Molecular Ecology (2004) **13**, 109–122 doi: 10.1046/j.1365-294X.2003.02022.x

Phylogeographical history of the sponge *Crambe crambe* (Porifera, Poecilosclerida): range expansion and recent invasion of the Macaronesian islands from the Mediterranean Sea

S. DURAN,* G. GIRIBET† and X. TURON*
*Department of Animal Biology (Invertebrates), University of Barcelona, 645 Diagonal Avenue, 08028 Barcelona, Spain;
†Department of Organismic & Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, USA

Abstract

We studied sequence variation in the nuclear ribosomal internal transcribed spacers (ITS-1 and ITS-2) in 111 individuals from 11 populations/localities of the sponge *Crambe crambe* across the core species range in the western Mediterranean Sea and Atlantic Ocean. We report the first confirmed instance of intragenomic variation in sponges. Phylogeographical, nested clade and population genetic analyses were used to elucidate the species' evolutionary history. The study revealed highly structured populations affected by restricted gene flow and isolation-by-distance. A contiguous range expansion in the whole distribution area of the sponge was inferred. Phylogenetic analyses indicate a recent origin of most sequence types that could be explained by a recent origin of the species or a by recent bottleneck event in the studied area. A recent expansion of the distribution range to the Macaronesian region from the Mediterranean Sea was also detected, suggesting that *C. crambe* was recently introduced from the Mediterranean Sea to the Atlantic Ocean via human-mediated transport, and that the pattern observed is not the result of a natural biogeographical relationship between these zones.

Keywords: Atlanto-Mediterranean distribution, internal transcribed spacers, invasion, nested clade analysis, phylogeography, Porifera

Received 2 July 2003; revision received 3 September 2003; accepted 24 September 2003

Introduction

Understanding the distribution of alleles throughout a species' range is fundamental to molecular ecologists, allowing inferences about how history influenced the spatial distribution of these particular genes. Phylogeography is seen as the bridge between population genetics and phylogenetic systematics (Avise *et al*. 1987; Avise 2000). Phylogenetic methods can be used to infer haplotype trees and estimate patterns of relatedness among haplotypes focusing on the historical relationships of gene lineages. By comparing the phylogenetic tree to the geographical structure of the data, one can infer historical patterns of population subdivision and understand the current distribution of the studied species.

Correspondence: Sandra Duran. Fax: +34 93 403 5740, E-mail: sandra@porthos.bio.ub.es

Population genetic parameters and analyses of population structure or demographic history (reviewed in Emerson *et al*. 2001) can provide information about processes driving observed patterns of genetic variation. Analyses partitioning molecular variation among populations and groups of populations are useful to examine patterns of geographical structure. The challenge is to determine what method is best to answer the questions at hand. The use of multiple approaches including phylogenetic inference, nested clade analysis and genetic diversity measures seems to be the most appropriate way for elucidating not only geographical structure, but also the evolutionary history that produced such structure (i.e. Bernatchez 2001; Tarjuelo *et al*. 2001; Althoff & Pellmyr 2002).

Sponges constitute a group of marine invertebrates whose larvae disperse over short distances (Borojevic 1970; for an exception see Vacelet 1999). They are one of the dominating benthic groups in terms of biomass and species

diversity, yet studies on structure and gene flow among conspecific populations are scarce. Until now, allozymes have been the most commonly applied markers in those studies (reviewed in Solé-Cava & Boury-Esnault 1999). Molecular data have shown the existence of cryptic species that had gone undetected by morphologists (Borchiellini *et al*. 2000), with the risk of potentially confounding studies of genetic connectedness among populations. A few recent phylogeographical studies have used sequence data from the nuclear ribosomal internal transcribed spacers ITS-1 and ITS-2 (Lopez *et al*. 2002; Wörheide *et al*. 2002b; van Oppen *et al*. 2002) demonstrating the utility of ITS in resolving phylogeographical relationships at large spatial scales in sponges.

The common encrusting sponge *Crambe crambe* (Schmidt 1862) is widely distributed along the sublittoral of the western Mediterranean Sea (Boury-Esnault 1971; Pulitzer-Finali 1983; Uriz *et al*. 1992). It has recently been found in the Canary (Maldonado & Uriz 1996) and Madeira (P. Wirtz, pers. commun. 2002) archipelagos in the eastern Atlantic Ocean (these archipelagos are part of the so-called Macaronesian islands). It was also reported once in the Adriatic Sea (Schmidt 1862), and along the eastern Mediterranean coast of Egypt and Turkey (Burton 1936; Saritas 1972). Its low abundance in the central Mediterranean (i.e. Sicily; E. Ballesteros, pers. commun. 2002) and the scarcity of reports from the eastern Mediterranean suggest that the species is less abundant in those regions than in the western Mediterranean.

