Delpinoa, n.s. **33-34**: 77-85. 1991-1992

Testing the usefulness of ITS1 sequence as a tool to infer relationships in *Orchis* L.

SALVATORE COZZOLINO*, PAOLO CAPUTO**, SERENA ACETO***, WALTER ROSSI****, PAOLO DE LUCA*.

*Orto Botanico e **Dipartimento di Biologia vegetale, Via Foria, 223, 80139 Napoli, ***Dipartimento di Genetica, Biologia generale e molecolare; Facoltà di Scienze, Università degli studi di Napoli Federico II. ****Dipartimento di Scienze ambientali, Università degli Studi de L'Aquila.

Riassunto

Lo spaziatore intragenico I (ITS1) del DNA nucleare di *Orchis coriophora*, *O. laxiflora*, *O. morio*, *O. purpurea* e *O. simia* è stato sequenziato per verificarne le potenzialità d'uso nello studio delle relazioni evolutive all'interno del genere. L'ITS1 di queste specie, risultato lungo 235-244 paia di basi e con un contenuto in GC del 41-47%, si è dimostrato sufficientemente variabile da permettere la risoluzione completa dei rapporti di affinità tra le specie in studio. Il fenogramma ottenuto mediante cluster analysis mostra due gruppi di specie; uno formato da *O. purpurea* e *O. simia*, vicinissime tra loro, e l'altro da *O. coriophora*, *O. morio* e *O. laxiflora*, con le prime due più vicine rispetto alla terza. Tali risultati sono in gran parte congruenti con dati morfologici, cariologici e molecolari disponibili in letteratura, e pertanto permettono di indicare l'ITS1 come un segmento di DNA fornito di variabilità appropriata per risolvere problemi filogenetici all'interno del genere *Orchis*.

INTRODUCTION

The genus *Orchis* L. (Orchidaceae) is generally regarded as not easily prone to phylogenetic inference. The great amount of parallel evolution which characterizes the genus and its allies (DRESSLER, 1981; 1993), the ecological convergence (DRESSLER, 1993), as well as the rather widespread tendency to form hybrids (e.g., EHRENDORFER, 1980) greatly complicate matters when attempts are made to infer historical propinquity in the genus by using morphological traits. As a consequence, and in spite of all the attention European orchids have attracted in the past two centuries, no general phylogenetic treatment of the species of *Orchis* is yet available, with the exclusion of a partial attempt by CAUWET-MARC and BALAYER (1984).

Key words: ITS1, Orchidaceae, *Orchis*, Ribosomal DNA.

Recently, novel approaches have started being used in inferring relationships among members of the genus *Orchis* (SCHLEGEL *et al.*, 1989; ROSSI *et al.*, 1994). Such methods mainly rely on starch gel enzyme electrophoresis and on computer algorithms to discriminate among different electromorphs. These techniques are doubtlessly of invaluable help in inferring affinities among these plants, but suffer from several methodological complications; among others, the necessity of testing all previous samples again whenever a new species is added to the analysis. Moreover, enzyme loci are employed in systematics under the assumption that their modifications reflect the underlying modifications occurring in the DNA. This is true in general terms, but changes in aminoacid sequence composition which do not alter catalytic function or mobility in starch gels, as well as silent mutations in the related DNA escape undetected to a study relying on enzymes. For these reasons, nucleic acids are considered a tool whose resolutory power is greater than that of proteins in drawing phylogenetic hypotheses. In recent years, DNA studies have proven to be extremely accurate in inferring phylogeny in many plant organisms at any level of taxonomic hierarchy (see for example OLMSTEAD and PALMER, 1994, for chloroplast DNA and BALDWIN *et al.*, 1995, for nuclear sequences). Among the variety of sequences employed in plant phylogeny, the internal transcribed spacers, ITS 1 and 2, of the nuclear ribosomal DNA (n-rDNA), have proven to be extremely useful for infrageneric comparisons (SUH *et al.*, 1993; KIM and JANSEN, 1994 and references therein), because of their fast mutation rate.

