ORIGINAL PAPER

Genetic variation in time and space: the use of herbarium specimens to reconstruct patterns of genetic variation in the endangered orchid *Anacamptis palustris*

Salvatore Cozzolino · Donata Cafasso · Giuseppe Pellegrino · Aldo Musacchio · Alex Widmer

Received: 27 January 2006/Accepted: 30 August 2006/Published online: 21 October 2006 © Springer Science+Business Media B.V. 2006

Abstract Habitat alteration, fragmentation and destruction as a consequence of human impact are a global phenomenon. Here we document changes in genetic variation in the marsh orchid Anacamptis palustris as a consequence of such habitat changes. We examined historical specimens that were collected during the nineteenth and early twentieth centuries, prior to the most recent massive habitat changes affecting this species. Sequences of a hypervariable region in the plastid DNA, located in the $tRNA^{LEU}$ intron, from herbarium vouchers were compared with those from a near-exhaustive survey of the extant A. palustris populations on the Italian peninsula. It was found that private haplotypes and alleles found in small, extant populations were once widespread and more common in historic populations and that alleles, once present in historic populations, have gone extinct. In addition, genetic differentiation among populations has increased over time and haplotype frequencies significantly differ among historic and extant populations. These results document that human induced habitat changes have reduced genetic diversity and

S. Cozzolino (⊠) · D. Cafasso · Dipartimento delle Scienze Biologiche, Università degli Studi di Napoli Federico II, Via Foria, 223, 80139 Naples, Italy e-mail: cozzolin@unina.it

G. Pellegrino · A. Musacchio Dipartimento di Ecologia, Università della Calabria, Arcavacata di Rende, 87036 Cosenza, Italy

A. Widmer

Institute of Integrative Biology, Plant Ecological Genetics, ETH Zurich, 8092 Zurich, Switzerland

increased the importance of random genetic drift in this species. It is concluded that the analysis of herbarium specimens may provide important insights into changes of genetic diversity over time and may be critical for correct inference of the evolutionary history of rare and endangered species.

Keywords Anacamptis palustris · Plastid DNA · Herbarium specimens · Historic haplotype diversity · Temporal variation

Introduction

The possibility to compare patterns of genetic diversity between extinct and extant populations is of great importance in conservation genetics because it allows the assessment of past population structures, and documents the consequences of random genetic drift in populations that declined as a consequence of either natural or human-induced processes (Roy et al. 1994). Until recently, genetic studies of small and endangered populations have suffered from the problem that reliable estimates of the pre-exploitation diversity were usually not available. Consequently, the low levels of genetic diversity observed in many extant populations could not be used to reliably date eventually occurred past population reductions (Fréville et al. 1998; Tessier and Bernatchez 1999; Pichler and Baker 2000).

Genetic studies of rare and endangered species can be extended by adding historical samples from museum (Thomas et al. 1990; Ellegren 1991; Nielsen et al. 1999; Rosembaum et al. 2000) and herbarium collections (Saltonstall 2002). This progress has been made possible by the technical advancement of PCR-based methods that now allow the analysis of ancient specimens from minute tissue samples that may not compromise the value of rare, often irreplaceable, historic herbarium specimens (Maunder et al. 1999; Drábková et al. 2002). However, DNA extracted from such historic specimens is often highly degraded as a consequence of long sample storage (Towne and Devor 1990; Roy et al. 1994; Savolainen et al. 1995; Drábková et al. 2002). Consequently, short DNA fragments from molecules that are present in high copy numbers, such as the mitochondrial (Thomas et al. 1990, Taberlet and Bouvet 1992) and plastid DNA genomes (cpDNA), may represent the best targets for studies on herbarium specimens (Savolainen et al. 1995; Fay and Cowan 2001).

The phylogeographic history of the rare marsh orchid Anacamptis palustris (Orchidaceae) has recently been reconstructed using a highly polymorphic plastid region with a complex molecular evolution. Cozzolino et al. (2003a) referred to the different sequence rearrangements observed as haplotypes, and to the length polymorphisms observed in some haplotypes as alleles. Patterns of haplotype and allele variation in this hypervariable plastid region were found to be very useful for documenting, at a remarkably fine geographic scale, patterns of genetic variation among extant populations located along the Italian peninsula (Cozzolino et al. 2003a). The phylogeographic scenario that has been deduced suggests that A. palustris has a center of diversification in north-eastern Italy, where the highest plastid diversity was observed and where the largest extant populations exist (Cozzolino et al. 2003b). However, comparisons of genetic diversity among north-eastern Italian populations and central and southern Italian populations revealed a striking pattern of discontinuity. While north-eastern Italian populations along the Adriatic sea were characterized by having a large number of alleles at the widespread haplotype N (according to Cozzolino et al. 2003a), the central and southern Italian populations along the coasts of the Tyrrhenian Sea mostly harbored different, exclusive haplotypes or only alleles with small repeat numbers (short alleles) at the N haplotype. This haplotype pattern may be a consequence of the relatively recent establishment of populations along the coasts of the Tyrrhenian Sea by seeds carrying short alleles, which is realistic, because short alleles are common in the potential source populations in north-eastern Italy. The absence of long alleles could then be a consequence of the young age of these populations that has not allowed the evolution of long alleles in the populations from the Tyrrhenian coast. Alternatively, the observed pattern of genetic variation in these populations may be an artifact, brought about by small population sizes as a probable consequence of human impact (Petit et al. 1998; Cozzolino et al. 2003b). Strong bottlenecks could lead to a severe loss of allelic variation and the survival of only the most common alleles (Ellegren 2000). At the same time, the presence of exclusive haplotypes in populations along the coasts of Tyrrhenian Sea may either be a consequence of severe reductions of haplotype diversity and distribution due to genetic drift, or may have evolved locally.

