentilla* among others) have been suggested to have a hybrid origin (Schulze-Menz, 1964). Also, polyploidy is widespread in these groups. For instance, *Potentilla* (in the wide sense of many authors) has ploidy levels ranging from diploids to 16× (Grant, 1981). Known ploidy levels of taxa included in this study are presented in Table 2.

The aims of this study were to (1) screen the clade Fragariinae for potential allopolyploidy speciation events. This was done by identifying topological incongruences between phylogenies based on DNA sequence data of plastid and nuclear origin, using a set

**Table 1**
Names of some species as used in this study, compared to names used in the study of Eriksson et al. (2003) and synonyms.

| Taxon name | Name in Eriksson et al. (2003) | Synonyms |
|---|---|---|
| *Comarum palustre* L. | *Potentilla palustris* Scop. | |
| *Comarum salesovianum* (Stephan) Bunge | *Potentilla salesowianum* [sphalm.] | |
| *Dasiphora fruticosa* (L.) Rydb. | *Potentilla fruticosa* L. | *Pentaphylloides* Duhamel |
| *Drymocallis agrimonioides* Rydb. | *Potentilla arguta* Pursh. | |
| *Sibbaldianthe bifurca* (L.) Kurtto & T. Erikss. | *Potentilla bifurca* L. | *Schistophyllidium* (Juz. ex Fed.) Ikonn. |
| *Sibbaldiopsis tridentata* (Aiton) Rydb. | *Potentilla tridentata* Aiton | |

**Table 2**
Taxon sampling, voucher specimen, GenBank accession numbers for ETS, ITS, *trn*L/F and *trn*S/G and information about chromosome number (Goldblatt and Johnson, 2008; Lid and Lid, 1994).

| Taxa | Voucher | GenBank accession | | | | Ploidy level |
|---|---|---|---|---|---|---|
| | | ETS | ITS | *trn*L/F | *trn*S/G | |
| *Alchemilla alpina* L. | T. Eriksson 805 (S) | FJ422342/ FJ422343 | | AJ512217[b] | FJ422305 | ca 15–18 ploid |
| | R. Eriksson s.n. (GH) | | U90816[a]/ U90817[a] | | | |
| *Alchemilla cryptantha* Steud. ex A. Rich | T. Eriksson 914 (S) | FJ422344 | FJ356153 | FJ422283 | FJ422306 | Decaploid |
| *Alchemilla mollis* (Buser) Rothm. | T. Eriksson s.n. (S) | FJ422345 | AJ511769[b] | AJ512218[b] | FJ422307 | ca 11–13 ploid |
| *Alchemilla pentaphylla* L. | B. Gehrke BG-E400 (ZH) | FJ422346 | FJ356154 | FJ422284 | FJ422308 | Not available |
| *Aphanes arvensis* L. | T. Eriksson s.n. (S) | FJ422347 | AJ511770[b] | AJ512234[b] | FJ422309 | Hexaploid |
| *Aremonia agrimonoides* (L.) DC. | Karlsson 94076 (LD) | Missing | U90799[a] | AJ512230[b]/ AJ512231[b] | FJ422310 | Hexaploid |
| *Chamaerhodos erecta* (L.) Bunge | Norlindh & Ahti 10161A (S) | FJ422348 | | AJ512219[b] | FJ422311 | Diploid |
| | Lackschewitz 11453 (GH) | | U90794[a] | | | |
| *Chamaerhodos mongolica* Bunge | E. Rosenius 1028 (S) | FJ422349 | FJ356155 | FJ422285 | FJ422312 | Not available |
| *Chamaerhodos nuttallii* Pickering ex Rydb. | J.W. Moore 23133 (S) | FJ422350 | FJ356156 | FJ422286 | Missing | Diploid |
| *Chamaerhodos sabulosa* Bunge | Joel Eriksson 618 (S) | FJ422351 | FJ356157 | Missing | Missing | Diploid |
| *Comarum palustre* L. | M. Lundberg 17 (S) | FJ422352 | FJ356158 | | FJ422313 | Hexaploid |
| | T. Eriksson 659 (GH, S) | | | AJ512237[b] | | |
| *Comarum salesovianum* (Stephan) Bunge | M. Lundberg 3 (S) | FJ422353 | | | FJ422314 | Not available |
| | Eriksson & Vretblad TE751 (S) | | AJ511779[b] | AJ512228[b] | | |
| *Dasiphora davurica* (Nestl.) Kom. & Aliss. | M. Lundberg 24 (S) | FJ422354 | FJ356159 | FJ422287 | FJ422315 | Not available |
| *Dasiphora fruticosa* (L.) Rydb. | T. Eriksson 806 (S) | FJ422355 | | AJ512233[b] | FJ422316 | Diploid/ tetraploid |
| | Karlsson 94074 (LD) | | U90808[a]/ U90809[a] | | | |
| *Dasiphora phyllocalyx* Juz. | T. Eriksson 757 (S) | Missing | FJ356160 | FJ422288 | FJ422317 | Not available |
| *Drymocallis agrimonioides* Rydb. | M Lundberg 15 (S) | FJ422356 | | FJ422289 | FJ422318 | Diploid |
| | Laferrière 2357 (A) | | U90787[a] | | | |
| *Drymocallis corsica* (Soleirol ex Lehm.) Kurtto | M. Lundberg 13 (S) | FJ422357 | FJ356161 | FJ422290 | FJ422319 | Not available |
| *Drymocallis glutinosa* Rydb. | M. Lundberg 5 (S) | FJ422358 | FJ356162 | FJ422291 | FJ422320 | Not available |
| *Drymocallis rupestris* (L.) Soják | M. Lundberg 6 (S) | FJ422359 | FJ356163 | FJ422292 | FJ422321 | Diploid |
| *Fragaria chiloensis* (L.) Mill. | M. Lundberg 14 (S) | FJ422360 | FJ356164 | FJ422293 | FJ422322 | Octoploid |
| *Fragaria moschata* Weston | T. Eriksson 826 (S) | FJ422361 | FJ356165 | FJ422294 | FJ422323 | Hexaploid |
| *Fragaria vesca* L. | Eriksson & Smedmark 43 (S) | FJ422362 | AJ511771[b] | AJ512232[b] | FJ422324 | Diploid |
| *Fragaria virginiana* Mill. | T. Eriksson s.n. (S) | FJ422363 | AJ511772[b] | AJ512220[b] | FJ422325 | Octoploid |
| *Fragaria viridis* Weston | M. Lundberg 16 (S) | FJ422364 | FJ356166 | FJ422295 | FJ422326 | Diploid |
| *Fragaria x ananassa* (Weston) Duchesne ex Rozier 'Pink Panda' | M. Lundberg 8 (S) | FJ422365 | FJ356167 | FJ422296 | FJ422327 | Octoploid |
| *Potaninia mongolica* Maxim. | Norlindh & Ahti 10384 (S) | FJ422366 | AM286742[c] | AM286743[c] | FJ422328 | Not available |
| *Potentilla alchemilloides* Lapeyr. | A. & A.-L. Anderberg 26 (S) | FJ422367 | FJ356168 | FJ422297 | FJ422329 | Not available |
| *Potentilla cuneata* Wall. & Lehm. | M. Lundberg 39 (S) | FJ422368 | FJ356169 | FJ422298 | FJ422330 | Tetraploid |
| *Potentilla caulescens* L. | T. Eriksson s.n. (S) | Missing | FJ356170 | Missing | FJ422331 | Diploid |
| *Potentilla lignosa* Willd. ex Schlecht. | M. Töpel MA132 (GB) | FJ422369 | FJ356171 | FJ422299 | FJ422332 | Not available |
| *Potentilla miyabei* Makino | Sten Bergman s.n. (S) | FJ422370 | FJ356172 | FJ422300 | Missing | Not available |
| *Rosa majalis* Herrm. | T. Eriksson 641 (GH, S) | FJ422371 | U90801[a] | AJ512229[b] | FJ422333 | Diploid |
| *Sanguisorba officinalis* L. | T. Eriksson 804 (S) | FJ422372 | | AJ416465[b] | FJ422334 | Tetra-/hexa-/ octoploid |
| | T. Eriksson s.n. (GH) | | U90797[a] | | | |
| *Sibbaldia cuneata* Hornem | Binns 5 (E) | FJ422373 | FJ356173 | FJ422301 | FJ422335 | Not available |
| *Sibbaldia parviflora* Willd. | M. Lundberg 4 (S) | FJ422374 | FJ356174 | FJ422302 | FJ422336 | Diploid |
| *Sibbaldia procumbens* L. | Eriksson 698 (S) | FJ422375 | | AJ512235[b] | FJ422337 | Diploid |
| | Aronsson s.n. (S) | | U90820[a]/ U90821[a] | | | |
| *Sibbaldia semiglabra* C.A. Mey. | J. Klackenberg 820621-11 (S) | FJ422376 | FJ356175 | FJ422303 | FJ422338 | Not available |
| *Sibbaldianthe adpressa* (Bunge) Juz. | V. A. Gusev 391 (S) | FJ422377 | FJ356176 | FJ422304 | FJ422339 | Not available |
| *Sibbaldianthe bifurca* (L.) Kurtto & T. Erikss. | M. Lundberg 1 (S) | FJ422378 | | | FJ422340 | Octoploid |
| | T. Eriksson 811 (S) | | | AJ512224[b] | | |

(*continued on next page*)

**Table 2** (*continued*)

| Taxa | Voucher | GenBank accession | | | | Ploidy level |
|---|---|---|---|---|---|---|
| | | ETS | ITS | *trn*L/F | *trn*S/G | |
| Karis 412 (S) | | U90786[a] | | | | |
| *Sibbaldiopsis tridentata* (Aiton) Rydb. | M. Lundberg 2 (S)<br>Eriksson & Smedmark 40 (S)<br>Hill 17146 (A) | FJ422379 | U90791[a] | AJ512236[b] | FJ422341 | Tetraploid |

[a] Eriksson et al. (1998).
[b] Eriksson et al. (2003).
[c] Potter et al. (2007).

of strictly applied criteria. (2) Expand the amount of data and sampling compared to earlier studies to test the monophyly of various genera. (3) Improve the phylogenetic backbone of Fragariinae. (4) Place the remaining *Potentilla* species that were suspected to be nested within Fragariinae.