This paper aims at testing the viability of ITS1 sequencing to infer relationships in the genus *Orchis*. For this purpose, species for which morphological (CAUWET-MARC and BALAYER, 1984), karyological (CAUWET-MARC and BALAYER, 1984), enzymatic (ROSSI *et al.*, 1994) and chloroplast DNA (CAPUTO *et al.*, 1995) data are available were chosen, taking care to select taxa which, founding on independent evidence, are either very close or very distant to each other, so as to define the upper and lower discrimination bounds for ITS1.

MATERIAL AND METHODS

Individual plants of *O. coriophora* L. subsp. *fragrans* (Pollini) Sudre (OCOR), *O. laxiflora* Lam. (OLAX), *O. morio* L. (OMOR), *O.*

purpurea Hudson (OPUR), and *O. simia* Lam. (OSIM) were collected at flowering time on the mountains W of Vallo di Diano (province of Salerno, Italy). Voucher specimens of all the examined plants are deposited at NAP.

Leaves (1 g per sample) were ground on liquid nitrogen and total DNA was extracted following the procedure described in CAPUTO *et al.* (1991).

ITS1 was amplified by using two primers which anneal respectively in the 3' region of the 18S DNA (5'-GGAGAAGTCGTAACAAGGTTTCCG-3') and in the 5' region of the 5.8S DNA (5'-ATCCTGCAATTCACACCAAGTATCG-3'). PCR reactions were carried out for 30 cycles. Initial conditions were as follows: 1 min denaturation at 94 °C, 1 min annealing at 55 °C, 45 sec extension at 72 °C. Samples were denatured for 5 min at 94 °C before the beginning of the first cycle; extension time was increased of 3 sec/cycle; extension was further prolonged for 7 min at the end of the last cycle.

PCR fragments were then directly sequenced in both directions by using a modification of the Sanger dideoxy method implemented in a Gibco-BRL dsDNA Cycle Sequencing System reaction kit. Preparation of the gel (5% acrylamide), electrophoresis (1800 V, 50 °C), gel drying and autoradiography were carried out as indicated in HILLIS *et al.* (1990).

Sequence alignments were accomplished by using the CLUSTAL V program (HIGGINS and SHARP, 1989) and distance computations by using the Jukes-Cantor method (JUKES and CANTOR, 1969) with the DNADIST software of the PHYLIP 3.57 package (FELSENSTEIN, 1993). The calculation of the dendrogram was carried out by employing an algorithm of cluster analysis (UPGMA) as implemented in the NEIGHBOR software in the above said package.

RESULTS

The ITS1's of the plants in study are 235-244 base pairs (bp) long, with a GC content ranging from 41 to 47%. The exact lengths and GC contents for each species are: OCOR 237 bp (42%), OLAX 235 bp (41%), OMOR 240 bp (47%), OPUR 244 bp (42%), and OSIM 244 bp (43%). The complete sequences are available upon request to the senior author.

The GC/ATGC plots for the sequences in study are shown in Fig. 1. OCOR and OLAX have the maximum GC concentration in the 125-150 bp range; OMOR has the maximum in the 110-150 range; OPUR and OSIM have their maxima in the 60-80 range.

The potentially informative/total aligned sites ratio (i.e., the different sites/total sites ratio), calculated on a consensus length of 251 bp is 0.36. When deletions and insertions (indels) are excluded from this computation, the above said ratio drops to 0.27.

