Plant Biosystems, 2009, 1–15, iFirst Article

Taylor & Francis
Taylor & Francis Group

PROGRESS IN MORPHOLOGICAL AND MOLECULAR SYSTEMATICS

Plant species delimitation: A comparison of morphological and molecular markers

J. DUMINIL[1,2] & M. DI MICHELE[3,4]

[1]*Université Libre de Bruxelles, Service Evolution Biologique et Ecologie, CP 160/12, 50 Av. F. Roosevelt, 1050 Bruxelles, Belgium,* [2]*Faculté Universitaire des Sciences Agronomiques de Gembloux, Laboratoire d'écologie, 2 Passage des Déportés 5030 Gembloux, Belgium,* [3]*Catholic University,"John Paul II" Center for High Technology Research and Education in Biomedical Sciences, Laboratory of Analytical Techniques and Proteomics, Largo A. Gemelli, 1, 8610, Campobasso, Italy and* [4]*University of Leuven, Centre for Molecular and Vascular Biology, Gasthuisberg Campus O&N I, Herestraat 49 - b911, 3000 Leuven, Belgium*

Abstract
Species delimitation is fundamental in many areas of biology. Despite its importance, there is no agreement on criteria for species delimitation mostly due to divergence on the point of view adopted by the different biological disciplines. Two main groups of diagnostic characters are commonly used to distinguish species: the traditional morphological ones and the molecular ones. Field species recognition and sampling are generally based on morphological characters, but they can either fail to discriminate species and mask the presence of cryptic species or discriminate different species while in reality there is only one. To overcome this problem it is common to compare clusters obtained on the basis of the observed polymorphism of both characters, and to analyse their agreement. Here we compile a set of studies that have examined species delimitation with both markers. This provides a review of the different morphological and molecular markers, and of the sampling strategy and clustering methodology generally employed to delimitate species. Some conclusions are drawn with regard to species delimitation, when comparing diagnostic morphological and molecular markers.

Keywords: *Cluster analysis, phylogenetic relationships, proteomics, species concept, systematics*

Introduction

The species is the fundamental unit of biological classification but it lacks a clear definition framework. The definition of species is still under debate (Sites & Marshall 2003; de Queiroz 2005; Hey 2006). Several species concepts were proposed that differ on the basis of the criteria used to define species integrity. The biological species concept identifies a species on the basis of the interfertility of individuals, that is in terms of reproductive isolation (Mayr 1942). According to this definition, testing species delimitation calls for a test of inbreeding between individuals that originate, or not, from the same population. Apart from the technical problems associated with testing this reproductive barrier (particularly for long-lived species like trees), it is necessary to establish the minimal level of gene flow between individuals, and if sporadic events of hybridisation are sufficient to unify individuals from two recognised species as a unique one (e.g. *Quercus robur* x *Q. petraea*, Petit et al. 1997). The phylogenetic species concept defines the species as a set of individuals that share a common ancestor and are monophyletic. The major problem related to this definition consists in the fact that all informative characters for species delimitation may not be in agreement. According to the ecological species concept, individuals are grouped together in a species if they compete for the same natural resources (Colinvaux 1986). A basic condition is that members of the same species do not present any polymorphism in their life history traits related to local adaptation.

Plant species delimitation is of central importance in many areas of biology, such as biogeography, ecology, population genetics, macroevolution,

Correspondence: J. Duminil, Université Libre de Bruxelles, Service Evolution Biologique et Ecologie – CP 160/12, 50 Av. F. Roosevelt, 1050 Bruxelles, Belgium. Tel: +32 (0)2 650 4511. Fax: +32 (0)2 650 2445. Email: jduminil@ulb.ac.be

ISSN 1126-3504 print/ISSN 1724-5575 online
DOI: 10.1080/11263500902722964

phylogenetic systematics, conservation biology and biodiversity (Vogler & DeSalle 1994; Davis 1996; Sites & Crandall 1997; Schluter 2001; Agapow et al. 2004; Hedren 2004; Isaac et al. 2004; Rodriguez et al. 2007). It is also of fundamental importance to infer the patterns and mechanisms of speciation and of hybridisation, that is, respectively, the evolutionary process by which new biological species arise and the occurrence of gene flow between closely related phylogenetic species. Indeed all these biological disciplines need to delimit species as an operational taxonomic unit that represents the fundamental entity of study (Peterson & Navarro-Siguenza 1999; Goldstein et al. 2000). Similarly, in population genetics, the measures of diversity and the study of the processes that act at the population level (e.g. adaptation, gene flow, selection) need to be placed in a species context (Hedren 2004; Mank & Avise 2004). Inaccurate species delimitation may therefore have a dramatic effect on scientific studies.

Plant species recognition has long been based on morphological characters. Plant morphology is highly polymorphic and phenotypic characters may, in principle, allow plant species classification. However, different individuals of the same species may present a variation in their morphology either naturally or in connection with local adaptations. This intra-species morphological variation could be at the origin of inflated species delimitation (e.g. Pratt & Clark 2004). Alternatively some species of the same genus can be morphologically very similar and may be grouped into the same species despite the fact that they represent separate taxonomic entities (Shaw 2000; Chan et al. 2002; Whittall et al. 2004). In this latter case the taxonomic group may contain cryptic species. Another drawback of morphological characters for the differentiation of species is based on their accessibility. Indeed it is often difficult to have access to the vegetative part of adult woody individuals, especially in tropical forest ecosystems (Duminil et al. 2006). This problem increases when the diagnostic morphological characters are reproductive traits that are absent during most of the year.

In the last few decades the use of molecular markers as tools for species delimitation has drastically increased. In principle all three plant genomes (nuclear, mitochondrial and chloroplastic) can be used to delimit species. However, the mitochondrial genome is generally insufficiently polymorphic among individuals to be informative, despite the existence of universal primers (Duminil et al. 2002). In plants, chloroplastic and nuclear genomes are therefore the most commonly used. The basic premise for the use of molecular markers for species delimitation is that the "species tree" should be inferred from a "gene tree". In other words, the evolutionary history of a phylum (the phylogenetic relationships of its species) is deduced from the evolutionary history of one or more DNA fragments. However, because of the existence of evolutionary processes like reticulate evolution, lineage sorting and introgression, the genomic regions of a given species may have different evolutionary histories, leading to a disagreement between gene tree and species tree (Brower et al. 1996; Page & Charleston 1997). For example, when using both nuclear and chloroplastic markers on gene trees obtained from the two types of genome are not always coherent (Tsitrone et al. 2003). This disagreement can be related to the fact that cytoplasmic genomes can be "captured" during events of hybridisation between individuals of two different species (Petit et al. 1997). This cytoplasmic capture or introgression disconnects the phylogenetic history of the genome and of the species, and may render cpDNA non-informative on the species status. To circumvent this problem, it is necessary to use as many loci, corresponding to different genomic regions, as possible and, if possible, loci that originate from different genomes (Wang et al. 2000).