Since Roman times, coastal marshes-the habitat of A. palustris-along the Italian peninsula have been extensively drained to reduce breeding grounds of the malaria vector Anopheles, and to recover suitable land for agriculture. Further massive reductions of these habitats have occurred in recent times, between the First and Second World War, when extensive drainage of coastal marsh habitats was promoted by the national government (Bevilacqua and Rossi-Doria 1984). Consequently, marsh habitats that formerly occurred over hundreds of kilometers along the Italian coasts, are now restricted to a few habitat patches that are often strongly isolated from neighboring patches. This strong anthropogenetic impact most likely has caused significant reductions of A. palustris population sizes along the Tyrrhenian coasts. In the absence of data from pre-drainage A. palustris populations it was unclear whether observed levels of genetic variation are a consequence of these most recent habitat destructions or date further back.

We therefore assessed genetic variation using historic herbarium specimens collected before the Second World War and compared the results with those previously reported for extant *A. palustris* populations. Specifically, we asked (a) whether levels of genetic diversity in extant populations are reduced compared to historic samples, (b) whether the rare and exclusive haplotypes found in extant populations from the Tyrrhenian coasts were once widespread, or whether they have evolved locally and failed to spread, and (c) whether long alleles at the N haplotype were once present in populations from the Tyrrhenian coasts and have been lost recently, or whether their presence in extant populations from the Tyrrhenian coast is due to recent recruitment from Adriatic populations.

Methods

DNA amplification and sequencing

Major Italian and Swiss herbarium collections were searched for historical specimens of *A. palustris* collected between the early nineteenth century and the Second World War (see Table 1 for the list of all historical specimens contributing data to this analysis). All herbarium specimens were visually inspected and all available information on the collection labels was recorded. For DNA analysis a small portion of a leaf tip from the herbarium specimen (approx. 5×5 mm) was sampled with a sterile forceps and DNA was extracted according to Doyle and Doyle (1987) and resuspended in 10 μ l of distilled water. PCR amplification of a part of the plastid *t*RNA^{LEU} intron was carried out using two specific primers and reaction conditions as described in Cozzolino et al. (2003c), with 5 μ l of resuspended DNA as PCR template.

Precautions were taken to guard against contaminations (Glenn et al. 1999; Miller and Waits 2003; Gilbert et al. 2005), although the use of species-specific PCR primers strongly reduced the likelihood of contaminations. DNA extractions and PCR amplifications were carried out in a separate building (University of Naples) where no A. palustris DNA had ever been investigated before and control extractions (without plant material) and control PCR reactions (without DNA) were done as negative controls. Amplification products were visualized on 2% Metaphore gels (FMS, Rockland, ME, USA), using a 50 bp ladder as standard (FMS) stained with ethidium bromide and photographed using a digital camera. Amplification products that did not produce visible bands on the gel were subject to ten additional amplification cycles by adding additional polymerase. Finally, the remaining 5 μ l of resuspended DNA were used either for independent verification of the results in an independent laboratory (University of Calabria) or as material for additional experiments for samples that did not successfully amplified in the first attempt. Approximately one-third of the samples were independently examined twice.

All PCR products were sequenced to check the repeat type (haplotype) and the nature of the length polymorphism (allelic variation). Out of 110 examined herbarium specimens, a total of 89 samples (81%), among which 85 samples from coasts of Tyrrhenian Sea, were successfully amplified and sequenced. PCR products were purified and cycle-sequenced using BigDye Terminator sequencing chemistry on an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Sequence Navigator software (Applied Biosystems) was used to inspect electropherograms and to align sequences.

Data analysis

Analyses were performed to determine if the pattern of genetic variation in the extant populations is different from that of the historic samples. Because herbarium specimens have typically not been collected for the purpose of conducting analyses of genetic diversity and population structure, sample sizes are often inadequate for a given locality and time period and the amount of information available for the different populations is usually limited (Nielsen et al. 1999). To partially overcome this problem in the present study we grouped individuals from geographically close populations together into regions. Thus we carried out population pooling both for historic samples and for extant samples in order to gather, wherever possible, comparable historic and extant samples for each region.

This methodological choice can be justified by the fact that most of the historic and extant populations (see Fig. 1) originated from the same, narrowly circumscribed geographic area (typically marshes along the coasts that are separated by a few kilometers). Consequently we do not expect that pooling samples from different, but geographically close, populations substantially distorts estimates of genetic diversity (H_T) and population differentiation (G_{ST}), which can occur when samples from distant and isolated populations are pooled.

The change over time in alleles frequencies can also be explained by the stochastic effects alone, as genetic drift and sampling errors. Thus, temporal differences in allelic frequencies were analyzed within regions following Waples (1989) in order to test whether significant temporal variation in haplotype frequencies are due entirely to stochastic process (as drift and sampling error) in a finite population. Analysis were conducted assuming 5 years per orchid generation and selection effects on haplotype frequencies were ignored here since we are concerned with neutral markers. A significant test (at the 0.05 level) implies that pure drift and sampling error was not sufficient to explain the difference in the haplotype frequencies (i.e., haplotype frequencies have changed greater than expected from drift alone) and that other effects, such as migrations or bottlenecks, can be considered to explain the changes.

