

Differentiation between Epipactis species from western Europe.

Daniel Prat

► To cite this version:

Daniel Prat. Differentiation between Epipactis species from western Europe.. 23rd World Orchid Conference, Taiwan Orchid Growers Association; World Orchid Conference Trust, Feb 2024, Tainan City, Taiwan. pp.384-390. hal-04845333

HAL Id: hal-04845333 https://hal.science/hal-04845333v1

Submitted on 18 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Differentiation between *Epipactis* species from western Europe

Daniel Prat Univ Lyon, Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR 5023 LEHNA, F-69622, Villeurbanne, France

ABSTRACT

In the Orchidaceae family, new species are regularly described, some remaining difficult to identify, due to few distinct morphological traits. Many genera, such as the genus Epipactis, are affected by this situation. This genus includes species with overlapping distributions, particularly in Europe. The relationships between these taxa do not remain completely clarified, with some of them yet to be studied. The objectives of the present study of the genus *Epipactis* were (1) to produce a phylogeny of the genus, including taxa not yet investigated, and (2) to test how RNAseg could provide markers suitable for phylogeny. Four markers, two from the plastid genome (matK and trnLF) and two from the nuclear genome (ITS and XdH) were selected based on their polymorphisms for their applications in phylogeny. The clear differentiation of Arthrochilium section, represented in Western Europe by E. palustris, is confirmed. The basal species of the section Euepipactis is E. kleinii, which is clearly differentiated. Its sister clade included E. lusitanica and E. tremolsii from Portugal in basal position, these two species were not distinct. In the remainder of the section, few other species or groups of species appeared well supported by the bootstrap values. The geographic distribution of maternally inherited polymorphisms revealed different patterns across species. The E. lusitanica-E. tremolsii complex was under discussion. The four markers tested in the present study provided results almost similar to published phylogeny even based on RADseq but remained less powerful. To obtain more suitable markers, polymorphisms have been assessed in transcripts assembled from a RNAseq experimentation involving three different plants belonging to two species of the unresolved clade. Among the numerous transcripts assembled, those presenting long and homologous sequences, common to the three plants, were selected. Most of these sequences, ca. 10 000, showed polymorphisms suitable for applications in phylogeny and population genetics. In genera like genus *Epipactis*, with a probable recent radiation, differentiation among taxa with overlapping distribution, could only be recorded with a high number of nuclear markers, particularly when they are not yet genetically separated and still capable to hybridizing,.

Corresponding author Email: daniel.prat@univ-lyon1.fr

INTRODUCTION

The Orchidaceae family is one of the largest plant families in terms of the number of genera and species. However, many species are threatened. Their protection is complicated due to frequent changes in nomenclature and the description of new species, characterized sometimes by few morphological traits, more or less influenced by their environment, really different or not from those already known. A good knowledge of the relationships and differentiation of species would be a good support for effective protection. In many genera such as the genus Epipactis, new species, varieties, etc. have been described in recent decades. Variations within species, possible hybridizations and environmental effects sometimes make species identification difficult. Epipactis Zinn., subfamily The genus Epipendroideae, tribe *Neottiae*, is widely distributed in the Northern Hemisphere. It is made up of a variable number of species according to taxonomists, it includes a few species with obvious geographic isolation but also a large group of taxa whose distributions overlap, particularly in Europe. The genus is characterized by the repetitive autogamy¹³. evolution towards Published phylogenies have revealed little differentiation

between most species^{1, 19, 21, 22}. Even the most comprehensive analysis, based on numerous markers obtained by RNAseq¹⁹, could not determine the position of all species studied.

The objectives of the present study of the genus *Epipactis* were (1) to produce a phylogeny including taxa not yet studied and (2) to evaluate the suitability of sequences obtained by a RNAseq experiment to provide a set of markers suitable to genetic analyzes in *Epipactis*.