Theoretically there is a strict relationship between a genotype and a phenotype. If this is generally true, exceptions are common. A remarkable example is related to plant phenotypic plasticity: genome expression differs in function of the environmental conditions leading to morphological differences. Neutral genetic markers are in principle not affected by environmental conditions (markers not involved in the selection process) and must be informative on the species barriers whatever the heterogeneity of environmental conditions over the distribution area of the species. Therefore the efficiency of morphological and molecular markers for species delimitation may not be the same.

The aim of this study was to define in practice the impact of the used characters, morphological or molecular, on species delimitation. On the basis of published examples that have empirically tested species boundaries it is possible:

(1) to present the main morphological characters usable for species delimitation;
(2) to give a brief overview of the different usable genetic markers and identify their usefulness at different hierarchical levels, from the population to the genus level and above;
(3) to present the different sampling strategy that could be adopted for delimiting species;
(4) to report the different statistical methodologies generally used to delimitate species; and
(5) to state if the classification obtained is stable independently from the *a priori* selected discriminatory markers.

Material and methods

A literature search was done to compile a representative view of studies dealing with species delimitation in plants using both morphological and molecular markers. Thirty-five studies were considered. For each study we compiled the name of the studied species, their taxonomic family, their distribution area at the continental level, the morphological and molecular markers used for species delimitation, the methodological approach used to cluster individuals, if a demonstration of hybridisation between species has been done or not (an absence of demonstration not signifying an absence of hybridisation), the sampling strategy adopted, and the main conclusion in terms of species delimitation (Table I). For the latter category, we made a simplified classification as follows: morphological and molecular markers give the same result (they provide the same information on species delimitation, noted as "morpho=genet"), morphological and molecular markers do not give the same result (one of the two classifies individuals in more species than the other). In the latter case, we distinguished two different cases, either species delimitation is higher with morphological characters than with molecular ones (noted as "morpho>genet") or the reverse (noted as "genet>morpho"). In some studies (noted as "morpho?genet") there were no clear tests of the congruence between morphological and molecular characters, or the same number of groups were obtained whatever the marker used, but individual assignments to these groups were different (e.g. Morrell & Rieseberg 1998).

Results and discussion

Most of the compiled studies were carried out on *Magnoliophyta* species (Table I), two on *Pteridophyta* species (Speer et al. 1998; Li & Haufler 1999) and one on a *Bryophyta* species (Shaw 2000). All but one of the species from the *Magnoliophyta* are terrestrial species (the genus *Potamogeton* is aquatic, Whittall et al. 2004). As a whole, 19 taxonomic families are represented (Table I). Although some of them seem over-represented (e.g. Asteraceae), we cannot conclude here that difficulties for species delimitation are higher for these families as (1) our compilation is far from be exhaustive, (2) the problem of species delimitation has not been addressed homogeneously across all taxonomic families. For example, very few studies have been carried on southern hemisphere species (Table I). Nevertheless we can note that the Asteraceae family is one of the largest plant families in terms of number of species. This indicates a high rate of speciation and, as a consequence, a higher probability of species delimitation problems.

Usable morphological markers for species delimitation

Morphological markers used for species delimitation are generally directly dependent on the species studied (Table I). However, it is noteworthy that markers for one system of species can also be employed for a sister taxon by means of the outline method. We do not present here an exhaustive list of usable morphological characters for species delimitation, but the main morphological characters used in our compilation and that have been compared with molecular markers. Two main categories of morphological characters are generally employed: characters that originate from the vegetative and the reproductive part of a plant. The former are generally related to leaf morphology, but can also consist of different size measures of the vegetative part of the plants (e.g. Hedren 2002). It is worth mentioning that two different morphometric approaches can be employed in relation to leaf morphology: the traditional (based on distances, ratios and/or angles measures) or the geometric morphometrics approach (coordinate-based methods, outline methods, fractals) (see the paper of Viscosi et al. in this issue). More global approaches based on the whole architecture of the plant can also be used, like plant branching patterns (Warwick & Gugel 2003) and whole tree silvic characters (Aldrich et al. 2003).

The sexual characters used are generally based on morphometric analyses of some specific anatomical parts of the flower (Sharitz et al. 1980; Pratt & Clark 2004), the fruit (Elisens & Nelson 1993) or the seed (Martinez-Castillo et al. 2007). A quantification of sexual organs can also be used, as for example the number of flowers per individual (Gemeinholzer & Bachmann 2005).

Life history traits or ecological traits of the species can also be used for species classification, like the environment where they live (Li & Haufler 1999; Rajakaruna et al. 2003).

Usable genetic markers for species delimitation

The molecular markers used in the compiled studies are of nine types: isozyme/allozyme, two-dimensional electrophoresis (2-DE) protein spots, Internal Transcribed Spacers (ITS) (ETS for one study, Chan et al. 2002), low-copy nuclear gene (e.g. nuclear *cyp73* intron, Triest et al. this issue), cpDNA sequences, Simple Sequence Repeats (SSR) (i.e. microsatellites), Amplified Fragment Length Polymorphism (AFLP), Restriction Fragment Length Polymorphism (RFLP) and Random Amplified Polymorphic DNA (RAPD) (Table I). The number of chromosomes can also be used as a discriminatory character (Warwick & Gugel 2003), but it is not considered here as a molecular marker as such. The ploidy level (number of

Table I. List of the studied species.

Species	Family	Distribution area	Morphological characters	Genetic markers[a]	Clustering methodology[b]	Presence of hybrid individuals[c]	Sampling strategy	Results[d]	Literature
Amaranthus rudis, A. tuberculatus	Amaranthaceae	North America	27 characters falling within 3 categories: 8 pistillate floral characters, 17 staminate floral characters and 2 vegetative characters	Isozymes/allozymes	2	No	Herbarium specimens and 40 populations	Morpho > genet	Pratt and Clark (2004)
Anchusa crispa complex	Boraginaceae	Europe	5 reproductive traits: corolla length and corolla diameter, lobe length, anther height and style length	Isozymes/allozymes	1, 2	No	9	Morpho > genet	Quilichini et al. (2004)
Arctium lappa, A. tomentosum, A. minus	Asteraceae	Europe	9 vegetative and floral characters: stem height, basal leaf length and width, pedicle length, capitulum length and width, petiole structure, involucre pubescence and structure	RAPD	2, 5	Yes	12	Morpho = genet	Repplinger et al. (2007)
Borrichia frutescens, B. arborescens	Asteraceae	North America	Leaf and flower morphology: diameter of the stem tip and length, width and thickness of the third leaf from the meristem, height and width of inflorescence, and length of the first pair of phyllaries	PCR-RFLP (cpDNA and nrDNA)	X (log-likelihood assignment tests, and maximum likelihood estimates of genealogical class frequencies)	Yes	27	Morpho = genet	Cattel and Karl (2004)
Bromus arvensis, B. japonicus, B. squarrosus	Poaceae	Mediterranean Basin	Anther length and lemma margin angle characters	Isozymes/allozymes	1	No	70 seed accessions	Genet > morpho	Oja et al. (2003)
Calliandra calothyrsus complex	Fabaceae	Central America and Southern Mexico	Not indicated	Isozymes/allozymes	4	No	17	Morpho = genet	Chamberlain (1998)
Carapa procera, C. guianensis	Meliaceae	South America	First leave morphology after germination	SSR	6	Yes	26	Morpho = genet	Duminil et al. (2006)
Carex flava complex	Cyperaceae	Europe	12 different characters: culm height from the ground to the uppermost female spike, male spike length (excluding and including peduncle), length of the bract to the first female spike including sheath, distance between first and second female spikes, length of the bract to the second female spike, distance between second and third female spikes, length of the bract to the third female spike, number of female spikes, perygium length, perygium beak length and perygium width	Isozymes/allozymes	1, 2, 3	Yes	135	Morpho = genet	Hedren (2002)

(Continued)

Table I. *(Continued)*.