Patterns of diversity and population differentiation were then compared between these historic samples and a comparable set of extant populations (a subset of those sampled in Cozzolino et al. 2003a) from the Tyrrhenian coast. We estimated population subdivision (G_{ST}) and genetic diversity (H_T) in the two datasets (historic and extant) and compared patterns of pairwise geographic differentiation among population pairs between historic and extant samples. Levels of population subdivision for plastid genomes using unordered alleles (G_{ST}) and genetic diversity within populations

Table 1	List	of	examined	herbarium	Α.	palustris	accessions
---------	------	----	----------	-----------	----	-----------	------------

	Sample	Herbarium	Collection locality	Region	Year	Haplotype
1	T 1A	University Federico II-NAP	Campi Flegrei	Campania-Tyrrhenian	Nineteenth century	D1
2	T 1B	University Federico II-NAP	Campi Flegrei	Campania-Tyrrhenian	Nineteenth century	D1
3	T 1C	University Federico II-NAP	Campi Flegrei	Campania-Tyrrhenian	Nineteenth century	D1
4	R29	University La Sapienza-ROMA	Monte Nuovo	Campania-Tyrrhenian	1883	D1
5	TO6	University Torino-TO	Paludi del Fusaro	Campania-Tyrrhenian	1897	No amplification
6	R43	University La Sapienza-ROMA	Ostia	Latium-Tyrrhenian	1832	D1
7	R45	University La Sapienza-ROMA	Fiumicino	Latium-Tyrrhenian	1864	D1
8	R47	University La Sapienza-ROMA	Fiumicino	Latium-Tyrrhenian	1896	A2
9	R48	University La Sapienza-ROMA	Fiumicino	Latium-Tyrrhenian	1896	A2
10	R49	University La Sapienza-ROMA	Flora romana	Latium-Tyrrhenian	Nineteenth century	D1
11	R50	University La Sapienza-ROMA	Flora romana	Latium-Tyrrhenian	Nineteenth century	D1
12	R51	University La Sapienza-ROMA	Flora romana	Latium-Tyrrhenian	Nineteenth century	D1
13	R52	University La Sapienza-ROMA	Flora romana	Latium-Tyrrhenian	Nineteenth century	D1
14	R53	University La Sapienza-ROMA	Flora romana	Latium-Tyrrhenian	Nineteenth century	DI
15	K54	University La Sapienza-ROMA	Flora romana	Latium-Tyrrhenian	Nineteenth century	DI
10	K33	University La Sapienza-ROMA	Flora romana	Latium-Tyrrhenian	Nineteenth century	DI D1
1/	K30 D57	University La Sapienza-ROMA	Flora romana	Latium-Tyrrnenian	Nineteenth century	DI D1
10	R3/ D59	University La Sapienza POMA	Flora romana	Latium-Tyrrhenian	Nineteenth century	DI D1
20	R30 P62	University La Sapienza ROMA	Fiumicino	Latium Tyrrhenian	1806	
20	R02 P63	University La Sapienza ROMA	Rami di Tivoli	Latium Tyrrhenian	1890	A2 D1
$\frac{21}{22}$	R63	University La Sapienza-ROMA	Bagni di Tivoli	Latium-Tyrrhenian	1896	D1
23	R65	University La Sapienza-ROMA	Bagni di Tivoli	Latium-Tyrrhenian	1896	A2
23	R66	University La Sapienza-ROMA	Bagni di Tivoli	Latium-Tyrrhenian	1896	A2
25	R70	University La Sapienza-ROMA	Località Bagni	Latium-Tyrrhenian	Nineteenth century	D1
26	R71	University La Sapienza-ROMA	Bagni ambula alla	Latium-Tyrrhenian	1943	D1
			Martellona			
27	R72	University La Sapienza-ROMA	Bagni ambula alla Martellona	Latium-Tyrrhenian	1943	A2
28	R73	University La Sapienza-ROMA	Bagni ambula alla Martellona	Latium-Tyrrhenian	1943	D1
29	R74	University La Sapienza-ROMA	Bagni di Tivoli	Latium-Tyrrhenian	1943	A2
30	R75	University La Sapienza-ROMA	Bagni di Tivoli	Latium-Tyrrhenian	1943	D1
31	R76	University La Sapienza-ROMA	Bagni di Tivoli	Latium-Tyrrhenian	1943	N1
32	R44	University La Sapienza-ROMA	Ostia	Latium-Tyrrhenian	1880	No amplification
33	R77	University La Sapienza-ROMA	Bagni di Tivoli	Latium-Tyrrhenian	1943	No amplification
34	GS 99	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	D1
35	GS 102 A	University Federico II-NAP	Spaccaforno	Sicily-Tyrrhenian	Nineteenth century	D1
36	GS 102 B	University Federico II-NAP	Milazzo	Sicily-Tyrrhenian	Nineteenth century	D1
37	GS 103 A	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	DI
38	GS 103 B	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	DI
39	GS 103 C	University Federico II-NAP	Sicilia	Sicily-Tyrrnenian	Nineteenth century	DI D1
40	GS 104	University Federico II-NAP	Sicilia	Sicily-Tyrrnenian	Nineteenth century	DI D1
41	GS 105 A	University Federico II-NAP	Spaceaforno	Sicily-Tyrrhenian	Nineteenth century	D1
42	GS 105 D	University Federico II-NAP	Spaceaforno	Sicily-Tyrrhenian	Nineteenth century	D1
43	GS 105 C	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	D1
45	GS 106 R	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	D1
46	GS 107 A	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	D1
47	GS 107 B	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	D1
48	GS 108	University Federico II-NAP	Mazzara del Vallo	Sicily-Tyrrhenian	Nineteenth century	D1
49	GS 109 A	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	D1
50	GS 109 B	University Federico II-NAP	Sicilia	Sicily-Tyrrhenian	Nineteenth century	D1
51	4496/22A	University Florence-FI	S.Ciro	Sicily-Tyrrhenian	1848	D1
52	4496/23A	University Florence-FI	S.Ciro	Sicily-Tyrrhenian	1898	D1
53	4496/24B	University Florence-FI	Siracusa	Sicily-Tyrrhenian	1898	D1
54	4496/24C	University Florence-FI	Siracusa	Sicily-Tyrrhenian	1898	D1
55	4496/25A	University Florence-FI	Terranova	Sicily-Tyrrhenian	1871	D1
56	4496/25B	University Florence-FI	Terranova	Sicily-Tyrrhenian	1871	D1
57	4496/26A	University Florence-FI	Terranova	Sicily-Tyrrhenian	Nineteenth century	D1
58	4496/24A	University Florence-FI	Siracusa	Sicily-Tyrrhenian	1898	No amplification