Several species growing in southwest Europe, such as E. kleinii have not yet been implicated in the phylogeny and their genetic relationships with other species remain unknown. The taxonomic statement of *E. fibri*, located in the Rhone valley is important to clarify for conservation issues. The presence of E. lusitanica in French eastern Pyrenees is controversial and requires clarification. Relationships between multiple species based on morphological systematics¹⁵ required genetic support. Thus, for this purpose, new taxa sampling is necessary and should include several plants from different stands to explore variation within and between taxa. This phylogeny would be established with sequences of markers from nuclear and plastid genomes. These markers might not provide a full resolved phylogeny. In addition, another set of markers more reproducible than RADseq markers and well suited to the *Epipactis* genus are nedeed for further studies. An RNAseq analysis already carried out in *Epipactis*¹⁸ could provide new markers. As it included three plants from two species, polymorphisms could be studied in these assembled transcripts sequences to select an appropriate set of markers.

MATERIALS AND METHODS

1) Plant material

Plant material was collected during the flowering season to verify plant identification. It includes wellestablished taxa of Epipactis and additional taxa described at the species or variety level. Some of remain controversial, considered them as synonyms of species described elsewhere or even local variation without taxonomic significance. This plant material has not been previously analyzed. To be sure of their identification, samples were collected in their respective locus classicus when possible. The nomenclature applied here is that of original description of the local taxon. In general, several stands of the same taxon were sampled, particularly in E. helleborine, a widely distributed species. Forty taxa were collected, and 138 plants analyzed (Table 1).

Leaves and possibly flowers or capsules were harvested, depending on the quantity of fresh intact leaves available. Immediately after collection, the plant material was frozen (at -18°C or in liquid nitrogen) until receipt in the laboratory. When transport conditions did not permit, the samples were dried with silica gel. In the laboratory, the samples were then stored in the freezer at -80°C until DNA extraction. In addition, *Cephalanthera* species were collected to serve as outgroup.

2) DNA extraction and sequence analyses

Total genomic DNA was extracted from ground samples in liquid nitrogen by the CTAB procedure, according to Doyle and Doyle (1987).

Genomic DNA was amplified by PCR using Eurobio Taq-polymerase and specific primers, following the manufacturer recommendations. Several genomic regions were tested and among them, four markers were selected based on their higher polymorphisms: two from the plastid genome (*matK* and *trnLF* spacer) and two from the nuclear genome (*ITS* and *XdH*). Specific primer pairs for these sequences were applied, the forward primer for *trnLF* was modified based on available sequences of *Epipactis* (Table 2). The PCR products were then sequenced by the Sanger method.

The sequences were aligned with ClustalX¹⁶ and then checked manually. Phylogenetic trees were produced using Seaview¹⁰ and the PhyML procedure¹¹; the distance HK85 was applied, and 1000 bootstraps were performed. Shared indels of more than one DNA base were coded as a single

base mutation, whatever their length. Phylogenetic trees were obtained separately on the four markers and on their complete concatenated sequence.

Table 1. Origin of the *Epipactis* plants analyzed. a: collection in the *locus classicus* or very close stands.

b: plants in culture.

Countries of plant collection: C (Czech Republic), F (France), H (Hungary), I (Italy), P (Portugal), S (Spain).

Таха	Stands and	
	Countries	of plants
E. albensis ª	1 (C)	2
E. atrorubens	7 (F)	7
E. atrorubens	2 (H)	2
var. borbasii	2 (11)	2
	1 (도)	1
E. atrorubens	1 (F)	I
var. <i>pallens</i>	4 (1)	4
E. autumnalis ª	1 (I)	1
E. bugacensis a	1 (H)	2
E. campeadori ª	1 (S)	1
E. cardina ^a	1 (S)	1
E. distans ^a	5 (F)	7
E. exilis	3 (F, S)	3
E. fageticola ª	4 (F, S)	4
E. fibri ª	3 (F)	4
E. gigantea	1 ^b	1
E. helleborine	14 (F, I)	15
<i>E. helleborine</i> subsp.	2 (F)	2
neerlandica		
<i>E. helleborine</i> var.	2 (F)	2
castanearum ^a		
E. helleborine var.	2 (F)	2
minor ^a	- (.)	-
E. hispanica ^a	1 (S)	1
E. kleinii	6 (F, S)	7
E. leptochila	3 (F)	3
		1
E. leptochila var.	1 (F)	I
neglecta		11
E. lusitanica ^a	5 (F, P)	11
E. maestrazgona	1 (S)	1
E. microphylla	5 (F, I)	5
E. molochina ª	1 (S)	1
E. muelleri ª	6 (F)	6
E. palustris	5 (F, I)	7
<i>E. palustris</i> var.	1 (F)	1
ochroleuca		
E. phyllanthes	4 (F)	4
<i>E. phyllanthes</i> var.	1 (F)	1
olarionensis ^a		
E. placentina ^a	3 (F, I)	5
E. provincialis ^a	3 (F)	3
E. purpurata	3 (F, I)	4
E. rhodanensis ^a	4 (F)	4
E. royleana	1 ^b	1
E. thesaurensis ^a	2 (I)	2
E. thunbergii	2 (1) 1 ^b	1
E. tremolsii	9 (F, P, S)	11
E. veratrifolia	9 (F, F, S) 1 ^b	1
	I	<u> </u>