Species	Family	Distribution area	Morphological characters	Genetic markers[a]	Clustering methodology[b]	Presence of hybrid individuals[c]	Sampling strategy	Results[d]	Literature
Carex flava, C. viridula	Cyperaceae	Europe	Not indicated	Isozymes/allozymes	4 and X (genetic identity)	No	34	Morpho > genet	Bruederle and Jensen (1991)
Cichorium intybus, C. spinosum	Asteraceae	Europe, Asia, Mediterranean Basin	Number of flowers per head, number and length of outer and inner phyllaries, presence or absence of glands on the outer and/or inner phyllaries, basal thickened peduncles and hairyness on the petals	AFLP, ITS, SSR	1	No	200 Cichorium accessions	Morpho > genet	Gemeinholzer and Bachmann (2005)
Collomia wilkenii, C. tinctoria, C. renacta , C. tenella, C. linearis	Polemoniaceae	North America	Calyx morphology, corolla morphology, stamen insertion and exertion, numbers of flowers in inflorescence clusters, and the kinds and distribution of glandular and eglandular trichomes	ITS, nuclear genes, cpDNA	1, 4	Yes	Herbarium specimens and 2 populations	Morpho = genet	Johnson and Johnson (2006)
Crambe abyssinica, C. hispanica, C. glabrata	Brassicaceae	Europe, Africa	Plant branching pattern, fruit articulation and colour, leaf pubescence and leaf shape	RAPD, ITS, cpDNA, chromosome number	1	Yes	Accessions and sources of seed	Genet > morpho	Warwick and Gugel (2003)
Dubautia arborea, D. ciliolata	Asteraceae	Hawaii	Plant height and leaf morphology	AFLP	3, 6	Yes	2	Morpho > genet	Remington and Robichaux (2007)
Dubautia arborea, D. ciliolata	Asteraceae	Hawaii	Leaf morphometric traits	SSR	6	No	5	Genet > morpho	Friar et al. (2007)
Gambelia speciosa, G. juncea, G. rupicola	Scrophulariaceae	Central and North America	20 vegetative, flower and fruit characters: petiole length, lamina apex, lamina length to width ratio, inflorescence type, branch and pedicel pubescence, bract length, pedicel length in flower and in fruit, upper and lower calyx lobes length to width ratio, calyx pubescence, corolla tube length and diameter at mouth, lower corolla lip length to upper corolla lip length ratio, palate pubescence, upper filament length to lower filament length ratio, ovary pubescence, style length, capsule height to width ratio, corolla type	Isozymes/allozymes	1, 2	No	9	Morpho = genet (except for *G. juncea*)	Elisens and Nelson (1993)

(Continued)

Table I. *(Continued)*.

Species	Family	Distribution area	Morphological characters	Genetic markers[a]	Clustering methodology[b]	Presence of hybrid individuals[c]	Sampling strategy	Results[d]	Literature
Gilia achilleifolia complex	Polemoniaceae	North America	Hybrid origin hypothesized based on traditional biosystematic evidence (i.e., morphological, cytological, and crossability data), sampled taxa identified according to the taxonomic treatments of Grant	ITS	1	No	32	Morpho ≠ genet	Morrell and Rieseberg (1998)
Hordeum marinum complex	Poaceae	Mediterranean Basin, Asia	Inner glume morphology of lateral spikelets	Isozymes/allozymes	5	Yes	78 accessions	Morpho = genet	Jaaska (1994)
Lasthenia californica	Asteraceae	North America	Flavonoid races previously described for the complex correspond to edaphic races	ITS	1	No	33	Morpho ≠ genet	Rajakaruna et al. (2003)
Lasthenia californica	Asteraceae	North America	Pappus morphological characters	ITS, ETS, cpDNA	1	No	Unique individual from 61 populations	Genet > morpho	Chan et al. (2002)
Lens culinaris	Fabaceae	Europe, West Asia, North Africa	Seed area, perimeter, major and minor axis length, roundness, 100-seed weight and volume and density	SSR/proteins	1, 2	Intrasp	13	Morpho = genet	Scippa et al. (2008)
Mielichhoferia elongata, M. mielichhoferiana	Bryaceae	North America, Europe	Not indicated	ITS	1	Yes	48 accessions	Genet > morpho	Shaw (2000)
Oxytropis campestris, O. arctica	Fabaceae	Alaska Canada	Flower size and color	ITS, RAPD	1	No	20	Morpho > genet	Jorgensen et al. (2003)
Phaseolus lunatus complex	Fabaceae	Central America	Seed morphology	SSR	6, X (genotype assignment methods and frequency methods to analyze gene flow)	Yes	24	/	Martinez-Castillo et al. (2007)
Polypodium pellucidum complex	Polypodiaceae	Hawaii	Epiphytic/terrestrial transparent veins	Isozymes/allozymes	1	No	16	Morpho > genet	Li and Haufler (1999)
Potamogeton clystocarpus, P. pusillus, P. foliosus	Potamogetonaceae	North America	Absence/presence of basal tubercles on fruits	AFLP, ITS, cpDNA	1	Yes	3	Genet > morpho	Whittall et al. (2004)
Potentilla delphinensis, P. grandiflora, P. thuringiaca	Rosaceae	Europe	Number of leaflets, size of the flower, hair on radical leaf, leaf size, plant height and volume	AFLP	1, 3, 6, X (interspecific crosses)	No	23	Genet > morpho	Nicole et al. (2007)

(Continued)

Table I. *(Continued)*.