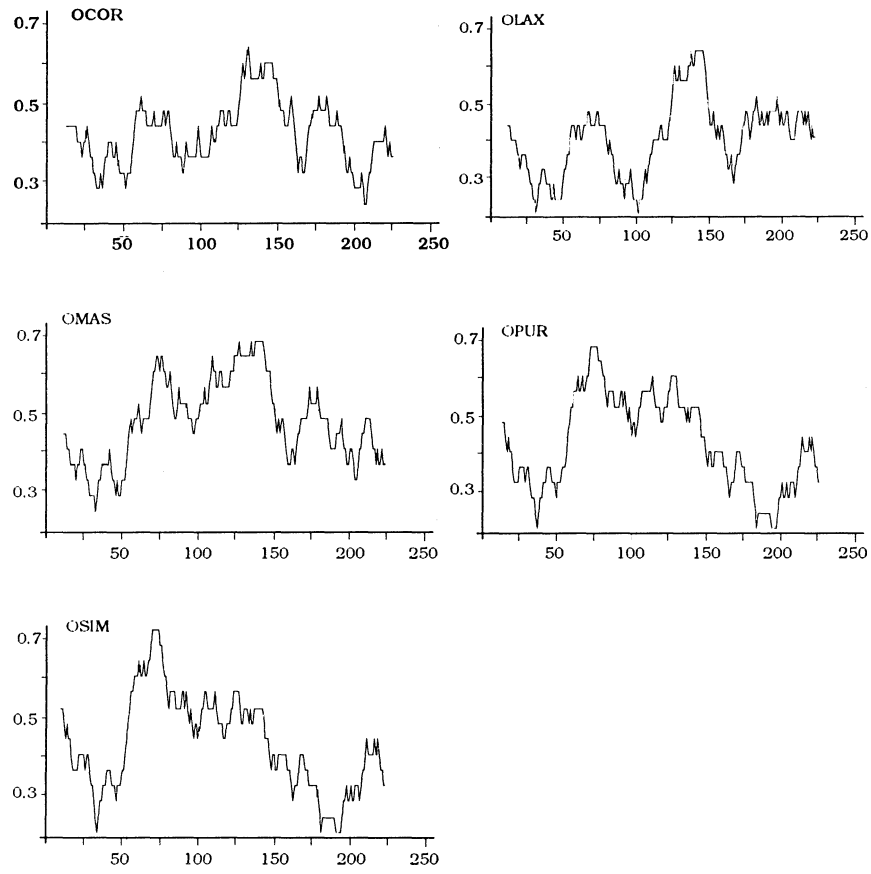

Fig. 1 - GC/ATGC plots for the species in study. See the text for acronyms.

The Jukes-Cantor distance for each pair of species is reported in Tab. I. The shortest distance separates OPUR and OSIM, and the largest OCOR and OSIM.

The UPGMA analysis on the distance matrix yielded the phenogram of Fig. 2, from which it is possible to desume that OPUR and OSIM form a cluster, well separated from the rest of the species. OCOR and OMOR also form a cluster, which is in turn contained in a larger cluster to which OLAX is added.

Tab. I - Jukes-Cantor distances calculated on the ITS1 of the species in study.

OCOR	0.0000	0.1932	0.1348	0.2140	0.2256
OLAX	0.1932	0.0000	0.1903	0.1950	0.2005
OMOR	0.1348	0.1903	0.0000	0.1714	0.1876
OPUR	0.2140	0.1950	0.1714	0.0000	0.0124
OSIM	0.2256	0.2005	0.1876	0.0124	0.0000

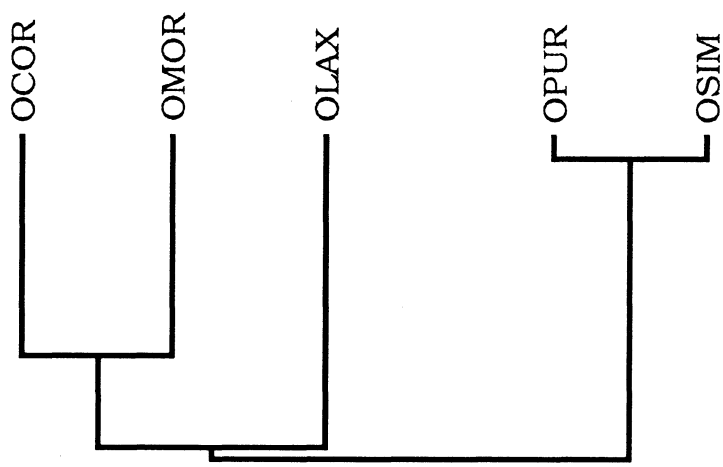

Fig. 2 - Dendrogram for the five species in study, based on an UPGMA analysis of the Jukes-Cantor distances. See the text for acronyms.