Table 1 continued

	Sample	Herbarium	Collection locality	Region	Year	Haplotype
59	TO1	University Torino-TO	Terranova	Sicily-Tyrrhenian	Nineteenth century	No amplification
60	TO2	University Torino-TO	Terranova	Sicily-Tyrrhenian	Nineteenth century	No amplification
61	TO3	University Torino-TO	Terranova	Sicily-Tyrrhenian	Nineteenth century	No amplification
62	R20	University La Sapienza-ROMA	Macciuccoli	Tuscany-Tyrrhenian	Nineteenth century	D1
63	R22	University La Sapienza-ROMA	Macciuccoli	Tuscany-Tyrrhenian	Nineteenth century	D1
64	Z 2A	University Zurich-Z	Marina dei Ronchi	Tuscany-Tyrrhenian	1948	N1
65	Z 2B	University Zurich-Z	Marina dei Ronchi	Tuscany-Tyrrhenian	1948	N1
66	Z 2C	University Zurich-Z	Marina dei Ronchi	Tuscany-Tyrrhenian	1948	N1
67	Z 2D	University Zurich-Z	Marina dei Ronchi	Tuscany-Tyrrhenian	1948	N1
68	Z 3B	University Zurich-Z	Marina dei Ronchi	Tuscany-Tyrrhenian	1948	N1
69	Z 3C	University Zurich-Z	Marina dei Ronchi	Tuscany-Tyrrhenian	1948	N1
70	4496/5A	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1881	D1
71	4496/5B	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1881	D1
72	4496/6A	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1881	D1
73	4496/6B	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1881	D1
74	4496/6C	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1881	A2
75	4496/7A	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1862	D1
76	4496/7B	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1862	D1
77	4496/7C	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1862	D1
78	4496/7D	University Florence-FI	Palude d'Asciano	Tuscany-Tyrrhenian	1862	D1
79	4496/82	University Florence-FI	Palude di Mondu-	Tuscany-Tyrrhenian	1860	N1
1)	4790/0a	Oniversity Therence-11	nito	ruseany-rynneman	1000	111
80	4496/9a	University Florence-FI	Orbetello-Anse- donia	Tuscany-Tyrrhenian	1892	D1
81	4496/10A	University Florence-FI	Castagnolo (PI)	Tuscany-Tyrrhenian	1891	D1
82	4496/10C	University Florence-FI	Castagnolo (PI)	Tuscany-Tyrrhenian	1891	D1
83	4496/10D	University Florence-FI	Castagnolo (PI)	Tuscany-Tyrrhenian	1891	D1
84	4496/12A	University Florence-FI	Palude di Minighi- ano	Tuscany-Tyrrhenian	1894	N2
85	4487/5A	University Florence-FI	Maremma orbetellana	Tuscany-Tyrrhenian	1892	D1
86	4487/6A	University Florence-FI	Marina di Massa	Tuscany-Tyrrhenian	1914	N2
87	4487/7C	University Florence-FI	Porto Ercole	Tuscany-Tyrrhenian	1856	D1
88	4487/7B	University Florence-FI	Porto Ercole	Tuscany-Tyrrhenian	1856	D1
89	4487/7A	University Florence-FI	Porto Ercole	Tuscany-Tyrrhenian	1856	N2
90	4487/8C	University Florence-FI	Marina di Pisa	Tuscany-Tyrrhenian	1916	D1
91	4487/8D	University Florence-FI	Marina di Pisa	Tuscany-Tyrrhenian	1916	D1
92	4487/8E	University Florence-FI	Marina di Pisa	Tuscany-Tyrrhenian	1916	D1
93	Z 3A	University Zurich-Z	Marina dei Ronchi	Tuscany-Tyrrhenian	1948	No amplification
94	R21	University La Sapienza-ROMA	Palude di Massac- ciuccoli	Tuscany-Tyrrhenian	Nineteenth century	No amplification
95	R34	University La Sapienza-ROMA	Pietra a palude	Tuscany-Tyrrhenian	Nineteenth century	No amplification
96	4487/8A	University Florence-FI	Marina di Pisa	Tuscany-Tyrrhenian	1916	No amplification
97	4487/8B	University Florence-FI	Marina di Pisa	Tuscany-Tyrrhenian	1916	No amplification
98	4496/14A	University Florence-FI	Pisa	Tuscany-Tyrrhenian	1838	No amplification
99	4496/3A	University Florence-FI	Litorale veneto	Adriatic	1845	D1
100	4496/4A	University Florence-FI	Estuario Venezia	Adriatic	1926	D1
101	4487/2A	University Florence-FI	Molino di Stabbe	Adriatic	1913	N6
102	4487/2B	University Florence-FI	Molino di Stabbe	Adriatic	1913	D1
103	4487/2C	University Florence-FI	Molino di Stabbe	Adriatic	1913	No amplification
104	4487/4 A	University Florence-FI	Parma	Northern-Italy	1872	D1
105	4487/4B	University Florence-FI	Parma	Northern-Italy	1872	D1
106	GG 33B	University Federico II-NAP	Parma	Northern-Italy	Nineteenth century	 D1
107	GG 33C	University Federico II-NAP	Parma	Northern-Italy	Nineteenth century	D1
108	GG33A	University Federico II-N ΔP	Parma	Northern-Italy	Nineteenth century	No amplification
100	4496/304	University Florence-FI	Tunisia	Northern-Africa	1888	D1
110	4496/30R	University Florence-FI	Tunisia	Northern-Africa	1888	D1
110	170,000	Chiversity i lorence-i i	i amora	rormorn / milea	1000	