Species	Family	Distribution area	Morphological characters	Genetic markers[a]	Clustering methodology[b]	Presence of hybrid individuals[c]	Sampling strategy	Results[d]	Literature
Pteridium aquilinum complex	Dennstaedtiaceae	North America	Frond morphology	Isozymes/allozymes	1, 3	No	11	Morpho ≠ genet	Speer et al. (1998)
Puccinellia nuttalliana complex (*P. lemmonii* and *P. parishii*)	Poaceae	North America	2 diploid taxonomic species recognized on the basis of morphology	Isozymes/allozymes	4	No	20	Morpho = genet	Davis and Goldman (1993)
Quercus petraea, *Q. robur*, *Q. pubescens* and *Q. pyrenaica*	Fagaceae	Europe	Petiole length, lamina length, lobe width, sinus width, length of the lamina at largest width, number of lobes, number of veins, basal shape of the lamina, pubescence for adaxial lamina, for central nervure and petiole, lamina shape, petiole ratio, lobe depth ration, percentage venation and lobe width ratio (morphological analysis of leaf variables), 11 leaves landmark (geometric mophometric analyses of leaf shape)	SSR	6	Yes	1	Morpho = genet	Viscosi et al. (this issue)
Quercus robur, *Q. petraea*, *Q. frainetto*, *Q. pubescen,* and *Q. virgiliana*	Fagaceae	Europe	Trichome types, epicuticular waxes and stomata, stomatal density	SSR	1, 5	No	15	Morpho = genet	Fortini et al. (this issue)
Quercus section lobatae (*Q. rubra, Q. shumardii, Q. palustris*)	Fagaceae	North America	Whole-tree silvic characters: bole and crown shape, bark, branchiness and leaf shape	SSR	1, 3, 6	Yes	1	Genet > morpho	Aldrich et al. (2003)
Salix alba L., *Salix fragilis* L. *(hybrid complex)*	Salicaceae	Europe	Leaves length and width, number of leaf teeth, leaf length at the position of the maximum width, degree of leaf pubescence, length of the leaf stalk, length of the catkin, the size of the stamen and the length of the catkin stalk for the male inflorescence. Length of the catkin, length of the ovary, size of the stigma and length of the catkin stalk for the female inflorescence	Nuclear *cyp73* intron- RFLP	1, 2	Yes	16	/	Trieste et al. (this issue)

(Continued)

Table I. *(Continued)*.

Species	Family	Distribution area	Morphological characters	Genetic markers[a]	Clustering methodology[b]	Presence of hybrid individuals[c]	Sampling strategy	Results[d]	Literature
Salix alba, S. fragilis, S. x rubens	Salicaceae	Europe	Not indicated	AFLP	1, 2	Yes	1	Genet > morpho	Beismann et al. (1997)
Trollius sp.	Renonculaceae	Asia, Europe and North America	Cauline leaf with basal sheath (absent/present), nectary shape (globular/shallow pit or nearly flat), staminodia type (thick and short/thin and long), style with a membranous sheath (absent/present), flower shape (flat/bowl or globe), floral scent (present/absent), sepals after flowering (persistent/deciduous)	cpDNA, ITS, AFLP	1	Yes	1–13 individuals of each of 14 *Trollius* species or subspecies	Morpho ≠ genet	Despres et al. (2003)
Typha latifolia, T. angustifolia, T. domingensis	Thyphaceae	North America	Floral morphology (presence of pistillate bracts, shape of the stigmas and pistillodia, and pollen grouping), gross characteristics of leaf and inflorescence morphology	Isozymes/allozymes	5	No	9	Genet > morpho	Sharitz et al. (1980)

[a]RFLP: restriction fragment length polymorphism; SSR: simple sequence repeats (microsatellites); RAPD: random amplified polymorphic DNA; AFLP: amplified fragment length polymorphism; STS: sequence tagged site; EST: expressed sequence tags.

[b]1: Phylogenetic analyses; 2: Multi-factorial analyses; 3: Population subdivision (*F*-statistics); 4: Population aggregation analyses; 5: Species-specific banding patterns or allele frequencies; 6: Clustering method based on genetic markers; X: other.

[c]"Yes" means that hybrid individuals have been found, "no" means either that no hybrid individuals have been found or no test of the presence of hybridization have been done, "intrasp" means that we are not in the context of an inter-species study but in the context of an inter-cultivar study.

[d]Morpho = genet: morphological and molecular markers tell the same (they provide the same information on species delimitation); morpho > genet: morphological and molecular markers do not tell the same (species delimitation is higher with morphological characters than with molecular characters); genet > morpho: morphological and molecular markers do not tell the same (species delimitation is higher with molecular characters than with morphological characters); morpho ≠ genet: they provide different information on species delimitation; /: not available.

homologous sets of chromosomes) could also be used in species delimitation as polyploidisation (genome doubling) is an important factor in speciation (see Soltis et al. 2007 for a discussion on this topic).

The markers can be classified as codominant (allelic variations of the loci can be defined and one locus at a time is analysed) or as dominant (allelic variation of the loci cannot be defined – that is heterozygous loci could not be scored – and many loci at a time can be analysed). AFLP, PCR-RFLP and RAPD are dominant, while the other classes are codominant markers.

Isozymes were the first molecular markers used. Note that isozyme and allozyme are generally used as synonyms in the context of species delimitation. However, a fundamental difference exists: isozyme designates different molecular forms of an enzyme encoded by different loci, while allozymes are different molecular forms of an enzyme specified by different alleles at the same locus (Crawford 2000). We use hereafter the term isozyme to designate isozyme as well as allozyme. The principle is based on the revelation of enzyme variation among individuals by electrophoresis. In principle the genotype-to-phenotype relationship is relatively simple and the resolution of these markers in terms of delimitation is relatively high (Crawford 2000). The attributions of isozymes should render them more useful in the search for fixed differences among population systems.

Protein molecular markers have been progressively replaced by DNA markers (see below the perspectives offered by the proteomic approach for species delimitation). As outlined before, these markers can be dominant or codominant. One of the main interesting properties of dominant markers (AFLP, PCR-RFLP and RAPD) compared to codominant markers is their high genome coverage: many loci are sampled for individuals at the same time with only one molecular reaction. They are therefore very informative and can be used to examine the problems associated with the incongruence between the gene and the species trees. At the same time, they reveal the existence of hybridisation events. Despres et al. (2003) outlined that the limits of these markers are related to their properties of dominance and homoplasy (markers that show similar characters but are not derived from a common ancestor; thus mutation does not inform on genealogy).

SSR markers and DNA sequences are codominant markers. SSR markers are highly polymorphic markers and should in principle be very informative on the species barriers. But a major drawback of these high mutation rates is that they are homoplasious markers. However, the study of multiple SSR loci checks for this effect. As for dominant markers, they provide, at the same time, information on the potential occurrence of hybridisation between species (Duminil et al. 2006). The principal drawback of SSR markers is linked to the availability of specific primers, as their definition is still costly and time-consuming.

ITS (18S rDNA) and chloroplastic genes *rbc*L and *atp*B are the main molecular markers employed in plant phylogenetic reconstruction (APG TAPG 2003). Their use for species discrimination depends on the time of divergence among species and their rate of molecular evolution in the studied phylum. If the divergence is too recent, these markers will not be informative. Moreover these markers present problems associated with introgression for cpDNA (Petit et al. 1997) or lineage sorting for ITS sequences (Cubas et al. 2006). They nevertheless can be easily used for testing species limits given the availability of universal PCR primers for their amplification (references available in Soltis et al. 2000).

Given the different properties of these molecular markers, they may provide contrasting information on species delimitation. In principle they will provide congruent information if time of divergence among taxa is ancient. Nevertheless, the problematic of species delimitation generally concerns recent divergence among taxa. In this case and according to the specific limitations of the different molecular markers discussed above, information on species delimitation can be divergent (Morrell & Rieseberg 1998). In the present issue, Triest et al. discuss this point for species delimitation in *Salix*.

The different groups of molecular markers present specific efficiencies for the detection of polymorphisms at different taxonomical levels (Table II). We believe that, despite the limits associated with dominant markers, they constitute ideal markers to provide information on species barriers. SSR

Table II. Efficiency of molecular markers for species delimitation.