## 2. Material and methods

### 2.1. Taxa

Taxon sampling was based on Eriksson et al. (2003) where the monophyly of Fragariinae was established with good support. Additional taxa were selected in order to increase the sample in general, and to increase the number of representatives of each genus in Fragariinae to be able to further test the monophyly of all Fragariinae genera. Twenty new ingroup taxa and three new outgroup taxa were added to a total of 40 taxa. Two species, *Potentilla cuneata* and *P. miyabei*, traditionally classified in *Potentilla* were included. In the monograph of *Potentilla* (Wolf, 1908) these two species were classified close to *Potentilla tridentata*, currently *Sibbaldiopsis*, shown by Eriksson et al. (1998) to belong to Fragariinae. In addition, *P. caulescens*, *P. lignosa* and *P. alchemilloides* from early diverging lineages of the *Potentilla* clade were sampled (Töpel et al., unpublished). Outgroup taxa were selected based on the results from Eriksson et al. (1998, 2003). Trees were rooted with *Rosa*, which is sister to the Sanpotina clade (= *Potentilla* + Fragariinae + Sanguisorbeae; Eriksson et al., 2003). Taxon names, voucher information, GenBank accession numbers and ploidy levels of taxa are given in Table 2.

### 2.2. Methods

We used 0.02–0.05 g plant material from silica dried leaves, herbarium material or fresh leaves. Plant material was homogenised in CTAB using a mini-beadbeater (Biospec products) and silica beads. The extraction procedure of total DNA followed a modified CTAB protocol (Doyle and Doyle, 1990) with minor modifications. Polymerase chain reactions (PCR) of 25 μL were set up by mixing 0.5 units (U) Taq DNA polymerase (Roche Applied Science, Germany), 1× PCR reaction buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.04% Bovine serum albumin (BSA), 0.3 μM of forward and reverse primers respectively, $ddH_2O$ and 1 μL of total DNA. PCR reactions were amplified through a 5 min initial denaturation at 95 °C, followed by 35 cycles of 94 °C denaturation for 1 min, annealing at 50 °C for 1 min, and extension of amplification products at 72 °C for 2 min. A final extension was performed at 72 °C for 7 min. Amplifications were carried out on an Eppendorf Mastercycler gradient or a Techne Genius PCR system. PCR reactions were checked for products by separation on a 1% agarose gel, stained in EtBr and visualised under UV-light. Amplification products were cleaned using Montage $PCR_{96}$ plates (Millipore) and a vacuum-manifold. Sequencing reactions of 10 μL, using BigDye V. 3.1 or BigDye V. 3.0 (Applied Biosystems) were set up by mixing 1 μL Terminator Ready Reaction mix, 1× BigDye Sequencing buffer, 0.16 μM primer $ddH_2O$ and 1–5 μL template. Sequence reactions were run on a GeneAmp PCR System 9700 (Applied Biosystems) with an initial denaturation at 96 °C for 1 min, followed by 25 cycles of 96 °C denaturation for 10 s, annealing at 50 °C for 5 s, and extension for 4 min at 60 °C. Extension products were visualised on a 3100 Genetic Analyzer (Applied Biosystems). The software 'Phred' version 0.020425.c (Ewing et al., 1998; Ewing and Green, 1998), run under the Staden package version 2003.0b1 (Staden, 1996) was used to obtain base scorings and 'Phrap' version 0.990319 (Green, 1996) was utilised to assemble contigs.

Problematic taxa were amplified with the high-fidelity and proof-reading Phusion™ high-fidelity DNA polymerase (Finnzymes, Finland). PCR reactions of 25 μL containing 0.5 U Phusion DNA polymerase, 1× Phusion HF Buffer, 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.3 μM of each primer, $ddH_2O$ and 1 μL of total DNA. PCR started with an initial denaturation of 98 °C for 2 min followed by 35 cycles of 10 s at 98 °C, 25 s at 60 °C and 45 s at 72 °C. A final extension for 10 min at 72 °C ended the amplification. When amplification yielded more than one product purification by gel based methods was applied (QiAquick Gel Extraction or Gel-M Gel extraction System) according to provided protocols.

Single nucleotide polymorphism was detected by studying chromatograms by eye. Observed polymorphisms were noted and treated as unknown characters in analyses. Sequence alignments of the separate regions were made by hand using Seaview (Galtier et al., 1996). All positions were included. Indels were not coded as separate characters.

### 2.3. Amplification and sequencing primers

The external transcribed spacer (ETS) region of 18S-26S nuclear ribosomal DNA was amplified and sequenced using the primers ETS1 and IGS6 (Oh and Potter, 2005). Amplification of the internal transcribed spacer (ITS) regions of the 18S-26S nuclear ribosomal DNA including the 5.8S subunit were performed using the primers ITS-I (Urbatsch et al., 2000) and ITS4 (White et al., 1990), and the amplification primers, along with the internal primers ITS2 (White et al., 1990) and ITS3B (Baum et al., 1994) when necessary, were used for sequencing. The spacer between *trn*L and *trn*F and the intron in the *trn*L gene of plastid DNA was amplified using the primers *trn*-C and *trn*-F and sequenced with the amplification primers and the internal primers *trn*-D and *trn*-E (Taberlet et al., 1991). The primer pair $trnS^{GCU}$/3′$trnG^{UUC}$ (Shaw et al., 2005) was used to amplify the *trn*S-*trn*G spacer and the *trn*G gene intron of plastid DNA. The internal primers 5′*trn*G2G and 5′*trn*G2S as well as the amplification primers were used for sequencing. Some taxa amplified poorly or not at all using the published primers. Therefore, an additional primer, *trn*Ga (TAGTCGACGTTGATTCATCA) was constructed based on existing sequences. The *trn*Ga primer is located 44 bp inside of the 3′$trnG^{UUC}$ primer. All regions were sequenced in both directions. All primers used were obtained from MWG Biotech AG, Germany.

### 2.4. Phylogenetic analyses

Each region (ETS, ITS, *trn*L/F and *trn*S/G) was analysed separately and in combination: nuclear (ETS + ITS), plastid (*trn*L/F + *trn*S/G), and a joint combined analysis of all four regions. In total, seven matrices were analysed. The model selection was based on the corrected AICc criterion (Sugiura, 1978; Hurvich and Tsai, 1989) in MrAIC 1.4.3 (Nylander, 2004) which uses PHYML (Guindon and Gascuel, 2003), and all matrices favoured the general time reversible model (GTR; Tavaré, 1986) with gamma distribution of rates (Yang, 1993). Phylogenies were estimated using Bayesian inference (Yang and Rannala, 1997) with MrBayes 3.1.2. (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). For each matrix, two parallel Bayesian analyses were performed starting from random trees. The default values in MrBayes were used for priors on model parameters and MCMC proposal parameters. Each analysis ran four MCMC chains for four million generations, sampling from the best chain every 1000 generation (4001 trees per analysis). A chain burnin of the initial 50% (see Section 3) of the samples was discarded (for each analysis) and a majority-rule consensus tree was obtained for the remaining 4002 trees, pooled from both analyses.

### 2.5. Comparison of topologies

Three criteria had to be fulfilled in order for considering topological incongruence between the nuclear and plastid phylogenies to potentially be caused by allopolyploid speciation: (1) Incongruences should be supported by a clade credibility of 0.95 or higher. (2) The strength of conflict should be at least 0.90. This was measured as the absolute value of the frequency of occurence of the node seen in one Mrbayes tree sample (after burnin), minus the frequency of occurence in the same tree sample for the conflicting node (i.e. the node seen in the other tree sample). The topological incongruences were checked in the Bayesian tree sample of each data set (i.e. nuclear and plastid). This estimated the support for the incongruence itself (strength of conflict), ruling out the possibility that it was caused by misleading "signal" in the data, or that there was substantial support also for the conflicting node. As an example, if the frequency for a node in the nuclear tree sample was 0.95 (CC = 0.95) and the frequency for the conflicting node (seen in the plastid tree sample) was 0.05, the strength of conflict would be 0.90—hence accepted. Calculations were performed by using a Perl script ('nodescan.pl') by T. Eriksson. (3) Convergence diagnostics should not indicate that the incongruences were caused by the lack of convergence of the individual MCMC runs. We monitored the split frequencies for the first and second run of the Bayesian analyses, and the potential scale reduction factor (PSRF; Gelman and Rubin, 1992) for substitution-model parameters as reported by MrBayes. In addition, we graphically monitored the $\hat{R}$-interval (Brooks and Gelman, 1998) for branch lengths using the software AWTY (Nylander et al., 2008).