Fig. 1 Geographic location of extant (*filled*) and historic (*empty*) populations from the Tyrrhenian coast

- 1. Lago Porta
- Macciuccoli, Palude di Masacciuccoli
 Torre Lago, Marina dei Ronchi
- Marina di Massa, Marina di Pisa, Pisa
- 4. Castagnolo, Palude d'Asciano, Palude di Mondunito, Palude di Minighiano, Pietra a palude
- 5. Orbetello, Orbetello-Ansedonia, Maremma orbetellana
- 6. Porto Ercole
- 7. Viterbo
- Maccarese, Fiumicino, Ostia
 Laghi del Vescovo
- Bagni di Tivoli, Località Bagni, Bagni alla Martellona
- 11. Castelvolturno
- 12. Campi Flegrei, Montenuovo
- Milazzo
- Mazzara del Vallo
 S. Ciro
- 16. Terranova
- 17. Siracusa
- 18. Spaccaforno

(H_T) were calculated according to Pons and Petit (1995) using the software HAPLODIV, and analysis of molecular variance (AMOVA, Excoffier et al. 1992) was calculated using ARLEQUIN version 3.0 (Excoffier et al. 2005) to estimate the partitioning of variance into within and among populations components. We tested isolation by distance using the Mantel test as implemented in the software TFPGA version 1.3 (Miller 1997) using as geographic distances beeline distances among pairs of populations and as genetic distances Nei's (1987) pairwise distances as calculated with ARLEQUIN version 3.0. Rarefaction was applied to correct diversity estimates for differences in sample sizes between extant and historic samples by employing the option RAREFACTION in the software CONTRIB (Petit et al. 1998). Finally, to discern whether a haplotype found in the historic samples but not in the samples of extant populations (and viceversa) was missed because of extinction or because if insufficient sampling, we used the approach given by

Table 2 Plastid haplotypes detected in extant (bold) and historic (italic) A. palustris populations

Regions	Sample sizes	Haplotypes						
		A	В	С	D	Е	N	
Tyrrhenian- Campania	70 , <i>4</i>				56 , <i>4</i>	4 , 0	10 , <i>0</i>	
Tyrrhenian- Latium	183 , <i>26</i>	0, 7	14 , <i>0</i>		138 , <i>18</i>		31 , <i>1</i>	
Tyrrhenian- Sicily	0 , <i>24</i>				0, 24			
Tyrrhenian- Tuscany	63 , <i>31</i>	7 , 1	10 , <i>0</i>		0 , <i>20</i>		46 , <i>10</i>	
Adriatic	322 , 4			1 , <i>0</i>	52 , 3		269 , 1	



Rosenbaum et al. (2000). By assuming that each herbarium specimens represented a random sample from the historic populations and that no new haplotype has arisen since the herbarium specimens were taken, the probability of observing a specimen with a rare haplotype is equal to the relative frequency of this rare haplotypes in the population. The probability of detecting a given haplotype was set at 95 and 99%, respectively and the maximal frequency of rare haplotypes in the historical (n = 85) and extant populations (n = 316) was estimated. The resulting maximal frequency values are in essence the frequency estimates for any possible rare haplotype that was not detected, which makes the observed data ≥ 95 and $\geq 99\%$ probable since the time when the specimens were collected.

Results

Three distinct haplotypes (A, D, and N) were detected among the 85 herbarium specimens from the Tyrrhe-

Table 3 Numbers and frequencies of haplotypes in populations from the Tyrrhenian coast

Haplotype	Histo	ric	Extant	t	χ^2
	n	%	n	%	
A	8	9.41	7	2.21	9.25*
В	0	0.00	24	7.59	6.37*
D	66	77.65	194	61.39	3.86*
Е	0	0.00	4	1.27	1.07
Ν	11	12.94	87	27.54	5.82*
Total	85		316		

Adjusted χ^2 values for test of the hypothesis that changes in haplotype frequency between historic and extant populations can be attributed to stochastic drift. Asterisks indicate significant chi-square test (P < 0.05; df = 1)

nian coast (Tables 2, 3, Fig. 1). Of these, the N haplotype showed allelic variation due to changes in repeat copy numbers $(N_1 \text{ and } N_2)$ (Table 4). A new allele of haplotype A was found in the historic samples. This allele differs from the original A allele, which was found in the extant Orbetello population (Cozzolino et al. 2003a), in the number of repeats: five repeats in Orbetello and two repeats only in the historic samples. The alleles at this haplotype will, from now on, be referred to as A_5 and A_2 alleles, respectively (Table 4). Haplotype D was found in all historic A. palustris populations from the Italian peninsula and Sicily and was the most common haplotype in the historic sample (Fig. 2). In Campania and Sicily, only haplotype D was found, while in Tuscany and Latium we also found haplotypes A (allele A_2) and N. Allele N_1 was found in both Tuscany and Latium, whereas allele N2 was found only in Tuscany. The two haplotypes B and E found in extant populations from the Tyrrhenian coast (Cozzolino et al. 2003a) were not found in the historic sample (Table 2, Fig. 2). All haplotypes but E significantly differed in their frequency between historic and extant samples (Table 3).

Genetic differentiation among regions, as estimated by G_{ST} , was significantly lower (P = 0.006, *t*-test) among the historic samples from the Tyrrhenian coast among extant populations than $(G_{ST} = 0.133)$ $(G_{ST} = 0.320)$. Accordingly, AMOVA revealed that only 18% of the variation was found among regions for the historic samples, whereas 40% of the variation was found among regions in the extant sample. The weak genetic differentiation among regions was reflected in the analysis of isolation by distance. No significant association between geographic and genetic distances was found for the historic samples (r = 0.656, P > 0.1) whereas evidence for isolation by distance was found among extant samples (r = 0.708, P < 0.01).