	Population level[a]	Genus level[b]	Family level[c]
Isozyme/allozyme	+++	+	–
Proteins spots	+++	+	–
ITS	–	+/–[d]	+++
SSR	+++	+++ (–[e])	–
PCR-RFLP	+	+	–
RAPD	+++	+++	–
AFLP	+++	+++	–
cpDNA	–	+	+++

[a]Polymorphism among individuals; [b]Polymorphism among species; [c]Polymorphism among genera; [d]Species delimitation power is variable from one study to another [(for Despres et al. (2003) not informative; for Warwich and Gugel (2003) informative)]; [e]Except in case of transferability of the markers among species of the same genus.

markers are also an ideal genetic tool. It is important to point out that both dominant and SSR markers can be used first to delimit species and then to study the genetic diversity and structure of the identified clades. In other words, information acquired by these markers allow species delimitation and subsequent characterisation of species genetic diversity. Indeed they can be very polymorphic at the population level (Table II). However, this requires an adequate sampling strategy (see below).

Proteomics for species delimitation

As outlined above, historically the first protein markers used for species delimitation (isozymes) were replaced by DNA markers. However, proteomics may offer new perspectives to determine the effective barriers of species.

Proteomics is a relatively new scientific discipline defined as the study of proteins expressed by a cell, tissue or whole genome of an organism. The most common methodology at the basis of proteomics is 2-DE coupled with mass spectrometry. 2-DE is a high-resolution technique able to separate thousands of genetic products (proteins) and detect post-translational modifications. Thus, resolved protein spots can be considered as useful monogenic and codominant markers. Therefore, 2-DE represents an important source of genetic markers and, as a consequence, proteomics may allow measurements of the genetic variation within and between populations. As individuals from different populations of a given species are primarily homogeneous, the majority of protein spots is expected to be similar between compared populations. Thus, population proteomics will be based on the analysis of the genetic polymorphisms detected by a qualitative (presence/absence) and a semi-quantitative analysis (relative volume) of the protein spots.

Through proteomics it is possible to investigate the differences between proteomes of ecotypes or populations from a same plant species, or distinguish different plant species; identify the major proteins which mostly contribute to the variations observed; establish whether differences originate from the expression of different protein isoforms with the same function or from the differential expression of proteins with different functions. Despite these considerable potentialities, only recently proteomics has been effectively applied to the fields of ecology and molecular evolution, thanks to the advances made in proteomics-related technologies during recent years.

There are several studies in which genetic markers obtained by 2-DE are used to study genetic variability between plant populations or species, and between varieties or ecotypes within the same species, such as in barley (Zivy et al. 1992), pine (Bahrman et al. 1994), wheat (Jacobsen et al. 2001), Brassicaceae family (Marquès et al. 2001), ginseng (Lum et al. 2002), Arabidopsis (Chevalier et al. 2004), strawberry (Alm et al. 2007) and lentil (Scippa et al. 2008). Other examples can be found in Thiellement et al. (1999), who extensively reviewed the application of the 2-DE technique to detect genetic diversity and to study phylogenetic relationships. One recent review describes how proteomics can provide useful information in phylogenetic analyses, allowing the detection and characterisation of specific proteins that have evolutionary value in terms of defining mono-, para- and polyphyly (Navas & Albar 2004). All these studies demonstrate the power and the usefulness of the proteomic approach for distinguishing lines, populations, varieties and even species.

Compared to sequence analysis, which usually provides neutral markers, the proteomic approach probably gives another kind of information, as thousands of expressed genes are taken into account, among which non-neutral polymorphisms are expected to be more frequent, because genetic variability of the amounts of proteins may have physiological consequences. Compared to the analysis of isozymes, which are the biochemical consequences of the substitution, deletion, or addition of amino acids colinearly to the nucleotide sequence, proteomics allows the examination of a broad spectrum of proteins and, consequently, a substantially increased number of protein-encoding loci.

On the other hand, one of the major limitations of proteomics is related to protein identification, because of the limited availability of sequence data. In fact, for plant species with unsequenced genomes, proteomics can be applied, but it will give a lower protein identification rate, based largely on homologous sequences. Moreover, although the proteome refers to the total protein complement of a given genome, the totality of this protein expression may not exist at any one given instant, and it is unlikely that all the translated proteins can be visualised, considering the detection threshold of 2-DE for less abundant proteins.

In conclusion, applying proteomics to evolutionary studies may provide essential tools to study the relationships between genotype and phenotype and to resolve issues specific to the ecology of plant species and evolution of populations. At the present, this promising area of research needs to be further developed. Potentially it represents a powerful tool to study phylogenetics, since 2-DE and mass spectrometry are better able to perform large-scale screening of genes than any other procedure, especially in view of the fact that the number of sequenced species is expected to increase.

Methodological approaches used to delimitate species

Six main methodological approaches were employed in the compiled studies: phylogenetic analyses, multifactorial analyses, population subdivision (*F*-statistics) analyses, population aggregation analyses, species-specific banding patterns, and the clustering method based on genetic markers (Table I).

Phylogenetic analyses are based on the principle that the gene tree corresponds to the species tree. Reticulate evolution and hybridisation are examples of evolutionary processes that implicate that this assumption is false (Brower 1999; Rosenberg 2002; Degnan & Rosenberg 2006). This explains why, despite strong advances in phylogenetics, a large amount of plant cladogenesis remains unresolved. This also represents the main limitation of barcoding methods. Nevertheless, if the gene tree corresponds to the species tree, these methods can contribute to species delimitation.

Multifactorial analysis (Principal Components Analysis) makes it possible to group individuals thanks to multiple informative characters. Both morphological and molecular markers can be used at the same time to assign individuals to different species. However, this methodological approach is generally used only for morphological markers. This method highlights the most informative morphological traits for species delimitation.

Population subdivision analyses are based on the repartition of the genetic diversity among individuals. If most of the genetic diversity between two populations is between-populations, this element supports the fact that the two populations are composed of individuals from different species. This has to be taken as an additional information used in combination with information obtained with other methods, but it could never constitute a valid and reliable test for species delimitation.

The population aggregation analysis relies on the criterion of fixation of alternate characters in different population systems (Davis & Nixon 1992). An important and interesting property here is that there are no *a priori* assignments of populations to species. Davis and Goldman (1993) discuss the details of the methodology.

The "species-specific banding patterns" approach is based on the existence of private alleles (a specific molecular tag) within a lineage or on differences in allelic frequencies between lineages. The use of private alleles to delimitate species is based on the same assumption as the barcoding method (Miller 2007). Species status is associated with one particular genotypic tag in a strict way. The genotypic tag is specific for one species. It is an allelic form of a molecular marker (revealed by sequencing or by electrophoresis, as for SSR and allozyme markers).

For a few years, clustering methods based on multilocus genotypes, that were originally developed to define a set of populations, are largely employed for studying species delimitation. Based on their multilocus genotype, individuals are clustered in different groups each characterised by a set of allele frequencies at each locus (Pritchard et al. 2000; Corander et al. 2003; Francois et al. 2006). See Sites and Marshall (2003) for a deeper discussion on the different methodologies used to delimit species.