Sequences obtained by RNAseq from flowers of three plants¹⁸, two of *E. helleborine* from the Alps region growing at contrasting altitudes and a plant of E. placentina from France, were analyzed to evaluate the available polymorphisms between and within the two species, belonging to the section Euepipactis. Sequences longer 500 bases were selected for high homology along their complete common sequence using rBlast¹². The functions of the genes expressed in flowers were inferred from the Uniprot protein sequence database using the Diamond tools of the Galaxy platform³. Sequences alignments were analyzed to assess the extent of variation between the three plants and within species. The same approach was applied to the full set of selected sequences and to some assembled transcripts used as examples. The level of SNPs could thus be compared to that obtained in the four sequences used for present phylogeny study.

Table 2. Primers and Tm conditions for PCR.

Gene	Primers 5'-3'	Tm (°C)
ITS	CCTTATCAYTTAGAGGAAGGAG ⁴ RGTTTCTTYTCCTCCGCTTA ⁴	54.0
matK	CGTTCTGRCCATATTGCACTATG ⁸ AACTAGTCGGATGGAGTAG ¹⁴	55.0
trn-F	GGGGATATGGCGAAATTGGTAG ATTTGAACTGGTGACACGAG ²⁰	59.0
Xdh	GAAGAGCAGATTGAAGAWWGCC ^S GWGAGAGAAAYTGGAGCAAC ⁹	9 65.0

RESULTS

1) Sequence polymorphisms

The four markers had a similar number of SNPs (Table 3) but the two plastid genome markers also showed variations in indels, particularly for the *trnLF* marker, and simple sequence repeats (SSRs) consisting of a single base repeats (A or T). The maximum variation in length of the SSRs was from 8 to 22 bases; little, if any, variation has been recoded within species. On average, a SNP is recorded every 30 DNA bases for nuclear genome markers and every 50 bases for plastid genome markers.

Assembled transcripts sequences, more than 500 bases long, with homology over the full length of the common fragment, revealed polymorphisms in more than 90% of them. Sequence identity was slightly higher between two plants of *E. helleborine* than among the three plants tested (Table 4). On average of more than 10.000 sequences selected without redundancy, one SNP was discovered each 125 DNA bases. The few sequence examples provided show a large variation of SNP frequency.

2) Species differentiation

The two sections of the genus *Epipactis* are well differentiated (Fig. 1). Some species or groups of

species were well separated while a large proportion of samples remained in an unresolved clade. The basal position of section *Euepipactis* was occupied by *E. kleinii* which has the largest genome size in the genus. *E. lusitanica* (P) and *E. tremolsii* from Portugal and a sample of *E. tremolsii* from France grouped together without separation of the two species. *E. microphylla*, as well as *E. exilis*, consisted of a well separated species.

Table 3. Polymorphisms observed in the four			
markers investigated in <i>Epipactis</i> for phylogeny			
reconstruction.			

Locus	Aligned length	SNPs	Indels	SSR (poly A/T)
ITS	725	28	1	-
matK	1226	27	3	2
trnLF	1401	26	11	2
XdH	1001	28	-	-

Table 4. Polymorphisms observed in homologous assembled transcripts of two plants of *E. helleborine* and one of *E. placentina*.

SNPs: number of SNPs observed per 1000 DNA bases.

Ehell: percent of sequence identity between two *E*. *helleborine* plants.

Spec: average percent of sequence identity between the three analyzed plants.