Table 4 Numbers and frequencies of alleles at differenthaplotypes in populations from the Tyrrhenian coast

Haplotypes	Alleles	Hist	oric	Extant	
		n	%	n	%
A	A_2	8	9.41	0	_
	$\overline{A_5}$	0	_	7	2.21
В	\mathbf{B}_n	0	_	24	7.59
D	\mathbf{D}_1	66	77.65	194	61.38
E	E	0	_	4	1.27
N	N_1	8	9.41	33	10.45
	N_2	3	3.53	43	13.62
	N_3	0	_	7	2.21
	N ₆	0	_	3	0.95
	N ₁₁	0	_	1	0.32
Total	_	85		316	

The estimate of gene diversity (based on all extant and historic sampled individuals) was significantly lower (P = 0.012, *t*-test) for the historic sample ($H_T = 0.352$) compared to the sample of extant populations ($H_T = 0.790$). However, rarefaction revealed that for the sample size achieved with the historic samples from the Tyrrhenian coast (85 individuals), only 3.6 haplotypes can be expected, which is close to the 3 haplotypes observed in the historic sample.

Using the approach implemented by Rosembaum et al. (2000) we estimated that we would have been able to detect a rare haplotype in the sample of extant populations from the Tyrrhenian coast with a probability of 95% if it had occurred as low as 0.95% in the historic sample, and with a probability of 99% if it had occurred at a rate as low as 1.45%.

At the same time, we would have been able to detect a rare haplotype in the sample of historic populations with a probability of 95% if it had occurred at a rate as low as 3.47% in the extant historic sample from the Tyrrhenian coast, and with a probability of 99% if it had occurred at a rate as low as 5.28%.

Discussion

Our comparison of sampling locations for herbarium specimens with the distribution of extant A. palustris populations (Cozzolino et al. 2003a) confirms that the distribution of A. palustris has been drastically reduced during the twentieth century (Fig. 1). Approximately 80% of collection sites described on the labels of the investigated herbarium specimens are either no longer marsh habitats or, if marsh habitats still exist, the populations of A. palustris are extinct. Our results demonstrate that this widespread deterioration, reduction, and destruction of suitable habitat and the concomitant loss of A. palustris populations has left its signature in the gene pool of the remaining extant populations. These populations have been sampled nearly exhaustively by Cozzolino et al. (2003a), with the only exception being one small population that is located in a bird reserve and cannot be visited during the orchid's flowering time which overlaps with bird nesting times in this area. Additionally, most of the extant populations correspond or are very close to the historic ones (see Fig. 1). It is thus unlikely that plastid haplotypes or alleles founds in herbarium specimens remained unsampled.

At first sight, evidence for a substantial loss of genetic variation in the recent history of the species may be absent. The analysis of the extant *A. palustris* populations has revealed a total of five haplotypes

Fig. 2 Detected geographic distribution of plastid haplotypes in extant and historic populations of *A. palustris* along the Italian peninsula



(Cozzolino et al. 2003a), in contrast to the three haplotypes detected in the historic sample (Table 2, Fig. 3). The apparent increase of haplotype number and gene diversity in extant populations could be due to new mutations that have arisen relatively recently or due to haplotypes that have been introduced into extant populations by recent immigrants, or could be a sampling artifact. The numbers of haplotypes detected depends strongly on sample sizes, which differed markedly between the study of extant populations from the Tyrrhenian coast (Cozzolino et al. 2003a, 316 examined specimens) and this study (85 specimens). The problem is aggravated when haplotypes occur at unequal frequencies, because the detection of rare haplotypes becomes less likely. For instance, the rare haplotype B has been found only in two extant populations, for which no historic samples from nearby sites were available, and the E haplotype, was found in four specimens only from Campania and thus occurred only in a frequency of 1.27% in the extant sample. Consequently, we propose that the two haplotypes E and B, which were rare in extant populations, were not found in the much smaller historic sample because of the limited sample size.

Generally, the limited sample size and the methodological bias in the collecting procedure for herbarium specimens (i.e., few individuals only from each locality) may render the herbarium collection less amenable for detecting rare haplotypes and alleles, but they are still useful for detecting variation in the distribution and frequency of common haplotypes and alleles. This is particularly true when several samples from the same (or from very close) locality are available in historic herbaria, as in the present case study of *A. palustris*.

The reverse pattern, i.e., the finding of a haplotype/ allele in the historic sample but not in the extant populations may again be due to insufficient sampling or could be the consequence of random genetic drift following population reductions and bottlenecks. The two alleles that were found in the historic sample, but were either absent (allele A₂) or rare (allele N₁) in the extant populations, were present in relatively high frequencies in the historic sample, which makes it highly unlikely that these haplotypes were still present in extant populations, but missed due to insufficient sampling. For instance, allele A₂, which occurred with a frequency of 9.4% in the herbarium samples (Tables 3, 4), would have remained undetected in the extant sample with a probability of 2.8×10^{-14} . This indicates that its absence from extant *A. palustris* populations is likely due to extinction.

Waples' test performed on haplotype frequencies was significant for all haplotypes but haplotype E. This implies that pure drift or sampling errors are not sufficient explanation of the observed temporal changes in haplotype frequencies and thus there are evidence of greater than expected change in haplotype frequencies due to population reduction.



Fig. 3 Relative haplotype frequencies in extant (*empty*) and historic (*filled*) populations from the Tyrrhenian coast

Population structure has equally been affected by population reduction. The much higher genetic differentiation among pools of extant populations, compared to historic populations, indicates that substantial fragmentation has occurred among populations along the Tyrrhenian coast. This effect may have been renforced by local random genetic drift which is expected to increase genetic differentiation among populations, even when population numbers remain constant. The increase in genetic differentiation strongly suggests that the remaining *A. palustris* populations are not sufficiently close or large to exchange migrants at a rate that could balance genetic drift.