Sampling strategy useful for delimiting species

The different methodological approaches to delimit species presented above are based on specific sampling strategies. Two main strategies are generally used: one individual per species, or many individuals per species. Different types of samples are also used: herbarium specimens, accessions (that is a record of additions to a collection), or direct sampling in the field of vegetative parts or seeds.

Phylogenetic analyses can be based either on only one individual per species or on many more. In both cases individuals are grouped on the basis of their shared similarities (cladistic) or divergence distance (phenetic). The taxonomic level of the study is generally related to the sampling strategy. If the taxonomic level is the family level or the genus level (in the case of a species-rich genus) one individual per species is sampled. However, the approach is more related to a traditional phylogenetic study than to species delimitation. In contrast, when phylogenetic approaches are used with a "multi-individual per species" sampling, their use can allow a comparison between morphological and molecular characters. The phylogenetic approach is generally based on the hypotheses developed by the phylogenetic species concept. When using a phylogenetic approach, two different options should be considered. The first one is based on the comparison of two distinct tree topologies, one obtained on the basis of morphological characters and the other on the basis of molecular markers. When both topologies are obtained, their congruence can be studied. The second one is based first on a phylogenetic reconstruction of individual phylogenetic relatedness based on data obtained by molecular markers, and, in a second step, morphological markers are mapped on the phylogenetic tree (Despres et al. 2003).

Multifactorial analyses require that the sampling covers the overall morphological trait variation used to discriminate groups. In principle, species delimitation efficiency is positively related with the number of discriminating characters. Nevertheless, in common with all but one (phylogenetic analyses) approaches, the sampling should cover as much as possible the distribution area of the studied species

with the aim of sampling the whole extent of variation of the characters.

Although the principle of the species-specific banding patterns approach is relatively simple, it needs an adequate sampling strategy to be efficient. Indeed, if the study is carried out on a restricted part of the distribution area of the studied species, discriminatory molecular tags may not be valid at a higher geographical scale. When dealing with a particular clade, the assumption of private alleles needs a widespread sampling of the studied individuals across the distribution area, in particular when using highly polymorphic markers. Indeed highly polymorphic markers present a strong restriction associated with their homoplasy.

Population aggregation analyses, as well as Bayesian clustering analyses, require "within-population" sampling that is many individuals per population and for as many populations as possible. However these methods can be quite informative using only one population with a larger sampling effort in this population (see Viscosi et al. this issue).

Comparison of morphological and molecular markers

Both correspondence and non-correspondence for species delimitation between morphological and molecular characters were demonstrated by experimental data (Table I). This conclusion on the congruence of species identification between markers is independent of the morphological or the molecular markers used, as well as of the sampling and clustering methodology. This correspondence depends directly on the specific system studied. Both cases of incongruence between morphological and molecular markers were found: higher species delimitation is found either with molecular markers in comparison with morphological markers or with morphological markers in comparison with molecular ones (Table I). The first case generally corresponds to studies on cryptic species where different species are similar morphologically, but where differences at the molecular level are sufficient to let the authors conclude that different species are present. The second case corresponds to species that present morphological variations that do not rely on species barriers but on different processes like local adaptation, phenotypic plasticity or simple neutral morphological polymorphism.

Conclusion

Genetic markers are useful tools for species delimitation. Nevertheless each genetic marker has its own properties that have to be taken into account. If the speciation event has taken place a sufficiently long time ago, ITS and cpDNA markers should be informative for species delimitation. However, in the case of recent divergence, the time of discriminatory molecular evolution should not provide any clues. The gene tree may not correspond to the species tree. This is particularly true for species that present some incidence of hybridisation. When dealing with closely related species, that have had the get time to accumulate informative characters on the species barrier, more polymorphic markers or high genome coverage markers have to be used. Therefore, most of the time, AFLP, RAPD and SSR markers should be more useful for species delimitation.

SSR markers present high mutation rates that allow species delimitation even in case of recent speciation. Moreover they can determine if hybridisation occurs among studied individuals. Coupled with assignment methods (e.g. Bayesian assignment method) they class individuals in groups that may correspond to species, and are able to detect intermediate genotypes that are probably hybrids. A test of the relative performance of these different Bayesian clustering methods can be found in two papers (Latch et al. 2006; Chen et al. 2007) for interested readers.

The main advantage of the Bayesian clustering method is that it allows species delimitation without *a priori* classification. The individuals are grouped together according to their multi-locus genotype. Therefore it is possible to sample the individuals "blindly", without any information on their distinctiveness as species, and to analyse their delimitation using molecular tools coupled with the Bayesian clustering method (Duminil et al. 2006). Indeed, if individuals are sampled with *a priori* classification based on morphological characters, and the correspondence is studied with molecular markers, there is a strong risk of obtaining an anthropogenic bias. The main difficulty is not to assign species *a priori*, but to let the data classify them. Molecular markers seem more suitable for this purpose.

Whatever the methods used to cluster individuals, how can we be confident that the groups obtained really correspond to species? As outlined in the Introduction, first it depends on the species concept adopted by the authors. Secondly we still currently fall in a vicious circle. When delimiting species on the basis of molecular markers, a validation is generally necessary from the information provided by morphological markers. If individuals are assigned to two different clusters based on their genotype, how can we be sure that the two clusters really correspond to two species? Does an agreement with discriminatory morphological markers have to be verified?

Let's imagine that we want to distinguish two species on the basis of molecular markers. We sample individuals blindly and apply cluster analyses that tell us that there are indeed two groups in our data. How

can we confirm that it really corresponds to the two species? It is quite common to check if there is an agreement with discriminating morphological markers. But if morphological markers are not in agreement with the group considered to be a species what then should we conclude? This leads to the question "which to trust?" Morphological or genetic markers? In some studies species differentiated on the basis of morphological characters could not be differentiated when using molecular tools (Gemeinholzer & Bachmann 2005). The reverse is also true. As stated above, the states of morphological markers can be closely dependent on environmental conditions. By contrast, neutral molecular markers are in principle, independent of environmental conditions, and should therefore be more reliable. The use of multiple morphological characters should, however, limit the problem associated with the environmental influence as all traits are unlikely to be affected. The problem can be, in this case, that the different molecular markers tell different stories, a problem that can be controlled by the use of different loci.

Acknowledgements

We are grateful to V. Viscosi for his critical comments on a previous version of the manuscript. JD is employed by the Belgian Funds for Scientific Research (F.R.S.-FNRS), as Postdoctoral Researcher. He was also funded by the Gembloux Agricultural University (FUSAGx, Belgium) via the project PPR 10.000. MDM is funded by a F+ fellowship by FWO (Fonds voor Wetenschappelijk Onderzoek).

References

Agapow PM, Bininda-Emonds OR, Crandall KA, Gittleman JL, Mace GM, Marshall JC, Purvis A. 2004. The impact of species concept on biodiversity studies. Q Rev Biol 79: 161–179.

Aldrich PR, Parker GR, Michler CH, Romero-Severson J. 2003. Whole-tree silvic identifications and the microsatellite genetic structure of a red oak species complex in an Indiana old-growth forest. Can J Forest Res 33: 2228–2237.