## 3. Results

We obtained 37 ETS sequences varying from 467 (*Aphanes arvensis*) to 484 (*Comarum salesovianum*) base pairs in length and the alignment consisted of 503 positions. We were unable to obtain sequences from *Aremonia agrimonoides*, *Potentilla caulescens* and *Dasiphora phyllocalyx*. Amplification of *Aremonia agrimonoides* produced several bands of unexpected length and ambiguous sequence similarity when blasted. We were unable to amplify *Potentilla caulescens* despite good quality DNA. A sequence of *Dasiphora phyllocalyx* was obtained, but with ambiguous quality and it was therefore excluded. Polymorphic sites were detected for *Dasiphora davurica* (1 site), *Potentilla alchemilloides* (1 site), *Potentilla lignosa* (7 sites) and *Sibbaldianthe bifurca* (3 sites). In total, 0.06% of the aligned ETS sites were polymorphic. For ITS, we obtained 24 sequences and used 16 from previous studies (Eriksson et al., 1998, 2003; Potter et al., 2007). They varied in length between 589 (*Aremonia agrimonoides*) and 627 (*Fragaria viridis*) base pairs. The aligned ITS dataset consisted of 40 taxa with 656 aligned positions. Polymorphic sites were detected for *Dasiphora davurica* (3 sites), *Fragaria chiloensis* (3 sites), *Fragaria moschata* (1 site), *Fragaria viridis* (1 site), *Fragaria x ananassa* (3 sites), *Potaninia mongolica* (1 site), *Potentilla alchemilloides* (2 sites), *Potentilla caulescens* (1 site) and *Potentilla lignosa* (4 sites). In total, 0.07% of the aligned ITS sites were polymorphic. For *trn*L/F, we obtained 22 sequences and used 16 from previous studies (Eriksson et al., 2003; Potter et al., 2007). The sequences varied between 915 (*Chamaerhodos nuttallii*) and 1047 (*Rosa majalis*) base pairs in length. We were unable to amplify *trnL/F* for *Chamaerhodos sabulosa* and *Potentilla caulescens*. The aligned *trn*L/F dataset consisted of 38 taxa with 1258 aligned positions. We obtained 37 *trn*S/G sequences. They varied between 878 (*Chamaerhodos mongolica*) and 1371 (*Fragaria moschata*) base pairs in length. *Chamaerhodos mongolica* had a 468 base pairs long deletion compared to *Chamaerhodos erecta*. We were unable to amplify *Chamaerhodos nuttallii*, *Chamaerhodos sabulosa* and *Potentilla miyabei*. The aligned *trn*S/G dataset consisted of 37 taxa and 1634 aligned positions.

### 3.1. Clades and incongruencies

When comparing nuclear and plastid phylogenies of Fragariinae, the topologies were congruent to a large extent. However, five well supported incongruencies were found. These involve *Fragaria vesca* and *F. moschata* (A-nodes in Fig. 1), *Drymocallis* (B-nodes in Fig. 1), *Aphanes arvensis* (C-nodes in Fig. 1), *Potentilla cuneata* (D-nodes in Fig. 1) and *Potentilla miyabei* (E-nodes in Fig. 1). All five nodes passed the three criteria for conflict including high scores for strength of conflict (Table 3). The MCMC convergence diagnostics indicated that the separate chains approximated the same target distribution (Fig. 2a). It is notable that $\hat{R}$-interval indicated by the treelengths (sum of all branch lengths) converged at a later stage (Fig. 2b and c) than would have been indicated by e.g. judging a traceplot of the log likelihood values (see also Nylander et al., 2008). Therefore, an initial burnin of 50% (2000 trees) was discarded from each of the nuclear, plastid and combined analyses, respectively.

Unless specified, all clades discussed below were supported by clade credibility (CC) values of 0.95 or higher. We consider a CC value of 0.95 or higher to indicate a well supported node when analyses converge properly (Alfaro and Holder, 2006).

#### 3.1.1. Nuclear sequence topology

The separate nuclear phylogenies of ETS and ITS (not shown) did not show any well supported incongruences. Hence, only the joint nuclear phylogeny (ETS + ITS; Fig. 1a) will be the subject of further discussion. The joint nuclear dataset consisted of 40 taxa and 1159 aligned DNA positions. Fragariinae is monophyletic with high support CC = 1.0 (node F, Fig. 1a), and consists of three main clades but the relationships among these three clades are not resolved. The three main clades are (1) *Fragaria*; (2) a clade comprising *Dasiphora*, *Potaninia*, *Drymocallis* and *Chamaerhodos*; and (3) a clade comprising the rest of Fragariinae: *Alchemilla*, *Aphanes*, *Comarum*, *Sibbaldia*, *Potentilla cuneata*, *P. miyabei*, *Sibbaldianthe* and *Sibbaldiopsis*.

*Fragaria* is monophyletic with well resolved internal relationships. In this tree, *F. chiloensis* and *F. x ananassa* form a clade together with *F. virginiana* as sister. *F. moschata* and *F. vesca* form a clade that has a sister relationship to the former clade. *F. viridis* is sister to the rest of *Fragaria* in the current sampling.

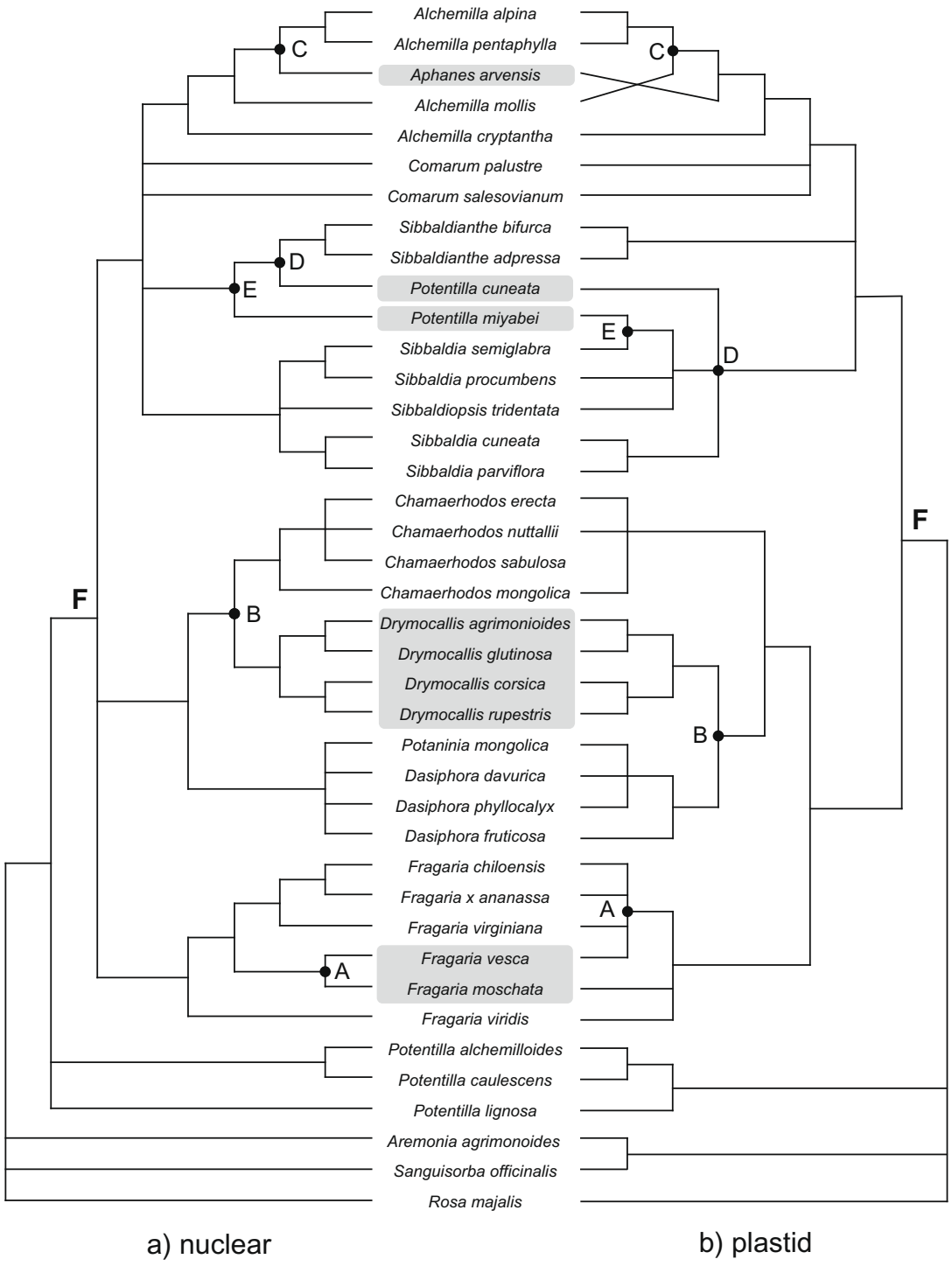


**Fig. 1.** (a) Majority-rule consensus tree, 95%, from Bayesian analysis of nuclear ETS and ITS in combination. The tree is based on 4002 trees pooled and sampled from two separate analyses after the chain burnin was removed. (b) Majority-rule consensus tree, 95%, from Bayesian analysis of plastid *trn*L/F and *trn*S/G in combination. The tree is based on 4002 trees pooled and sampled from two separate analyses after the chain burnin was removed. Nodes marked with capital letters are discussed in the text and F indicates the clade Fragariinae.