Phylogeographic studies often have problems to explain the occurrence of exclusive haplotypes in some isolated populations (Schaal et al. 1998). The haplotype pattern found in our earlier analysis of extant populations left open several possible hypotheses for the origin of these isolated populations or for the origin of the exclusive haplotypes. The addition of data from historic samples may help distinguishing between alternative scenarios. For example, the extant population from Viterbo (Fig. 1) (Cozzolino et al. 2003a), which grows in a sulphuric inland marsh and not in coastal marshes, is presently small (50 flowering individuals in 2002) and fixed for allele N_1 at haplotype N. In the extant sample, this allele has been found outside of Viterbo only in two individuals from a population from the Tyrrhenian coast (Castelvolturno) and three individuals from the Adriatic Sea (Cozzolino et al. 2003a). The frequency of this allele in the historic sample is in stark contrast to its rarity in present-day populations. Allele N₁ has occurred in the past also in populations Marina dei Ronchi, Palude di Mondunico and Bagni di Tivoli (Fig. 1). All these populations are nowadays extinct. Thus, the historic samples demonstrate that allele N₁ has not evolved locally in Viterbo and thus does not flag an evolutionary uniqueness of that particular sulfuric inland marsh population; instead, the current distribution of N₁ is relictual. Similarly, haplotype A, which was found to occur among extant populations exclusively in the small population of Orbetello in Tuscany, has now been found, in the form of allele A_2 , to be relatively common in the historic sample from the Tyrrhenian coast (Table 4). This finding further allows to reject the earlier hypothesis (Cozzolino et al. 2004) that the Orbetello population originated from a Balearic population of A. robusta, where a similar haplotype (the D_n haplotype in Cozzolino et al. 2004) occurs.

The three remaining extant populations from Tuscany carry haplotypes A, B, and N, but lack haplotype D, which separates them from all other, more southern populations from the Tyrrhenian coast, where D is common. This genetic discontinuity is again the result of a loss of haplotype D from Tuscany, as evidenced by the frequency of this haplotype in the historic sample from Tuscany. Thus, populations from Tuscany are not derived from Adriatic populations, but clearly belong to a lineage that is distributed along the Tyrrhenian coast and extends into Sicily and northern Africa as confirmed by the analyses of few additional historic samples from Tunisia (see Table 1).

Extant populations from the Adriatic coast were found to be highly variable at haplotype N, where a large number of alleles with different repeat numbers were observed (Cozzolino et al. 2003b) and a long N allele has also been found in corresponding historic samples (Table 2, Fig. 2).

Among extant populations from the Tyrrhenian coast, the diversity at haplotype N was much lower. This pattern was confirmed by the results from the historic sample. Historic populations from the Tyrrhenian coast did not harbor, at a level that could be detected with our sampling, long alleles at haplotype N (Tables 2, 4, Fig. 4). In fact, given our sample size (85), we would have detected with a probability of 99% a long allele at haplotype N if long alleles had occurred with a frequency of at least 5.28%. When present, haplotype N occurs in the historic sample from the Tyrrhenian coast only in the form of short alleles, N₁ or N_2 (Table 4, Fig. 4), which suggests that the current sporadic presence of long alleles in the extant populations from the Tyrrhenian coast, Torre Lago and Castelvolturno (Fig. 1), is likely due or to the reduction of a previously larger allelic variation at this haplotype, or to long distance dispersal events from Adriatic populations, where long alleles presumable evolved and are still common. Although much can be learned from the comparison of extant and historic samples, one has to keep in mind that in a geographic area such as Italian peninsula, that has been heavily influenced by human presence during the last 3,000 years,



Fig. 4 Relative allele frequencies at haplotype N in extant (*empty*) and historic (*filled*) populations from the Tyrrhenian coast

numerous unrecorded events have occurred prior to the collection of our historic samples, and these may already have shaped the genetic diversity of populations.

The comparison of present day and historic haplotype distributions helps us not only to better understand the evolutionary history of A. palustris, but may also provide guidelines for the conservation or reintroduction of the species. Although the species has minute wind-dispersed seeds, which are typically found in orchids (Arditti and Ghani 2000), and can occasionally disperse to relatively distant sites, the low number of long alleles at haplotype N that are found outside of the Adriatic coast, where they most likely evolved, indicates that long distance recruitment events are nevertheless rare. The populations that are able to send out migrants are the only remaining large populations from the Adriatic coast. The small extant populations from the Tyrrhenian coast are no longer able to exchange sufficient migrants to prevent genetic drift and thus genetic differentiation among populations is increasing. Where suitable habitat still exists, reintroduction projects should be carried out, because a natural colonization of habitat patches is unlikely. For this purpose, the value of analyzing historic samples lies in identifying suitable source populations.

For example on Sicily, where *A. palustris* is now extinct (Fig. 1), all populations harbored haplotype D (Table 2, Fig. 2). If the species were to be reintroduced to Sicily, extant populations from southern Italy that carry the D haplotype may be used as genetic resources. From these, individuals should be selected that currently grow under similar ecological conditions as those present in the reintroduction area.

Overall, the combined analysis of genetic variation in historic and extant populations of *A. palustris* has revealed haplotype extinctions, a low frequency of long distance dispersal events in a time window of approximately one and a half centuries, and the disappearance of a substantial number of the sites from which the plant has been reported in the past. Together, these findings strongly emphasize that *A. palustris* populations from the Tyrrhenian coast are under sever risk of extinction and must be protected.