Predicted gene	SNPs	Ehell	Spec	
Anthocyanidine	12.8	99.8	98.7	
glucosyl-transferase				
Anthocyanin synthase	18.5	98.7	98.2	
Chalcone isomerase	3.5	99.9	99.6	
Chalcone synthase	6.3	99.6	99.4	
Dihydroflavonol	7.5	99.8	99.3	
4-reductase				
Geranyl diphosphate	7.5	100.0	99.3	
phosphohydrolase				
geranylgeranyl	0.6	99.9	99.9	
diphosphate reductase				
Geranylgeranyl	12.1	99.9	98.8	
pyrophosphate synthase				
Isoflavone 2'-hydroxylase		100.0	99.8	
Ultrapetala 1	2.6	99.7	99.7	
Full set of selected	2.0 7.9	99.4	99.2	
	1.9	55.4	99.Z	
transcripts (>10 000)				

In the large unresolved clade, some groups could be however identified: *E. bugacensis*, *E. rhodanensis* and *E. hispanica* grouped together. Few differences could be recorded in the sequences of these species. *E. muelleri* and the sample of *E. placentina* from France were separated from other samples. *E. fageticola* and *E. phyllanthes* constituted a well- differentiated clade. All samples of *E. purpurata* grouped together. The significant cluster of *E. distans* on one hand and of *E. helleborine*, on the other hand, only included part of the samples. The two samples of the variety *E. helleborine* var. *minor* clustered together in a homogenous group of *E. helleborine* samples without any statistical support.

The samples of *E. fibri* and *E. albensis* showed identical sequences for all four markers but were not really differentiated from the remaining species. The two samples of the variety *E. atrorubens* var. *borbasi* clustered with the other *E. atrorubens* samples, without any bootstrap support.

Samples presumed to be *E. lusitanica* from France, marked (F) in Fig. 1, grouped with a sample of *E. tremolsii* from Spain, one from France and *E. cardina* from Spain. These plants shared a specific insertion in the plastid genome. Although this group was not supported by bootstrap values, it is however clearly separated from the group including *E. lusitanica* from Portugal (where the species was described). They shared only part of nuclear polymorphisms with plants of Portugal.

DISCUSSION

1) Phylogeny

The resulting cladogram showed clear similarities whit that of Sramkó et al. (2019) which shared multiple species and had stronger statistical support. In the absence of E. kleinii, E. lusitanica occupied the basal position of section *Euepipactis*. E. atrorubens and E. phyllanthes were closely related but proximity to E. exilis was not observed in the present study. E. phyllanthes and E. fageticola could be considered two subspecies of the same species. E. muelleri and E. purpurata were well differentiated in both studies, but their position, as for the *E. bugacensis* group, could not be compared. E. leptochila and its variety (considered as species in Sramkó et al., 2019) cluster together. Autogamous species did not constitute a clade.

The main species, *E. helleborine*, with a wide distribution area, shows a large genetic variation as shown by Evans *et al.* (2023) and did not consist of a monophyletic clade in the study of Sramkó *et al.* (2019). In the present study, samples distributed across the unresolved clade appeared to show reduced extent. These samples came from Southeast France and northern Italy; this limited geographical distribution could explain their reduced diversity. Sampling remains a crucial point in phylogeny, especially when species are poorly separated. Variation within species must be captured to avoid overestimation of species differentiation.

2) Taxonomic considerations

The five sections of the genus proposed by Klein (2005) did not receive support from phylogenies, with the exception of sections *Arthrochilium* and possibly *Cymbochilum* (consisting of *E. veratrifolia*). For example, *E. atrorubens*, *E. kleinii*, and *E.*

microphylla which were supposed to belong to the same section, did not group together.

E. albensis and *E. fibri*, with very similar morphology and ecology, showed no differences in the markers analyzed. As several organizations have already done, they can not be considered as two different species despite the distance between the populations of the Rhône valley and the Elbe valley, the name *E. albensis* has the priority.

E. bugacensis and *E. rhodanensis* have similar habits and morphology, few differences in the four markers were observed. Among the four samples from France, the one from Pyrenees was genetically closest to the Hungarian population. They are not two different species.

The situation is more controversial for the presence of *E. lusitanica* in France. The plants growing there are clearly genetically different from those growing in Portugal. They shared an insertion in plastid genome with plants known as *E. tremolsii* in Spain and France. Even in their nuclear genome, some differences with plants from Portugal have been identified. This situation could result from past hybridizations between different lineages.