DISCUSSION AND CONCLUSIONS

The ITS1's of the species in study, which are the first sequences of this kind obtained from *Orchis*, range from 235 to 244 bp, and therefore are well included in the range of length variation reported for angiosperms (187-298 bp, BALDWIN *et al.*, 1995). On the contrary, their GC content (41 to 47%) is comparatively low; only within Polemoniaceae, Scrophulariaceae and Viscaceae such low or lower GC contents are reported (BALDWIN *et al.*, 1995).

The informative sites/total sites ratio in the studied orchids (both with and without indels) is the highest yet recorded for angiosperms (BALDWIN *et al.*, 1995). Comparable values are found only in the papilionoid tribe Galegaeae (SANDERSON and WOJECHOWSKI, 1993), and higher are found among Polemoniaceae (PORTER, 1993). Regardless, these values were obtained by comparing sequences at the tribal or even familial levels. The amount of indels (23) is also high in our species. Even if no group-wide comparison is available, it is notable that the number of indels in the majority of infrageneric studies (e.g., KIM and JANSEN, 1994) is less than half the number reported in our five *Orchis* species.

The variability of ITS1 sequence in our species of *Orchis* influences also the distribution of GC rich zones (Fig. 1). Indeed, even a qualitative estimate of Fig. 1 would suggest a pattern of affinities similar to that indicated by distance analysis on the whole sequences.

The closest pair of species in study, OPUR and OSIM, showed differences in three positions; the second closest pair, OCOR and OMOR, showed differences in 34 positions. However, in spite of the said high interspecific variability in the ITS1 of the plants in study, preliminary experiments carried out on *O. simia* (data not shown) indicated that samples originating from two different Italian localities (approx. 300 Km apart from each other) showed no ITS1 sequence difference.

The phenetic relationships depicted by the UPGMA dendrogram of Fig. 2 are congruent with other evidence. In particular, the separation in two groups of *Orchis* species is supported by morphological (CAUWET-MARC and BALAYER, 1984), karyological (CAUWET-MARC and BALAYER, 1984), enzymatic (ROSSI *et al.*, 1994) and chloroplast DNA data (CAPUTO *et al.*, 1995). Notably enough, coincidence between our results and

previous evidence occurs irrespectively of the method used in data analysis: in fact, the pattern of relationships obtained with our sequence data correspond to that obtained for chloroplast DNA restriction fragment polymorphisms (CAPUTO *et al.*, 1995) by using cladistics (which was not used here for the lack of an appropriate outgroup).

However, the closer affinity between OCOR and OMOR than between them and OLAX (Fig. 1), which is asseverated on karyological, enzymatic and chloroplast DNA grounds (see references above), is contradicted by the dendrogram, based both on morphological and karyological traits, present in CAUWET-MARC and BALAYER (1984). Regardless, for the reasons exposed in the introduction, we believe that morphology is not necessarily an appropriate tool to infer relationships in a group notoriously beset with parallelisms.

In conclusion, ITS1 proved to be a suitable molecule to infer historical propinquity in *Orchis*: its variability range within this genus is appropriate to draw hypotheses both on distant and closely related species. On the contrary, a more variable molecule (or as an alternative the inclusion of ITS2 sequences in the analysis) may prove more useful for the study of problems pertaining to lower taxonomic hierarchies.

The initial results reported here have stimulated further research on more species of *Orchis*, utilizing also ITS2. DNA sequencing is being employed to extend our studies to the whole genus and to its relationships with various satellite group, in the hope of elucidating its phylogeny.

Abstract

The intragenic spacer I of the nuclear DNA (ITS1) was sequenced in *O. coriophora*, *O. laxiflora*, *O. morio*, *O. purpurea*, and *O. simia*, in order to test its suitability as a phylogenetic tool within the genus. ITS1 (length 235-244 base pairs, GC content 41-47%) appeared variable enough as to grant full resolution of the relationships among the species in study. The resulted UPGMA phenogram indicates that *O. coriophora* and *O. morio* form a cluster, included in a larger cluster in which *O. laxiflora* is also present. *O. purpurea* and *O. simia*, very close to each other, form another cluster. These results are congruent to a great extent with morphological, karyological and molecular data, and therefore allow to indicate ITS1 as an appropriate sequence to solve phylogenetic problems in *Orchis*.