Phylogeographic and population genetic studies of other plant species, especially those that have experienced population declines and range reductions, could greatly benefit from including herbarium specimens into their sampling. Otherwise, the risk of incorrectly inferring the species' evolutionary history from observed patterns of genetic diversity is high. Acknowledgments The authors thank the herbarium institutions (NAP, FI, ROMA, TO, Z/ZT) for providing access to their historic collections and for permitting sampling of specimens. In particular we thank the herbarium curators, namely M. Baltisberger, P. Cuccuini, L. Guglielmone, A. Millozza, and A. Santangelo for their invaluable help in specimen identification, label translation, and sample shipment. The authors also thank three anonymous reviewers for their comment that contributed to improve the MS. The present study was supported by the PRIN program of the Italian Ministry of the University and Scientific research and by a Swiss Federal Institute of Technology (ETH) grant to AW.

References

- Arditti J, Ghani AKA (2000) Tansley review No. 110—numerical and physical properties of orchid seeds and their biological implications. New Phytol 145:67–421
- Bevilacqua P, Rossi-Doria M (1984) Lineamenti per una storia delle bonifiche in Italia dal '700 a oggi. Laterza, Roma-Bari
- Cozzolino S, Cafasso D, Pellegrino G, Musacchio A, Widmer A (2003a) Fine-scale phylogeographical analysis of Mediterranean *Anacamptis palustris* (Orchidaceae) populations based on chloroplast minisatellite and microsatellite variation. Mol Ecol 12:2783–2792
- Cozzolino S, Noce ME, Musacchio A, Widmer A (2003b) Variation at a chloroplast minisatellite locus reveals the signature of habitat fragmentation and genetic bottlenecks in the rare orchid *Anacamptis palustris* (Orchidaceae). Am J Bot 90:1681–1687
- Cozzolino S, Cafasso D, Pellegrino G, Musacchio A, Widmer A (2003c) Molecular evolution of a plastid tandem repeat locus in an orchid lineage. J Mol Evol 57:41–49
- Cozzolino S, Cafasso D, Pellegrino G, Musacchio A, Widmer A (2004) Hypervariable plastid locus variation and intron evolution in the *Anacamptis palustris* lineage. Genome 47:999–1003
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11– 15
- Drábková L, Kirschner J, Vlcek C (2002) Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens of Juncaceae. Plant Mol Biol Rep 20:161–175
- Ellegren H (1991) DNA typing of museum birds. Nature 354:113
- Ellegren H (2000) Heterogeneous mutation processes in human microsatellite DNA sequences. Nat Genet 24:400–402
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online
- Fay M, Cowan RS (2001) Plastid microsatellites in *Cypripedium* calceolus (Orchidaceae): genetic fingerprints from herbarium specimens. Lindleyana 16:151–156
- Fréville H, Colas B, Ronfort J, Riba M, Olivieri I (1998) Predicting endemism from population structure of a widespread species: case study in *Centaurea maculosa* Lam. (Asteraceae). Conserv Biol 12:1269–1278
- Gilbert MTP, Bandelt H, Hofreiter M, Barnes I (2005) Assessing ancient DNA studies. Trends Ecol Evol 20:541–544

- Glenn TC, Stephan W, Braun MJ (1999) Effects of population bottleneck on whooping crane mitochondrial DNA variation. Conserv Biol 13:1097–1107
- Maunder M, Culham A, Bordeu A, Allainguillaume J, Wilkinson M (1999) Genetic diversity and pedigree for Sophora toromiro (Leguminosae): a tree extinct in the wild. Mol Ecol 8:725–738
- Miller MP (1997) Tools for population genetic analyses (TFPGA) 1.3: a windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author
- Miller CR, Waits LP (2003) The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): implications for conversation. Proc Natl Acad Sci USA 100:4334–4339
- Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York
- Nielsen EE, Hansen MM, Loeschcke V (1999) Genetic variation in time and space: microsatellite analysis of extinct and extant populations of Atlantic salmon. Evolution 53:261–268
- Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12:844–855
- Pichler FB, Baker GS (2000) Loss of genetic diversity in the endemic Hector's dolphin due to fisheries-related mortality. Proc R Soc B Biol Sci 267:97–102
- Pons O, Petit RJ (1995) Estimation, variance and optimal sampling of gene diversity. 1. Haploid locus. Theor Appl Genet 90:462–470
- Rosenbaum HC, Egan MG, Clapham PJ, Brownell RL Jr, Malik S, Brown MW, White BN, Walsh P, Desalle R (2000) Utility of North Atlantic right whale museum specimens for

assessing changes in genetic diversity. Conserv Biol 14:1837-1842

- Roy MS, Girman DJ, Taylor AC, Wayne RK (1994) The use of museum specimens to reconstruct the genetic variability and relationships of extinct populations. Experientia 50:551–557
- Saltonstall K (2002) Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. Proc Natl Acad Sci USA 99:2445–2449
- Savolainen V, Cuénoud P, Spichiger R, Martinez MDP, Crèvecoeur M, Manen JF (1995) The use of herbarium specimens in DNA phylogenetics: evaluation and improvement. Plant Syst Evol 197:87–98
- Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7:465–474
- Taberlet P, Bouvet J (1992) Bear conservation genetic. Nature 358:197
- Tessier N, Bernatchez L (1999) Stability of population structure and genetic diversity across generations assessed by microsatellites among sympatric populations of landlocked Atlantic salmon (*Salmo salar* L.). Mol Ecol 8:169–179
- Thomas WK, Pääbo S, Villablanca FX, Wilson AC (1990) Spatial and temporal continuity of kangaroo rat populations shown by sequencing mitochondrial DNA from museum specimens. J Mol Evol 31:101–112
- Towne B, Devor EJ (1990) Effect of storage time and temperature on DNA extracted from whole blood samples. Hum Biol 62:301–306
- Waples RB (1989) Temporal variation in allele frequencies: testing the right hypothesis. Evolution 43:1236–1251