*Chamaerhodos* is monophyletic with *C. mongolica* as sister to the rest of the included species. *Drymocallis* is also monophyletic and divided into two clades consisting of *D. agrimonioides* and *D. glutinosa* in one clade, and *D. corsica* plus *D. rupestris* in the other. The sampled species of *Dasiphora* form an unresolved clade together with *Potaninia mongolica*. Here, the *Dasiphora + Potaninia* clade is sister to a *Drymocallis + Chamaerhodos* clade.

The third main clade contains most of the Fragariinae species. It consists of three subclades plus two unresolved *Comarum* species. *Aphanes*, here represented by a single species only, is nested inside *Alchemilla* as sister to *A. alpina* and *A. pentaphylla* comprising a clade. *A. mollis* is sister to the previous three species, and the Malagasy species *A. cryptantha* is sister to the rest of the *Alchemilla + Aphanes* species sampled here. Another well supported clade contains *Sibbaldianthe* with *Potentilla cuneata* and *Potentilla miyabei* as a "sister grade" outside. The included *Sibbaldia* species form a well supported clade together with the monotypic *Sibbaldiopsis*,

**Table 3**
Frequencies for presence of conflicting nodes in the tree samples of separate Bayesian analyses of nuclear and plastid data sets, respectively. Trees sampled during MCMC chain burnin were excluded.

| Clade | Nodes (cf. Fig. 1) | Occurence | Strength of conflict |
|---|---|---|---|
| *Fragaria* | | | |
| Nuclear trees | 1a A/1b A | 0.996/0.002 | 0.994 |
| Plastid trees | 1b A/1a A | 0.979/0.000 | 0.979 |
| *Drymocallis* | | | |
| Nuclear trees | 1a B/1b B | 0.996/0.000 | 0.996 |
| Plastid trees | 1b B/1a B | 1.000/0.000 | 1.000 |
| *Aphanes* | | | |
| Nuclear trees | 1a C/1b C | 0.994/0.004 | 0.990 |
| Plastid trees | 1b C/1a C | 0.990/0.005 | 0.985 |
| *Sibbaldianthe* | | | |
| Nuclear trees | 1a D/1b $D_2$ | 1.000/0.000 | 1.000 |
| Plastid trees | 1b $D_2$/1a D | 1.000/0.000 | 1.000 |
| *Potentilla miyabei* | | | |
| Nuclear trees | 1a E/1b E | 1.000/0.000 | 1.000 |
| Plastid trees | 1b E/1a E | 0.957/0.000 | 0.957 |

but the internal relationships are not fully resolved. *Sibbaldia cuneata* and *S. parviflora* group together as well as *S. procumbens* and *S. semiglabra*. Collapsed here, but with a CC of 0.94 is a more inclusive clade joining *Sibbaldia, Sibbaldiopsis, Sibbaldianthe* and the two orphan *Potentilla* species.

### 3.1.2. Plastid sequence topology

As in the nuclear analyses, the two plastid regions did not have any well supported incongruences. We therefore only discuss the joint plastid phylogeny (*trn*L/F + *trn*S/G; Fig. 1b). The joint plastid data set consisted of 39 taxa *(Chamaerhodos sabulosa* missing) and 2892 aligned DNA positions. Fragariinae was found to be monophyletic and well supported, CC = 1.0 (node F, Fig. 1b). The resolution near the root of Fragariinae is somewhat better than in the nuclear tree. The resolution in the rest of the tree is very similar to that of the nuclear tree, but here *Fragaria* is sister to the clade of *Dasiphora*, *Potaninia*, *Drymocallis* and *Chamaerhodos*.

*Fragaria* is monophyletic but less well resolved internally than in the nuclear tree. *Fragaria chiloensis*, *F. vesca*, *F. virginiana*, and *F. x ananassa* constitute a collapsed clade with *F. moschata* and *F. viridis* as unresolved sisters. The nuclear tree had a well supported clade of *F. vesca + F. moschata* which is contradicted here.

*Chamaerhodos* is monophyletic while lacking internal resolution, and in conflict with the nuclear tree, here as sister to a clade of *Drymocallis + Dasiphora + Potaninia*. *Drymocallis* forms a clade that is sister to that of *Dasiphora + Potaninia*. The internal relationships of *Drymocallis* are identical to those of the nuclear tree. In this tree, the monotypic *Potaninia mongolica* forms a clade together with two *Dasiphora* species (*D. davurica* and *D. phyllocalyx*), with *Dasiphora fruticosa* as sister.

Like in the nuclear tree, the rest of the Fragariinae species form a well supported clade, but the internal relationships differ. Here, the two *Comarum* species join the *Alchemilla + Aphanes* clade as unresolved sisters. The resolution of the *Alchemilla + Aphanes* clade is similar to that in the nuclear tree, but although *Aphanes* is still nested within, it has switched places with *A. mollis*. *Sibbaldianthe* is monophyletic and is here found in an unresolved trichotomy with the previously described clade and the following, that comprises *Sibbaldia, Sibbaldiopsis, Potentilla cuneata*, and *P. miyabei*. In the latter clade, *P. miyabei* and *Sibbaldia semiglabra* are sisters, and together with *S. procumbens* and *Sibbaldiopsis tridentata* they form an unresolved clade. The relationship of *Potentilla cuneata* differ significantly in this plastid tree compared with the nuclear tree whereas *Sibbaldia cuneata* and *S. parviflora* are closest relatives and form a trichotomy with *Potentilla cuneata* and the previous clade.

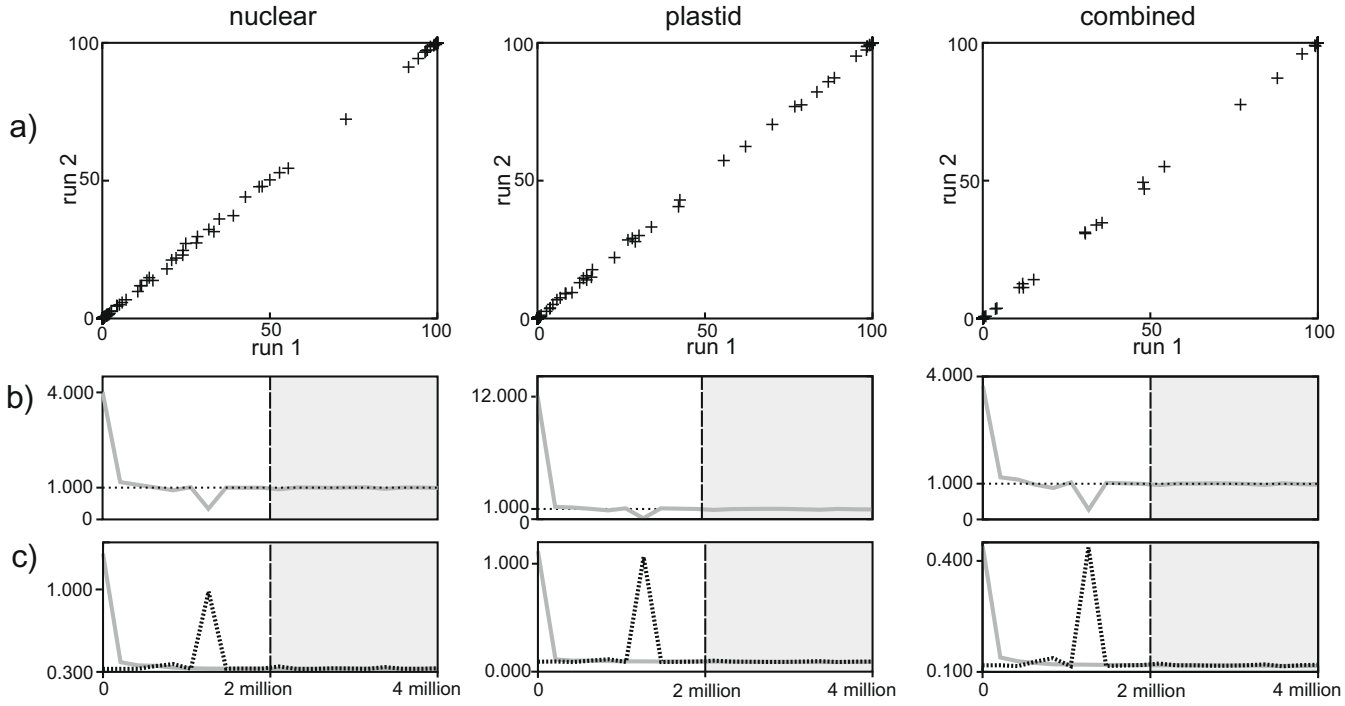


**Fig. 2.** Graphical output from AWTY analysis. Graphs show output from the three different data sets analysed by MrBayes: nuclear (left column), plastid (middle column), and combined (right column). Row (a) shows a bivariate plot of the split frequencies from the two separate runs. Row (b) shows the $\hat{R}$-interval, which compares the 95% interval of credibility between the runs with the 95% interval of credibility of the pooled runs. As $\hat{R}$-interval approach 1 (dotted line), the chains converge. Units of the *Y*- and *X*-axis are $\hat{R}$ (Treelength)-interval and MCMC generations, respectively. The white area marks the bounds of burnin (50%) of the tree sample discarded. Row (c) grey line shows the length of total chain interval while the dotted line show the mean length of the within chain interval. Units of the *Y*- and *X*-axis are $\hat{R}$-interval and MCMC generations, respectively. The white area indicate the bounds of the burnin (50%) of the tree sample discarded.