Alm R, Ekefjard A, Krogh M, Hakkinen J, Emanuelsson C. 2007. Proteomic variation is as large within as between strawberry varieties. J Proteome Res 6: 3011–3020.

APG TAPG 2003. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. Bot J Linn Soc 141: 399–436.

Bahrman N, Zivy M, Damerval C, Baradat PH 1994. Organization of the variability of abundant proteins in several geographical origins of maritime pine. Theor. Appl. Genet. 88: 407–411.

Beismann H, Barker JHA, Karp A, Speck T. 1997. AFLP analysis sheds light on distribution of two *Salix* species and their hybrid along a natural gradient. Mol Ecol 6: 989–993.

Brower AVZ, DeSalle R, Vogler A. 1996. Gene trees, species trees, and systematics: A cladistic perspective. Ann Rev Ecol Syst 27: 423–450.

Brower AVZ. 1999. Delimitation of phylogenetic species with DNA sequences: A critique of Davis and Nixon's population aggregation analysis. Syst Biol 48: 199–213.

Bruederle LP, Jensen U. 1991. Genetic differentiation of *Carex flava* and *Carex viridula* in West Europe (Cyperaceae). Syst Bot 16: 41–49.

Cattell MV, Karl SA. 2004. Genetics and morphology in a *Borrichia frutescens* and *B. arborescens* (Asteraceae) hybrid zone. Am J Bot 91: 1757–1766.

Chamberlain JR. 1998. Isozyme variation in *Calliandra calothyrsus* (Leguminosae): Its implications for species delimitation and conservation. Am J Bot 85: 37–47.

Chan R, Baldwin BG, Ornduff R. 2002. Cryptic goldfields: A molecular phylogenetic reinvestigation of *Lasthenia californica* sensu lato and close relatives (Compositae: Heliantheae sensu lato). Am J Bot 89: 1103–1112.

Chen C, Durand E, Forbes F, Francois O. 2007. Bayesian clustering algorithms ascertaining spatial population structure: A new computer program and a comparison study. Mol Ecol Notes 7: 747–756.

Chevalier F, Martin O, Rofidal V, Devauchelle A-D, Barteau S, Sommerer N, Rossignol M. 2004. Proteomic investigation of natural variation between *Arabidopsis* ecotypes. Proteomics 4: 1372–1381.

Colinvaux P. 1986. Ecology. New York: John Wiley & Sons. p. 152.

Corander J, Waldmann P, Sillanpaa MJ. 2003. Bayesian analysis of genetic differentiation between populations. Genet 163: 367–374.

Crawford DJ. 2000. Plant macromolecular systematics in the past 50 years: One view. Taxon 49: 479–501.

Cubas P, Pardo C, Tahiri H. 2006. Morphological convergence or lineage sorting? The case of *Cytisus purgans* (Leguminosae). Taxon 55: 695–704.

Davis JI. 1996. Phylogenetics, molecular variation, and species concepts. BioSci 46: 502–511.

Davis JI, Goldman DH. 1993. Isozyme variation and species delimitation among diploid populations of the *Puccinellia nuttalliana* complex (Poaceae): Character fixation and the discovery of phylogenetic species. Taxon 42: 585–599.

Davis JI, Nixon KC. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. Syst Biol 41: 421–435.

Degnan JH, Rosenberg NA. 2006. Discordance of species trees with their most likely gene trees. PLoS Genet 2: 762–768.

de Queiroz K. 2005. Different species problems and their resolution. BioEssays 27: 1263–1269.

Despres L, Gielly L, Redoutet B, Taberlet P. 2003. Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. Mol Phylogenet Evol 27: 185–196.

Duminil J, Caron H, Scotti I, Cazal S-O, Petit RJ. 2006. Blind population genetics survey of tropical rainforest trees. Mol Ecol 15: 3505–3513.

Duminil J, Pemonge MH, Petit RJ. 2002. A set of 35 consensus primer pairs amplifying genes and introns of plant mitochondrial DNA. Mol Ecol Notes 2: 428–430.

Elisens WJ, Nelson AD. 1993. Morphological and isozyme divergence in *Gambelia* (Scrophulariaceae): Species delimitation and biogeographic relationships. Syst Bot 18: 454–468.

Fortini P, Viscosi V, Maiuro L, Fineschi S, Vendramin GG. This issue. Comparative leaf surface morphology and molecular data of five oaks of subgenus *Quercus oerst*. (Fagaceae). Plant Biosyst.

Francois O, Ancelet S, Guillot G. 2006. Bayesian clustering using hidden Markov random fields in spatial population genetics. Genet 174: 805–816.

Friar EA, Cruse-Sanders JM, McGlaughlin ME. 2007. Gene flow in *Dubautia arborea* and *D. ciliolata*: The roles of ecology and isolation by distance in maintaining species boundaries despite ongoing hybridization. Mol Ecol 16: 4028–4038.

Gemeinholzer B, Bachmann K. 2005. Examining morphological and molecular diagnostic character states of *Cichorium intybus* L. (Asteraceae) and *C. spinosum* L. Plant Syst Evol 253: 105–123.

Goldstein PZ, DeSalle R, Amato G, Vogler AP. 2000. Conservation genetics at the species boundary. Conserv Biol 14: 120–131.

Hedren M. 2002. Patterns of allozyme and morphological differentiation in the *Carex flava* complex (Cyperaceae) in Fennoscandia. Nord J Bot 22: 257–301.

Hedren M. 2004. Species delimitation and the partitioning of genetic diversity – an example from the *Carex flava* complex (Cyperaceae). Biodivers Conserv 13: 293–316.

Hey J. 2006. On the failure of modern species concepts. Trends Ecol Evol 21: 447–450.

Isaac NJB, Mallet J, Mace GM. 2004. Taxonomic inflation: Its influence on macroecology and conservation. Trends Ecol Evol 19: 464–469.

Jaaska V. 1994. Isoenzyme evidence on the systematics of *Hordeum* section *Marina* (Poaceae). Plant Syst Evol 19: 213–226.

Jacobsen S, Nesic L, Petersen M, Sondergaard I. 2001. Classification of wheat varieties: Use of two-dimensional gel electrophoresis for varieties that cannot be classified by matrix assisted laser desorption/ionization-time of flight-mass spectrometry and an artificial neural network. Electrophoresis 22: 1242–1245.

Johnson LA, Johnson RL. 1994. Morphological delimitation and molecular evidence for allopolyploidy in *Collomia wilkenii* (Polemoniaceae), a new species from northern Navada. Syst Bot 31: 349–360.

Jorgensen JL, Stehlik I, Brochmann C, Conti E. 2003. Implications of ITS sequences and RAPD markers for the taxonomy and biogeography of the *Oxytropis campestris* and *O. arctica* (Fabaceae) complexes in Alaska. Am J Bot 90: 1470–1480.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conserv Genet 7: 295–302.

Li J, Haufler CH. 1999. Genetic variation, breeding systems, and patterns of diversification in Hawaiian *Polypodium* (Polypodiaceae). Syst Bot 24: 339–355.