3) Development of markers

A large amount of polymorphisms is necessary to obtain a resolved phylogeny in *Epipactis*. Next generation sequencing significantly increases the number of markers available for systematic and evolutionary studies, including in *Epipactis*². Various recent developments⁶ provide many more markers that the usual few ones used in phylogeny. The assembled transcripts evaluated in the present study from RNAseq could also be a good approach to obtain many markers suitable for the phylogeny of *Epipactis* and perhaps other orchid species.

CONCLUSION

The present study confirms the low differentiation between western European *Epipactis* species that exhibit overlapping zones. These data could be used to establish evolutionary trends and possible migration routes for *Epipactis*. Further developments with more powerful markers are still needed to answer all questions about the evolution of *Epipactis* and of orchids in general. The results also have taxonomic consequences.

ACKNOWLEDGMENTS

I would like to thank Javier Benito Ayuso, Joris Bertrand, Pierre-Michel Blais, Cyril Blanc, Martine Breret, Pierre Brothier, Roselyne Buscail, Mickael Busi, Léa Charvin, Christian Dirwimer, Irena Formanová, Olivier Gerbaud, Alain Gévaudan, Pascal Lambé, Michel Nicole, Christian Noël, Helmut Pressler, Gil Scappaticci, Daniel Tyteca, Yves Wilcox for the samples they provided or their help in the collecting plants in native populations, and Lara Konecny for her welcome and support at the molecular laboratory of LEHNA.

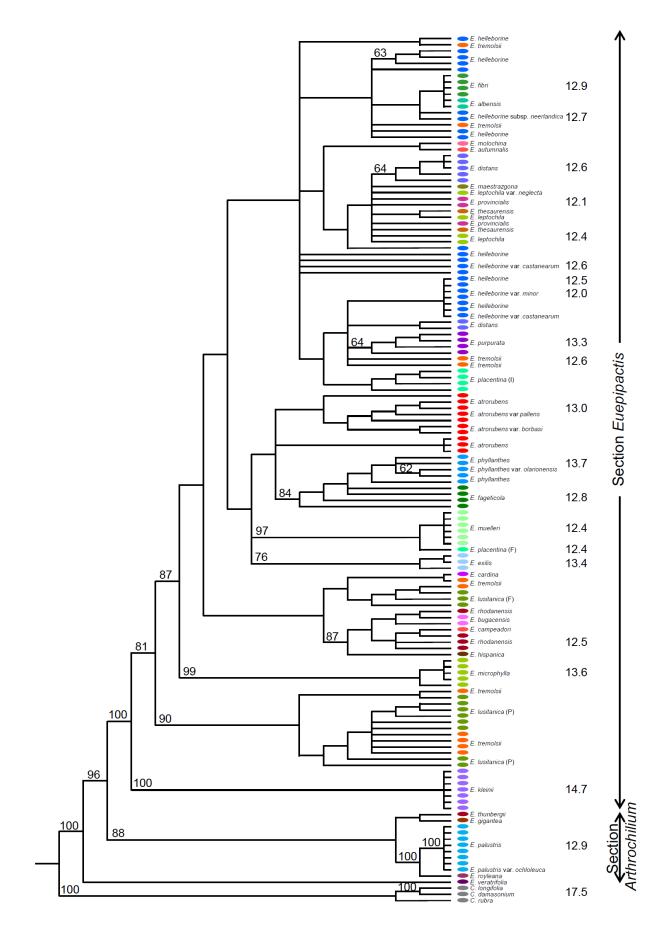


Fig. 1. Cladogram of the *Epipactis* taxa analyzed in the present study.

Values of bootstraps higher than 50% are provided; branches with bootstrap values lower than 10% were merged. 1C DNA amount in Gbases is provided on the right of the figure (from Prat *et al.* 2014).