*3.1.3. Combined nuclear and plastid data*

The topology from the combined analysis has with a few exceptions strongly supported nodes throughout (Fig. 3). The combined data set consisted of 40 taxa and 4051 aligned DNA positions. The combined tree has the same main clades as the nuclear and the plastid trees. There are some improvements in resolution and support as well as nodes that suffer from conflict. The main notable difference in resolution is found in the clade of *Sibbaldia* et al. Fragariinae as a whole is monophyletic with representatives of *Potentilla* in the strict sense as sister clade.

*Fragaria* is resolved as in the nuclear tree which means that the position of *F. vesca* (or *F. moschata*) contradict the plastid resolution of *Fragaria*. As in the plastid tree, *Fragaria* is sister to a clade comprising *Chamaerhodos, Drymocallis*, *Dasiphora*, and *Potaninia*.

*Chamaerhodos* is resolved like in the nuclear tree (the plastid tree was unresolved). Also in accordance with the nuclear tree but contrary to the plastid tree, *Chamaerhodos* is sister to *Drymocallis*. *Dasiphora* includes *Potaninia* and is resolved similarly to the plastid tree. This clade is sister to that of *Chamaerhodos* + *Drymocallis*.

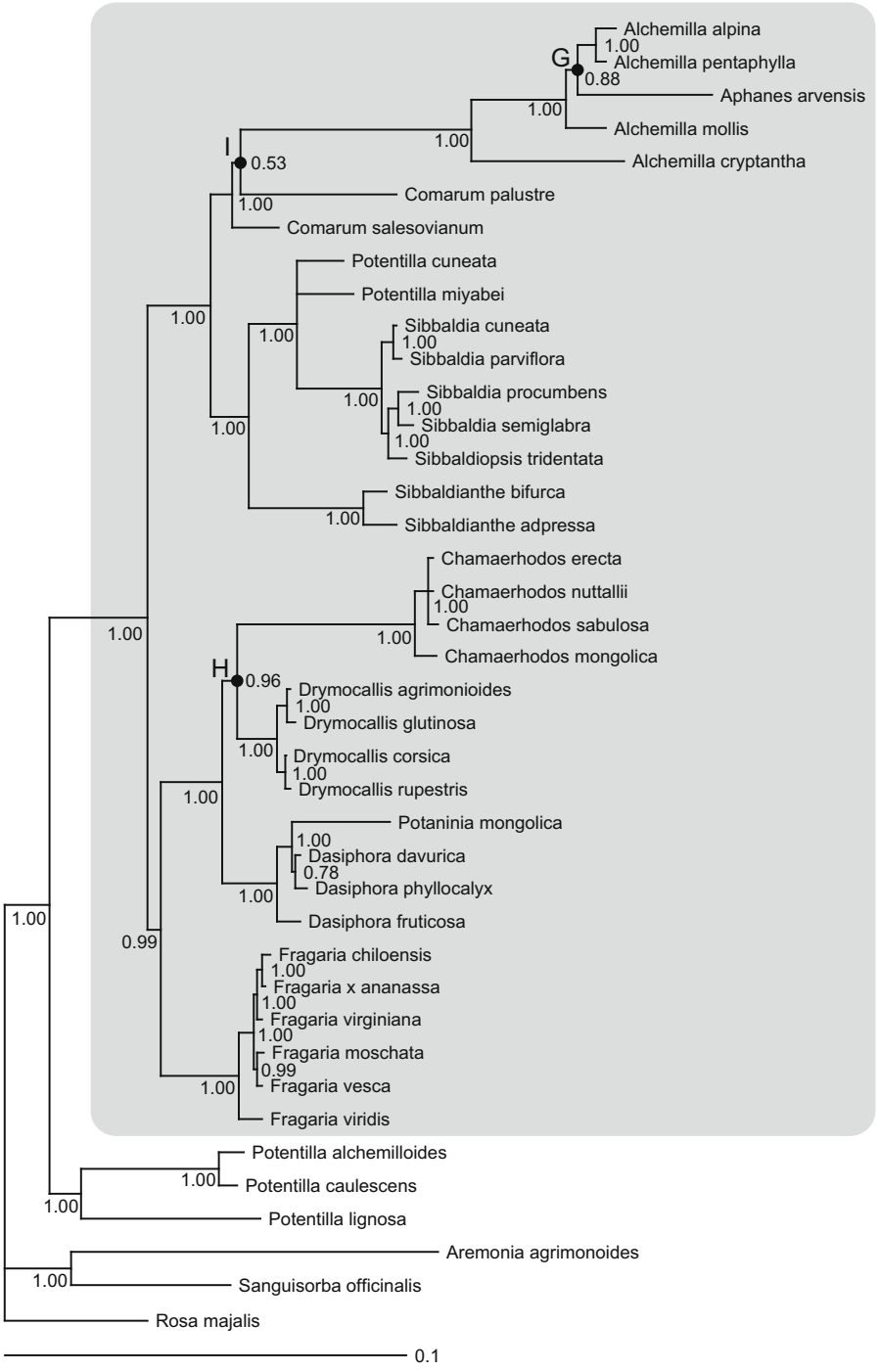


**Fig. 3.** Majority rule consensus tree (50%) from the Bayesian inference analysis of the four regions ETS, ITS, *trn*L/F and *trn*S/G in combination. Shaded area delimits the clade Fragariinae. Bayesian clade credibilities are indicated at the nodes. G-I refer to nodes mentioned in the discussion. The branch lengths are mean lengths calculated from 4002 trees. Scale bar = 0.1 substitutions per site.

*Alchemilla* includes *Aphanes* and is resolved like in the nuclear tree. However, the support for node G (Fig. 3) has decreased.

The resolution within the clade of *Sibbaldia* et al. differs from both the nuclear and the plastid tree. Similarly to the plastid tree, *Sibbaldianthe* is outside of the other species. In the combined tree, though, it is sister to a clade of *Sibbaldia, Sibbaldiopsis*, and *Potentilla cuneata* plus *P. miyabei*. This is not in contradiction to the plastid tree because it is merely unresolved. The nuclear tree, however, has a different topology. In the combined tree, the *Sibbaldia* species form a clade with *Sibbaldiopsis* inside, and the two *Potentilla* species as unresolved sisters. The resolution of *Sibbaldia* resembles that of the nuclear tree but the position of *Sibbaldiopsis* is here resolved inside *Sibbaldia*.

## 4. Discussion

### 4.1. Incongruences between nuclear and plastid phylogenies

The main purpose of the present study was to detect potential events of allopolyploidy in the phylogenetic history of Fragariinae by contrasting phylogenies based on maternally inherited plastid sequence data (*trn*L/F plus *trn*S/G regions), and biparentally inherited ribosomal DNA sequence data (ITS plus ETS). Results from this study will subsequently allow a more detailed investigation of the most promising subclades using low copy nuclear sequence data. We found five incongruences, that meet our criteria of incongruence. Due to the processes of concerted evolution in nuclear ETS and ITS regions, the sole use of topological incongruence for detecting allopolyploid speciation will not guarantee detection of all allopolyploid candidates. Regions affected by concerted evolution may homogenise gene copies to one of the parental sequences, or occasionly to a mix. If the process of homogenisation has converged to the paternal sequence, phylogenies will indicate incongruence but if the maternal sequence is fixed no incongruence can be observed. Topological differences can also be caused by incomplete lineage sorting, independent gene duplications, random loss in multiple genes, horizontal gene transfer, heterotachy and selection, but although such incongruences might be hard to distinguish from allopolyploidy in the current setup, further studies using low copy genes may enable us to do this. Insufficient sampling of taxa and the amount of data might also affect the topology or the support for any potential incongruences. The nucleotide polymorphisms detected in the nuclear ribosomal DNA sequences were treated as unknown character states (“N”). Because of this and since they were very limited in number (in total 0.06% of nuclear DNA), we do not expect them to have affected the tree topologies. After performing convergence diagnostics we conclude that none of the incongruences detected are the result of inaccurate Bayesian MCMC analyses. Output from AWTY analyses in this study show that the independent runs of each dataset are close in parameter (tree) space, thus we can conclude that the two separate runs approximated the same target (tree) distribution.