Lum JH, Fung KL, Cheung PY, Wong MS. 2002. Proteome of oriental ginseng *Panax* ginseng C.A. Meyer and the potential to use it as an identification tool. Proteomics 2: 1123–1130.

Mank JE, Avise JC. 2004. Individual organisms as units of analysis: Bayesian-clustering alternatives in population genetics. Genet Res 84: 135–143.

Marquès K, Sarazin B, Chané-Favre L, Zivy M. 2001. Comparative proteomics to establish genetic relationships in the Brassicaceae family. Proteomics 11: 1457–1462.

Martinez-Castillo J, Zizumbo-Villarreal D, Gepts P, Colunga-GarciaMarin P. 2007. Gene flow and genetic structure in the wild-weedy-domesticated complex of *Phaseolus lunatus* L. in its Mesoamerican center of domestication and diversity. Crop Sci 47: 58–66.

Mayr E. 1942. Systematics and the origin of species. New York: Columbia University Press.

Miller SE. 2007. DNA barcoding and the renaissance of taxonomy. Proc Natl Acad Sci USA 207: 4775–4776.

Morrell PL, Rieseberg LH. 1998. Molecular tests of the proposed diploid hybrid origin of *Gilia achilleifolia* (Polemoniaceae). Am J Bot 85: 1439–1453.

Navas A, Albar JP. 2004. Application of proteomics in phylogenetic and evolutionary studies. Proteomics 4: 299–302.

Nicole F, Tellier F, Vovat A, Till-Bottraud I. 2007. Conservation unit status inferred for plants by combining interspecific crosses and AFLP. Conserv Genet 8: 1273–1285.

Oja T, Jaaska V, Vislap V. 2003. Breeding system, evolution and taxonomy of *Bromus arvensis*, *B. joponicus* and *B. squarrosus* (Poaceae). Plant Syst Evol 242: 101–117.

Page RDM, Charleston MA. 1997. From gene to organismal phylogeny: Reconciled trees and the gene tree species tree problem. Mol Phylogenet Evol 7: 231–240.

Peterson AT, Navarro-Siguenzay AG. 1999. Alternate species concepts as bases for determining priority conservation areas. Conserv Biol 13: 427–431.

Petit RJ, Pineau E, Demesure B, Bacilieri R, Ducousso A, Kremer A. 1997. Chloroplast DNA footprints of postglacial recolonization by oaks. Proc Natl Acad Sci USA 94: 9996–10001.

Pratt DB, Clark LG. 2004. *Amaranthus rudis* and *A. tuberculatus*, one species or two? J Torrey Bot Soc 128: 282–296.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genet 155: 945–959.

Quilichini A, Debussche M, Thompson JD. 2004. Geographic differentiation of morphological traits and isozymes in the Mediterranean island endemic *Anchusa crispa*: Implications for the conservation of a protected species. Biol Conserv 118: 651–660.

Rajakaruna N, Baldwin BG, Chan R, Desrochers AM, Bohm BA, Whitton J. 2003. Edaphic races and phylogenetic taxa in the *Lasthenia californica* complex (Asteraceae: Heliantheae): An hypothesis of parallel evolution. Mol Ecol 12: 1675–1679.

Remington DL, Robichaux RH. 2007. Influences of gene flow on adaptive speciation in the *Dubautia arborea-D. ciliolata* complex. Mol Ecol 16: 4014–4027.

Repplinger M, Johannesen J, Seitz A, Comes HP. 2007. Morphological and molecular evidence for hybridization and introgression in Central European Arctium (Asteraceae). Plant Syst Evol 268: 75–95.

Rodriguez J, Balch J, Rodriguez-Clark K. 2007. Assessing extinction risk in the absence of species-level data: Quantitative criteria for terrestrial ecosystems. Biodivers Conserv 16: 183–209.

Rosenberg NA. 2002. The probability of topological concordance of gene trees and species trees. Theor Popul Biol 61: 225–247.

Schluter D. 2001. Ecology and the origin of species. Trends Ecol Evol 16: 372–380.

Scippa GS, Trupiano D, Rocco MP, Viscosi V, Di Michele M, D'Andrea A, Chiatante DH. 2008. An integrated approach to the characterization of two autochthonous lentil (*Lens culinaris*) landraces of Molise (South-Central Italy). Hered 101: 136–144.

Sharitz RR, Wineriter SA, Smith MH, Liu EH. 1980. Comparison of isozymes among *Typha* species in the Eastern United States. Am J Bot 67: 1297–1303.

Shaw AJ. 2000. Molecular phylogeography and cryptic speciation in the mosses, *Mielichhoferia elongata* and *M. mielichhoferiana* (Bryaceae). Mol Ecol 9: 595–608.

Sites JW, Crandall KA. 1997. Testing species boundaries in biodiversity studies. Conserv Biol 11: 1289–1297.

Sites JW, Marshall JC. 2003. Delimiting species: A Renaissance issue in systematic biology. Trends Ecol Evol 18: 462–470.

Soltis DE. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbc*L, and *atp*B sequences. Bot J Linn Soc 133: 381–461.

Soltis DE, Soltis PS, Schemske DW, Hancock JF, Thompson JN, Husband BC, Judd WS. 2007. Autopolyploidy in angiosperms: Have we grossly underestimated the number of species? Taxon 56: 13–30.

Speer WD, Werth CR, Hilu KW. 1998. Relationships between two infraspecific taxa of *Pteridium aquilinum* (Dennstaedtiaceae). II. Isozyme evidence. Syst Bot 23: 313–325.

Thiellement H, Bahrman N, Damerval C, Plomion C, Rossignol M, Santoni V, de Vienne D, Zivy M. 1999. Proteomics for genetic and physiological studies in plants. Electrophoresis 20: 2013–2026.

Triest L, Qang Trung L, Haque Talukder M, Van Puyvelde K. 2009. Nuclear *cyp73* intron fragment length polymorphism supports morphological analysis of *Salix* species and hybrids. Plant Biosyst 143: 555–563.

Tsitrone A, Kirkpatrick M, Levin DA. 2003. A model for chloroplast capture. Evol 57: 1776–1782.

Viscosi V, Lepais O, Gerber S, Fortini P. 2009. Comparison of leaf morphological analysis in four species and hybrids of the European white oak complex. Plant Biosyst 143: 564–574.

Vogler AP, DeSalle R. 1994. Diagnosing units of conservation management. Conserv Biol 8: 354–363.

Wang X-Q, Tank DC, Sang T. 2000. Phylogeny and divergence times in Pinaceae: Evidence from three genomes. Mol Biol Evol 17: 773–781.

Warwick SI, Gugel RK. 2003. Genetic variation in the *Crambe abyssinica – C. hispanica – C. glabrata* complex. Genet Resour Crop Evol 50: 291–305.

Whittall JB, Hellquist CB, Schneider EL, Hodges SA. 2004. Cryptic species in an endangered pondweed community (*Potamogeton*, Potamogetonaceae) revealed by AFLP markers. Am J Bot 91: 2022–2029.

Zivy M, Devaux P, Blaisonneau J, Jean R, Thiellement H. 1992. Segregation distortion and linkage studies in microspore-derived double haploid lines of *Hordeum vulgare* L. Theor Appl Genet 83: 919–924.