REFERENCES

- Bateman, R., P.M. Hollingsworth, J. Squirrell and M.L Hollingsworth (2005) Tribe Neottieae. In: A.M. Pridgeon, P.J. Cribb, M.W. Chase and F.N. Rasmussen (eds). Genera Orchidacearum, 4. p. 487-495. Oxford University Press.
- 2. Bateman, R.M. (2020) Implications of nextgeneration sequencing for the systematics and evolution of the terrestrial orchid genus *Epipactis*, with particular reference to British Isles. Kew Bull. 75:4.
- 3. Buchfink, B., C. Xie and D.H. Huson (2014) Fast and sensitive protein alignment using DIAMOND. Nature Methods 12:59-60.
- Cheng, T., C. Xu, L. Lei, C.H. Li, Y. Zhang and S.L. Zhou (2016) Barcoding the kingdom Plantae: new primers for *ITS* regions of plants with improved universality and specificity. Mol. Ecol. Res. 16:138-149.
- 5. Doyle, J.J. and J.L. Doyle (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19:11-15.
- Eserman, L.A., S.K. Thomas, E.E.D. Coffey and J.H. Leeben-Mack (2021) Target sequence capture in orchids: developing a kit to sequence hundreds of single-copy loci. App. Plant Sci. 9:e11416.
- Evans, A., H. de Kort, R. Brys, K.J. Duffy, J. Jersáková, T. Kull, M.A. Selosse, S. Tsiftsis, J. Minasiewicz and H. Jacquemyn (2023) Historical biogeography and local adaptation explain population genetic structure in a widespread terrestrial orchid. Ann Bot. 28:623-634.
- Goldman, D.H., J.V. Freudenstein, P.J. Kores, M. Molvray, D.C. Jarrell, W.M. Whitten, K.M. Cameron, R.J. Jansen and M.W. Chase (2001) Phylogenetics of Arethuseae (Orchidaceae) based on plastid *matK* and *rbcL* sequences. Syst. Bot. 26:670-695.
- Górniak, M., O. Paun and M.W. Chase (2010) Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene, *Xdh*: congruence with organellar and nuclear ribosomal DNA results. Mol. Phyl. Evol. 56:784-795.
- 10. Gouy, M., S. Guindon and O. Gascuel (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27:221-224.
- 11. Guindon, S., J.F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk and O. Gascuel (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59:307-321.

- 12. Hahsler, M and A. Nagar (2019) rBLAST: R interface for the basic local alignment search tool. R package version 0.99.2, URL: https://github.com/mhahsler/rBLAST.
- Hollingsworth, P.M., J. Squirrell, M.L. Hollingsworth, A.J. Richards and R.M. Bateman (2006) Taxonomic complexity, conservation and recurrent origins of self-pollination in *Epipactis* (Orchidaceae). In J. Bailey and R.G. Ellis (eds). Current taxonomic research on the British & European flora, 27. p. 27-44. BSBI.
- Johnson, L.A., E. Douglas and D.E. Soltis (1994) *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. Syst. Bot. 19:143-156.
- Klein, E. (2005) Versuch enier Gliederung der Gattung *Epipactis* Zinn. (Orchidaceae-Neottiae). J. Eur. Orchid. 37:121-130.
- Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson and D.G. Higgins (2007) Clustal W and Clustal X version 2.0. Bioinformatics, 23:2947-2948.
- 17. Prat, D., S. Brown and A. Gévaudan (2014) Evolution des Neottieae, apport de la cytométrie en flux. Cah. Soc. Fr. Orch. 8:125-133.
- Prat, D., J. Briolay and V. Navratil (2018) RNAseq as a source of genetic polymorphisms in *Epipactis* for molecular marker development. Cah. Soc. Fr. Orch. 9:91-93.
- Sramkó, G., O. Paun, M.K. Brandrud, L. Laczkó, A. Molnár and R.M. Bateman (2019) Iterative allogamy–autogamy transitions drive actual and incipient speciation during the ongoing evolutionary radiation within the orchid genus Epipactis (Orchidaceae). Ann. Bot. 124:481-497.
- Taberlet, P., L. Gielly, G. Pautou and J. Bouvet (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17:1105-1109.
- 21. Xiang, X.G., D.Z. Li, W.T. Jin, H.L. Zhou, J.W Li and X.H. Jin (2012) Phylogenetic placement of the enigmatic orchid genera *Thaia* and *Tangtsinia*: evidence from molecular and morphological characters. Taxon 61:45-54.
- 22. Zhou, T. and X.H. Jin (2018) Molecular systematics and the evolution of mycoheterotrophy of tribe Neottieae (Orchidaceae, Epidendroideae). PhytoKeys 94:39-49.