In the following sections we discuss the clades causing the observed incongruence between nuclear and plastid topologies (nodes A–E in Fig. 1).

A previous phylogeny of 14 *Fragaria* species based on ITS and *trn*L/F (Potter et al., 2000) is almost fully congruent with our *Fragaria* clade. Our plastid tree (node A, Fig. 1b) like their *trn*L/F tree indicates that *F. vesca, virginiana, chiloensis* and *x ananassa* form a clade with *F. moschata* and *viridis* unresolved outside. In our ETS + ITS tree as well as in the ITS tree of Potter et al. (2000) *F. viridis* is sister to the rest of the species (based on our sample). There are differences in resolution between the ITS tree of Potter et al. (2000) and our ETS + ITS tree that might be an effect of differential sampling and/or different amounts of data. It is clear, however, that both studies show a well supported incongruence involving *F. vesca* or *moschata* (node A, Fig. 1a). In the *Fragaria* clade, the ploidy level of all 20 species is known and every ploidy level from diploid to octoploid is represented (Table 2; Potter et al., 2000). *Fragaria* is well known morphologically and geographically, and there are numerous reports of hybrid varieties in nature and cultivated strawberries are known to be the result of interspecific hybridisation (Staudt, 1962; Bringhurst and Gill, 1970). The incongruence in *Fragaria* combined with the ploidy levels makes this clade highly interesting for further study using nuclear single/low copy genes. Some phylogenetic results based on low copy genes are currently being published (T. Davis, pers. comm.; Rousseau-Gueutin et al., in press).

In the nuclear tree, the *Chamaerhodos* clade is sister to the *Drymocallis* clade (node B, Fig. 1a) whereas in the plastid tree, the *Drymocallis* clade is sister to the *Dasiphora + Potaninia* clade (node B, Fig. 1b). As far as is known, there are only diploid species in *Chamaerhodos* and *Drymocallis*. In *Dasiphora* diploid and tetraploid taxa are naturally occurring while among cultivated Asian individuals, also hexa- and octoploidy have been reported (Table 2; Klackenberg, 1983). The *Dasiphora fruticosa* sampled in this study is assumed to be tetraploid because of dioecous flowers and its geographical origin in Sweden (Klackenberg, 1983). In order to explain the incongruence involving *Drymocallis* by allopolyploid speciation, an allopolyploid ancestor should have existed that went through extensive diploidisation, finally reaching back to a diploid ploidy level. Another explanation might be that the incongruence is due to a homoploid hybridisation event involving two diploid species. For instance, a regular sexual event between taxa of *Chamaerhodos* and *Dasiphora* took place in which each gamete had a haploid complement of the nuclear chromosomes of each parent, but gametes that formed the zygote came from different genera. In order for the hybrid diploid to persist, the hybrid between *Chamaerhodos* and *Dasiphora* should have had a partial fertility and been isolated from parental species in order not to backcross. In a diploid hybridisation event, the diversification of *Drymocallis* took place after a probable single hybridisation. Yet another potential scenario might occur if the female parent was an ancestor of *Dasiphora* and the male parent was an ancestor of *Chamaerhodos*. Following extensive (male) backcrossing of *Chamaerhodos* into the hybrid population, the male nuclear DNA might be homogenised towards *Chamaerhodos*, while the plastid would still be from the female parent. In the light of such complex hypothetical explanations, this clade would not be among our first choices for further studies using nuclear single/low copy gene phylogenies.

In a recent study of *Alchemilla* in the wide sense, Gehrke et al. (2008) presented relationships of *Alchemilla, Aphanes*, and *Lachemilla*. Their analyses, based on sequence data from ITS and *trn*L/F regions for a sample of 85 ingroup taxa, supported five clades: *Aphanes, Lachemilla* and three major clades of *Alchemilla* species. Our small set of taxa from this group has representatives from four of these clades (not *Lachemilla*): *A. alpina* and *pentaphylla* from the “dissected clade”, *A. mollis* from the “lobed clade”, *A. cryptantha* from the “Afromilla clade”, and *Aphanes arvensis* from the *Aphanes* clade. Disregarding *Lachemilla*, the relationships of these major clades in their phylogenies were the same and congruent with our plastid phylogeny where *Aphanes arvensis* is sister to a clade including *Alchemilla alpina*, *A. pentaphylla* and *A. mollis* (node C, Fig. 1b). However, our ITS + ETS phylogeny differs in the position of *Aphanes* (node C, Fig. 1a). Although our analyses suffer from a small *Alchemilla* sample as compared to the sample of Gehrke et al., we do have more sequence data. The basal chromosome number of the *Alchemilla* clade including *Aphanes* and *Lachemilla* is $x = 8$. There are no known diploid species in the three *Alchemilla* clades, only higher ploidy levels from tetraploid up to $2n = ca224$ (28-ploid) (Table 2; Kalkman, 2004). The *Aphanes* clade includes representatives that are diploid, tetraploid and hexaploid while

ploidy levels are not known in *Lachemilla*. Although many ploidy levels are represented in *Alchemilla*, recent hybridisation events are considered rare because of the absence of diploids and that they reproduce through autonomous apomixis (Fröhner, 1995). It is notable in the analysis of Gehrke et al. (2008) that the position of *Lachemilla* is incongruent when comparing nuclear and plastid phylogenies. This, and the presence of diploids in an otherwise polyploid clade makes us consider this clade to be of great potential from an allopolyploidy point of view.

In the nuclear tree, *Potentilla cuneata* is sister to *Sibbaldianthe* (node D, Fig. 1a), but in the plastid tree *Potentilla cuneata* is instead sister to a clade which includes *P. miyabei* and *Sibbaldia* species (node D, Fig. 1b). *Potentilla cuneata* is a probable candidate for being of allopolyploid origin due to this incongruence and because of established polyploidy. *Potentilla cuneata* and its relationship to *Sibbaldianthe* would be an intriguing set of species to investigate using nuclear low copy gene phylogenies. Especially since the pattern of conflict is complex.

*Potentilla miyabei* is a rare endemic from Japan, sister to *Sibbaldia semiglabra* in the plastid tree (node E, Fig. 1b) but forming a clade with, and sister to, *Sibbaldianthe* and *Potentilla cuneata* in the nuclear tree (node E, Fig. 1a). *Potentilla cuneata* is a known tetraploid (Table 2), hence there is some evidence that polyploidy has influenced this group. Taking the incongruent topologies into account, allopolyploidy seems probable. Better sampling is difficult to achieve since there are not many species in this group. Knowledge of the ploidy level of *Potentilla miyabei* and all the *Sibbaldia* taxa including close allies would help elucidate potential patterns of incongruences and allopolyploid evolution in the group. It would be a good group to focus on to perform a nuclear single/low copy phylogeny study.

*Sibbaldiopsis tridentata* is a tetraploid and nested within *Sibbaldia* where two incongruences were found in this study. We think that it may be necessary to involve the entire *Sibbaldia* clade to be able to elucidate the evolution of these species.

Even though five well supported incongruences were found, we might expect to find more if our taxon sampling were denser and if primers that targets easily amplified nuclear markers unaffected by concerted evolution were known to us. It is a problem that for many of the species investigated the ploidy level is not yet known. For instance, there is evidence of polyploidy in *Comarum palustre*, but due to low resolution in that part of the tree it is difficult at this stage to hypothesise through which pathway *Comarum palustre* reached a hexaploid level. Using low copy sequences of *Comarum* might indicate possible allopolyploid evolution. Another example is the *Dasiphora* clade where diploid to octoploid ploidy levels are known but due to low resolution no hypothesis involving possible allopolyploid speciation could be drawn. In addition, the inclusion of the morphologically distinct *Potaninia mongolica* in *Dasiphora* would be especially interesting if *Potaninia* was shown to be polyploid.

### 4.2. Phylogeny of combined analysis and taxonomic implications

We decided to estimate a joint tree based on a combined data set of all our data despite the fact that the separate phylogenies show five well supported incongruences. This decision is partly based on the idea that most nodes are congruent between the separate trees, and for such nodes a combined approach is reasonably the best. A combined analysis may also provide additional support for increased resolution (e.g. Olmstead and Sweere, 1994) and allow us to compare how support for nodes changes—as compared to support in the separate trees. If the support for a node involved in an incongruence decreases in the combined analysis, the confidence that the conflict in the data is real is strengthened, for example caused by allopolyploidy. If, on the other hand, one of the conflicting nodes shows increased support in the combined analysis it may be reasonable to question the conflict as perhaps the cause of sampling effects (data or taxa).

Our combined analysis indicate three nodes with lower support as compared to corresponding nodes in the separate data partitions. *Aphanes arvensis* is sister to a clade containing *Alchemilla alpina* and *Alchemilla pentaphylla*, CC = 0.88 (node G, Fig. 3) in the combined phylogeny, and CC = 0.99 in the nuclear phylogeny. The node joining the clades of *Chamaerhodos* and *Drymocallis* has a CC = 0.96 in the combined phylogeny (node H, Fig. 3), and CC = 1.00 in the nuclear phylogeny. The third is the node where *Comarum palustre* is sister to the *Alchemilla* clade has a CC = 0.53 in the combined phylogeny (node I, Fig. 3), and CC = 0.57 in the plastid phylogeny. The last of these three nodes is far from being well supported and the difference between the analyses is not large. However, for the two first clades, we think that the results of the analysis increase our confidence in the presence of conflict.