REFERENCES

BALDWIN B.G., SANDERSON M.J., PORTER M.J., WOJCIECHOWSKI M.F. CAMPBELL C.S. and DONOGHUE M.J., 1995. *The ITS region of nuclear ribosomal DNA: a valuable source of evidence in angiosperm phylogeny.* Ann. Missouri Bot. Gard., 82: 247-277.

CAPUTO P., ACETO S., COZZOLINO S., NAZZARO R. and GAUDIO L., 1995. *Relazioni tra* Orchis *L. (Orchidaceae) ed alcuni generi affini.* Giorn. Bot. Ital., 129(1): 116.

CAPUTO P., STEVENSON D.W. and WURTZEL E.T., 1991. *A phylogenetic analysis of American cycads (Cycadales) using chloroplast DNA restriction fragment polymorphisms.* Brittonia, 43: 135-145.

CAUWET-MARC A.M. and BALAYER M., 1984. *Les genres* Orchis *L.,* Dactylorhiza *Necker ex Newski,* Neotinea *Reichb. et* Traunsterinera *Reichb.: Caryologie et proposition de phylogénie et d'évolution.* Botanica Helvetica, 94: 391-406.

DRESSLER R.L., 1981. *The orchids: natural history and classification.* Harvard University Press, Cambridge (Mass.)

DRESSLER R.L., 1993. *Phylogeny and classification of the orchid family.* Cambridge University Press, Cambridge (England).

EHRENDORFER F., 1980. *Hybridisierung, Polyploidie und Evolution bei europaischmediterranen Orchideen.* Jahresber. Naturwiss. Ver. Wuppertal, 33: 15-34.

FELSENSTEIN J., 1993. *PHYLIP (Phylogeny Inference Package) version 3.5c.* Distributed by the author. Department of Genetics, University of Washington, Seattle.

HIGGINS D.G. and SHARP P.M., 1989. *Fast and sensitive multiple sequence alignments on a microcomputer.* CABIOS, 5: 151-153.

HILLIS D.M., LARSON A., DAVIS S.K. and ZIMMER E.A., 1990. *Nucleic acids III: sequencing.* In: HILLIS D.M. and MORITZ C. (Eds.) Molecular systematics. Sinauer Associates, Sunderland (Ma).

JUKES T.H. and CANTOR C.R., 1969. *Evolution of protein molecules.* pp. 21-132 In: MUNRO H.N. (Ed.) Mammalian Protein Metabolism. Academic Press, New York.

KIM Y.-D. and JANSEN R.K, 1994. *Comparison of phylogenetic hypotheses among different data sets in dwarf dandelions (*Krigia, *Asteraceae): additional information from internal transcribed spacer sequences of nuclear ribosomal DNA.* Pl. Syst. Evol., 190: 157-185.

OLMSTEAD R.G. and PALMER J.D, 1994. *Chloroplast DNA systematics: a review of methods and data analysis.* Amer. J. Bot., 81: 1205-1224.

PORTER J.M., 1993. *Phylogeny of Polemoniaceae based on nuclear ribosomal internal transcribed spacer sequences.* Amer. J. Bot., 80 (supplement, abstracts): 171.

ROSSI W., CORRIAS B., ARDUINO P., CIANCHI R. and BULLINI L., 1994. *Multilocus electrophoresis and European orchid systematics: the genus* Orchis *and related genera.* Proc. 14th World Orchid Conference. HMSO, Glasgow. Pp. 78-83.

SANDERSON M.J. and WOJECHOWSKI M.F., 1993. *Phylogenetic analysis of rates of diversification in Galegaeae (Fabaceae): Are there so many species of* Astragalus? Amer. J. Bot., 80 (supplement, abstracts): 174.

SCHLEGEL M., STEINBRÜCK G., HAHN K. and RÖTTGER B., 1989. *Interspecific relationships of ten European orchid species as revealed by enzyme electrophoresis.* Pl. Syst. Evol., 163: 107-119.

SUH Y., THIEN H.E. and ZIMMER E.A., 1993. *Molecular evolution and phylogenetic implications of internal transcribed sequences of ribosomal DNA in Winteraceae.* Amer. J. Bot., 80: 1042-1055.

Finito di stampare nel marzo 1996.