Five clades that received high support (CC = 1.0) in all our analyses, including the combined analysis, are notable because they have been classified as genera. The monophyly of these genera is not affected by any of the conflicts mentioned above: *Alchemilla* (when including *Aphanes* as suggested by Gehrke et al. (2008)), *Sibbaldianthe*, *Chamaerhodos*, *Drymocallis* and *Fragaria*. On the other hand, the increased sampling of Fragariinae in the present analysis, and the addition of the two DNA sequence regions ETS and *trn*S/G, made it possible to point out other "genera" in Fragariinae that may be polyphyletic. Also, two species that are still classified in *Potentilla* (*P. cuneata* and *miyabei*) were shown to belong to Fragariinae.

The combined phylogeny (Fig. 3) is with one exception congruent with the Fragariinae of Eriksson et al. (2003), but the tree in the present study has generally better resolution. Fragariinae as a whole is well supported in both studies. The only incongruence is that *Aphanes arvensis* is sister to *Alchemilla alpina* and *Alchemilla mollis* in the Eriksson et al. (2003) study similarly to the results of Gehrke et al. (2008), while in our study *Aphanes arvensis* is sister to *Alchemilla alpina* (and *Alchemilla pentaphylla*; the "dissected" clade) with *Alchemilla mollis* as sister to both of them but with low support (CC = 0.88). Regarding the generic classification of *Alchemilla* and close allies, there is nothing in our analyses that contradict Gehrke et al. (2008) who suggested to include *Aphanes* and *Lachemilla* in *Alchemilla*.

The resolution of the *Sibbaldia* clade (*Sibbaldia*, *Sibbaldiopsis*, *Sibbaldianthe*, *Potentilla cuneata* and *P. miyabei*) in the combined tree is different from either of the separate trees. The genus *Sibbaldianthe* consists as far as known only of two species, *S. adpressa* and *S. bifurca*, both sampled in this study and sometimes classified *Sibbaldia* and *Potentilla* (*Schizophyllidium*) (Kurtto and Eriksson, 2003; Eriksson et al., in preparation). The species of *Sibbaldianthe* differ from *Sibbaldia* and *Potentilla* based on the number of stamens. They form a well supported clade based on available sequence data. There seems to be little doubt that *Sibbaldianthe* is monophyletic and that *Sibbaldiopsis* is nested within *Sibbaldia*. It may be a good solution to include *Sibbaldiopsis* in *Sibbaldia*. However, in order to resolve the position of the two *Potentilla* orphans within this clade we need further data, preferably low copy nuclear sequence data. The combined phylogeny makes sense from a classification standpoint, giving *Sibbaldia* (including *Sibbaldiopsis*) good support. However, describing a new genus for the two unresolved *Potentilla* orphans would be improper since they do not form a clade. Also, taking the separate trees into account may indicate a hybrid origin for *Potentilla miyabei* and *P. cuneata*. Although it is not logical to keep classifying these two species in *Potentilla*, we think that it is prudent to resolve the question of hybrid origin before making classification changes. For example, it might be a relevant possibility to expand *Sibbaldia* to encompass *Potentilla cuneata* and

*miyabei*, but it depends upon the resolution of an analysis that take reticulations into account. A more thorough study of the *Sibbaldia* clade will be published elsewhere (Eriksson et al., in preparation).

*Comarum* forms a grade with respect to *Alchemilla* with *Comarum palustre* very weakly supported as closest sister to *Alchemilla*. The two *Comarum* species have many similarities, but also notable differences, and due to the lack of a morphological phylogenetic analysis it is not clear which of these are potential synapomorphies for *Comarum* (as a clade). The differences have resulted in *Comarum salesovianum* being classified as a separate unispecific genus, *Farinopsis* (Soják, 2004). If better sampling was possible, it might increase the resolution in this part of the tree, but due to the lack of species, adding more sequence data may be the only hope for resolving the issue of whether the species form a clade or not.

We note that although the small sample of *Drymocallis* species, only four out of about 30 species, the clade is divided into two geographically separate clades: *Drymocallis agrimonioides* and *Drymocallis glutinosa* (North America), and *Drymocallis corsica* and *Drymocallis rupestris* (Europe and North Africa). The unispecific *Potaninia* is nested within the *Dasiphora* clade. *Potaninia* is endemic to Mongolia and has distinctive trimerous flowers, distinct enough so that the genus never has been considered even close to *Dasiphora* (but see Potter et al., 2007). According to our results, classifying *Potaninia* in a separate genus would make *Dasiphora* paraphyletic, even though *Dasiphora* is morphologically homogenous. It therefore seems reasonable based on our current sample that *Potaninia mongolica* is included in *Dasiphora*.

## Acknowledgments

This study was funded by grants provided by Stockholm University to Magnus Lundberg, to Torsten Eriksson by the Swedish Research Council, VR grant 2004-1698. We acknowledge A. Khodabandeh for lab support, B. Gehrke for providing plant material, P. Östensson for unpublished sequences and J. Schönenberger for valuable comments of the manuscript. We are grateful to the following institutions for granting access to living and dry collections, as well as allowing us to sample the collections for DNA extraction purposes: The Bergius Botanic Garden, The Swedish Museum of Natural History (S) and Gothenburg Botanical Garden.

## References

Alfaro, M.E., Holder, M.T., 2006. The posterior and the prior in Bayesian phylogenetics. Annu. Rev. Ecol. Evol. Syst. 37, 19–42.

Alice, L.A., Eriksson, T., Eriksen, B., Campbell, C.S., 2001. Hybridization and gene flow between distantly related species of *Rubus* (Rosaceae): evidence from nuclear ribosomal DNA internal transcribed spacer region sequences. Syst. Bot. 26, 769–778.

Bandelt, H.J., Dress, A.W.M., 1992. Split decomposition: a new and useful approach to phylogenetic analysis of distance data. Mol. Phylogen. Evol. 1, 242–252.

Baum, D.A., Sytsma, K.J., Hoch, P.C., 1994. A phylogenetic analysis of *Epilobium* (Onagraceae) based on nuclear ribosomal DNA sequences. Syst. Bot. 19, 363–388.

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.

Bowers, J.E., Chapman, B.A., Rong, J., Paterson, A.H., 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422, 433–438.

Bringhurst, R.S., Gill, T., 1970. Origin of *Fragaria* polyploids. II. Unreduced and doubled-unreduced gametes. Am. J. Bot. 57, 969–976.

Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. 7, 434–455.

Darlington, C.D., 1937. Recent advances in Cytology, second ed. Churchill, London. pp. 183–229.

Doyle, J.J., Doyle, J.L., 1990. Isolation of plant DNA from fresh tissue. Focus 12, 13–15.

Eckhardt, N., 2001. A sense of self: the role of DNA sequence elimination in allopolyploidization. Pl. Cell 13, 1699–1704.

Eriksson, T., Donoghue, M.J., Hibbs, M.S., 1998. Phylogenetic analysis of *Potentilla* using DNA sequences of nuclear ribosomal internal transcribed spacers (ITS), and implications for the classification of Rosoideae (Rosaceae). Pl. Syst. Evol. 211, 155–179.

Eriksson, T., Hibbs, M.S., Yoder, A.D., Delwiche, C.F., Donoghue, M.J., 2003. The phylogeny of Rosoideae (Rosaceae) based on sequences of the internal transcribed spacers (ITS) of nuclear ribosomal DNA and the trnL/F region of chloroplast DNA. Int. J. Pl. Sci. 164, 197–211.

Ewing, B., Hillier, L., Wendl, M.C., Green, P., 1998. Base calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8, 175–185.

Ewing, B., Green, P., 1998. Base calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8, 186–194.

Focke, W.O., 1894. Rosaceae. In: Engler, A., Prantl, K. (Eds.), Die naturlichen Pflanzenfamilien III/3. Engelmann, Leipzig, pp. 32–36.

Fröhner, S.E., 1995. Alchemilla. In: Scholz, H., Conert, H.J., Jäger, E.J., Kadereit, J.W., Schultze-Motel, W., Wagenitz, G., Weber, H.E. (Eds.), Hegi: Illustrierte Flora von Mitteleuropa, 4 Teil 2B. Verlag Paul Parey, Berlin, Germany, pp. 13–242.

Galtier, N., Gouy, M., Gautier, C., 1996. SeaView and Phylo_win, two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12, 543–548.

Gehrke, B., Bräuchler, C., Romoleroux, K., Lundberg, M., Heubl, G., Eriksson, T., 2008. Molecular phylogenetics of *Alchemilla*, *Aphanes* and *Lachemilla* (Rosaceae) inferred from plastid and nuclear intron and spacer DNA sequences, with comments on generic classification. Mol. Phylogen. Evol. 47, 1030–1044.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences (with Discussion). Statist. Sci. 7, 457–511.

Goldblatt, P., Johnson, D.E., 2008. Index to plant chromosome numbers. Published on the Internet (http://mobot.mobot.org/W3T/Search/ipcn.html), accessed March 2008.

Grant, V., 1963. The Origin and Adaptations. Columbia University Press, New York. pp. 486.

Grant, V., 1981. Plant Speciation, second ed. Columbia University Press, New York. pp. 293–297, 324–328.

Green, P., 1996. Phrap. Computer software distributed by the author. Available from: <http://www.phrap.org/>.

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

Huber, K.T., Oxelman, B., Lott, M., Moulton, V., 2006. Reconstructing the evolutionary history of polyploids from multi-labelled trees. Mol. Biol. Evol. 23, 1784–1791.

Huelsenbeck, J.P., Larget, B., Miller, R.E., Ronquist, F., 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Syst. Biol. 51, 673–688.

Huelsenbeck, J.P., Ronquist, F.R., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Hurvich, C.M., Tsai, C.-L., 1989. Regression and time series model selection in small samples. Biometrika 76, 297–307.

Huson, D.H., 1998. SplitsTree: a program for analyzing and visualizing evolutionary data. Bioinformatics 14, 68–73.

Hutchinson, J., 1964. The genera of flowering plants. Dicotyledons, vol. 1. Clarendon Press, Oxford. pp. 195–208.

Kalkman, C., 1988. The phylogeny of Rosaceae. Bot. J. Linn. Soc. 98, 37–59.

Kalkman, C., 2004. Rosaceae. In: Kubitzki, K. (Ed.), The Families and Genera of Vascular Plants, Flowering plants—Dicotyledons: Celastrales, Oxalidales, Rosales, Cornales, Ericales, vol. 6. Springer, Berlin, pp. 343–386.

Kihara, H., Ono, T., 1926. Chromosomenzahlen und systematische Gruppierung der Rumex-Arten. Z. Zellf. Mikroskop. Anat. 4, 475–481.

Klackenberg, J., 1983. The holarctic complex *Potentilla fruticosa* (Rosaceae). Nordic J. Bot. 3, 181–191.

Kurtto, A., Eriksson, T., 2003. Atlas Florae Europaeae notes. 15. Generic delimitation and nomenclatural adjustments in Potentilleae (Rosaceae). Ann. Bot. Fenn. 40, 135–141.

Levin, D.A., 1983. Polyploidy and novelty in flowering plants. Am. Nat. 122, 1–25.

Lid, J., Lid, D.T., 1994. Norsk Flora, sixth ed. Det Norske Samlaget, pp. 267–328.

Lopez, P., Casane, D., Philippe, H., 2002. Heterotachy, an important process of protein evolution. Mol. Biol. Evol. 19, 1–7.

Mallet, J., 2007. Hybrid speciation. Nature 446, 279–283.

Mason-Gamer, R.J., 2001. Origin of North American *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. Syst. Bot. 26, 757–768.

Masterson, J., 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science 264, 421–423.

Morgan, D.R., Soltis, D.E., Robertson, K.R., 1994. Systematic and evolutionary implications of rbcL sequence variation in Rosaceae. Am. J. Bot. 81, 890–903.

Müntzing, A., 1936. The evolutionary significance of autopolyploidy. Hereditas 21, 263–378.

Nylander, J.A.A., 2004. MrAIC.pl. Computer software distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24, 581–583.

Oh, S.-H., Potter, D., 2005. Molecular phylogenetic systematics and biogeography of tribe Neillieae (Rosaceae) using DNA sequences of cpDNA, rDNA, and LEAFY. Am. J. Bot. 92, 179–192.

Olmstead, R.G., Sweere, J.A., 1994. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. Syst. Biol. 43, 467–481.

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

Popp, M., Erixon, P., Eggens, F., Oxelman, B., 2005. Origin and evolution of a circumpolar polyploid species complex in Silene (Caryophyllaceae) inferred

from low copy nuclear RNA polymerase introns, rDNA and chloroplast DNA. Syst. Bot. 30, 302–313.
Potter, D., Eriksson, T., Evans, R.C., Oh, S., Smedmark, J.E.E., Morgan, D.R., Kerr, M., Robertson, K.R., Arsenault, M., Dickinson, T.A., Campbell, C.S., 2007. Phylogeny and classification of Rosaceae. Pl. Syst. Evol. 266, 5–43.
Potter, D., Gao, F., Bortiri, P.E., Oh, S., Baggett, S., 2002. Phylogenetic relationships in Rosaceae inferred from chloroplast matK and trnL–trnF nucleotide sequence data. Pl. Syst. Evol. 231, 77–89.
Potter, D., Luby, J.J., Harrison, R.E., 2000. Phylogenetic relationships among species of *Fragaria* (Rosaceae) inferred from non-coding nuclear and chloroplast DNA sequences. Syst. Bot. 25, 337–348.
Rauscher, J.T., Doyle, J.J., Brown, A.H.D., 2002. Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. Mol. Ecol. 11, 2691–2702.
Richardson, A.O., Palmer, J.D., 2007. Horizontal gene transfer in plants. J. Exp. Bot. 58, 1–9.
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.
Rokas, A., Williams, B.L., King, N., Carroll, S.B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425, 798–804.
Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Rousseau-Gueutin, M., Gaston, A., Aïnouche, A., Aïnouche, M.L., Olbricht, K., Staudt, G., Richard, L., Denoyes-Rothan, B., in press. Tracking the evolutionary history of polyploidy in *Fragaria* L. (strawberry): new insights from phylogenetic analyses of low-copy nuclear genes. Mol. Phylog. Evol. doi:10.1016/j.ympev.2008.12.024.
Sang, T., Zhang, D., 1999. Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five *Paeonia* species based on Adh gene phylogenies. Syst. Bot. 24, 148–163.
Schranz, M.E., Osborn, T.C., 2000. Novel flowering time variation in the resynthesized polyploid *Brassica napus*. J. Heredity 91, 242–246.
Schulze-Menz, G.K., 1964. Rosaceae. In: Melchior, H. (Ed.), A. Engleŕs Syllabus der Pflanzenfamilien, II. Borntraeger, Berlin, pp. 209–242.
Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E., Small, R.L., 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92, 142–166.
Simillion, C., Vandepoele, K., Montagu, M.C.E., Zabeau, M., Van de Peer, Y., 2002. The hidden duplication past of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 99, 13627–13632.
Smedmark, J.E.E., Eriksson, T., 2002. Phylogenetic relationship of *Geum* (Rosaceae) and relatives inferred from the nrITS and trnL–trnF regions. Syst. Bot. 27, 303–317.
Smedmark, J.E.E., Eriksson, T., Evans, R.C., Campbell, C.S., 2003. Ancient allopolyploid speciation in Geinae (Rosaceae): evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. Syst. Biol. 52, 374–385.
Soják, J., 2004. *Potentilla* L. (Rosaceae) and related genera in the former USSR (identification key, checklist and figures) Notes on *Potentilla* XVI. Bot. Jahrb. Syst. 125, 253–340.
Song, K., Lu, P., Tang, K., Osborn, T.C., 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proc. Natl. Acad. Sci. USA 92, 7719–7723.
Staden, R., 1996. The Staden sequence analysis package. Mol. Biotechnol. 5, 233–241.
Staudt, G., 1962. Taxonomic studies in the genus *Fragaria*. Typification of *Fragaria* species known at the time of Linnaeus. Can. J. Bot. 40, 869–886.
Stebbins, G.L., 1950. Variation and Evolution in Plants. Columbia University Press, New York. pp. 300.
Sugiura, N., 1978. Further analysis of data by Akaikes information criterion and finite corrections. Comm. St. A. 7, 13–26.
Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Pl. Mol. Biol. 17, 1105–1109.
Tavaré, S., 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lect. Math. Life Sci. 17, 57–86.
The Arabidopsis Genome Initiative, 2000. Analysis of the genome of the flowering plant *Arabidopsis thaliana*. Nature 408, 796–814.
Urbatsch, L.E., Baldwin, B., Donoghue, M.J., 2000. Phylogeny of coneflowers and relatives (Heliantheae: Asteraceae) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chloroplast DNA restriction site data. Syst. Bot. 25, 539–565.
Vision, T.J., Brown, D.G., Tanksley, S.D., 2000. The origins of genomic duplications in *Arabidopsis*. Science 290, 2114–2117.
Wendel, J.F., Schnabel, A., Seelenan, T., 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc. Natl. Acad. Sci. USA 92, 280–284.
White, T.J., Burns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR Protocols: A Guide to Methods and Applications. Academic Press, London, pp. 315–322.
Winge, Ö., 1917. The chromosomes: their number and general importance. Compt. Rend. Trav. Lab. Carlsberg 13, 131–275.
Winkler, H., 1916. Über die experimentelle Erzeugung von Pflanzen mit abweichenden Chromosomenzahlen. Z. Bot. 8, 417–531.
Wolf, T., 1908. Monographie der Gattung Potentilla. Biblioth. Bot. 16 (Heft. 71), 1–714.
Wolfe, K.H., 2001. Yesterday's polyploidization and mystery of diploidization. Nat. Rev. Genet. 2, 333–341.
Yang, Z., 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Mol. Biol. Evol. 10, 1396–1401.
Yang, Z., Rannala, B., 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain monte carlo method. Mol. Biol. Evol. 14, 717–724.