C. crambe is one of the best-known sponges from biological and ecological viewpoints (Becerro *et al*. 1997; Turon *et al*. 1998; Uriz *et al*. 1998) and is a medically important organism, producing numerous bioactive metabolites that are interesting from a pharmacological standpoint (Jares-Erijman *et al*. 1991; Berlinck *et al*. 1992). In addition, *C. crambe* is virtually free of microsymbionts (Becerro 1994; Galera *et al*. 2000), minimizing the effect of exogenous DNA in genetic studies (Lopez *et al*. 2002).

Field and laboratory studies on the swimming behaviour and dispersal abilities of its larvae (Uriz *et al*. 1998) indicate that the pelagic phase before settlement is short (≈ 48–72 h). This suggests a small dispersal potential between areas separated by open sea without a continuum of rocky littoral habitats. A study based on the DNA sequence data of the mitochondrial gene cytochrome *c* oxidase subunit I showed homogeneity across sponge populations (Duran *et al*. 2003) with scant variability and no geographical resolution, whereas a microsatellite study of one Atlantic and one Mediterranean population showed significant differentiation between them (Duran *et al*. 2002). In this study, we expanded our genetic sampling to include nuclear markers of the region comprising the ribosomal ITS-1, 5.8*S* and ITS-2 for the same individuals analysed in our previous studies, with the addition of individuals from other populations. In all, we included populations covering most of the species range, allowing us to investigate the role of evolutionary and ecological processes, such as restricted gene flow and population history in shaping the distribution of alleles from this locus throughout the species range.

In this article, we provide clear evidence that *C. crambe* has experienced a recent demographic expansion of its distribution area invading the Canary and Madeira archipelagos in the Atlantic Ocean, a colonization most probably mediated by human-related activities.

Materials and methods

Study area and sample collections

Individuals of *Crambe crambe* were sampled from nine populations from the western Mediterranean and two from the Macaronesian archipelagos in the Atlantic Ocean (Fig. 1, Table 1). Asexual fissiparous reproduction is known to occur in this sponge but both fission rates and

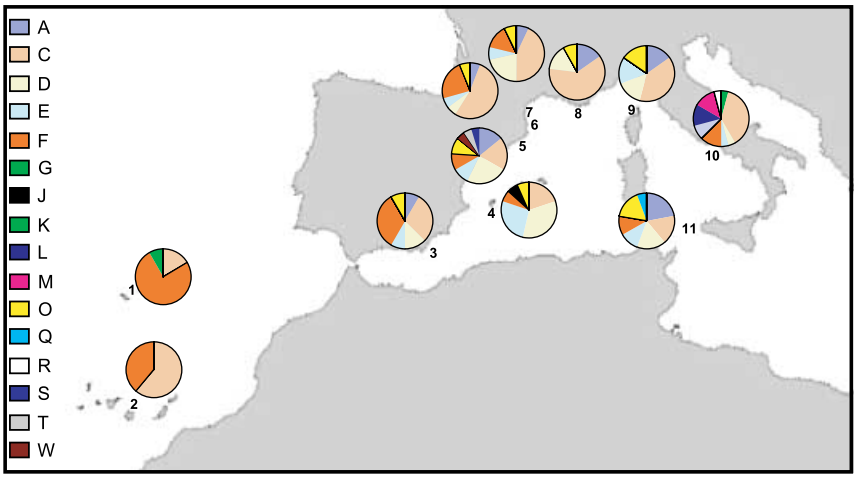

Fig. 1 Map showing the localities sampled (see Table 1 for details) with the geographical distribution of the sequence types and sequence type frequencies per population.

Table 1 Population code, sample size (*N*), number of sequence types (*Ns*), nucleotide (π) and sequence type (h) diversity, and uncorrected p-distance (p-dist) for each population studied. (SD) Standard deviation

Population	Code	*N*	*Ns*	π (SD)	h (SD)	p-dist (SD)
Madeira	1	10	3	0.000726 (0.000764)	0.4394 (0.1581)	0.469697 (0.438800)
Canaries	2	12	2	0.000778 (0.000774)	0.5033 (0.0639)	0.503268 (0.447965)
Cabo de Gata	3	11	6	0.002089 (0.001493)	0.8007 (0.0497)	1.351449 (0.866471)
Balearic Is.	4	10	6	0.002473 (0.001740)	0.8190 (0.0636)	1.600000 (1.004233)
Tossa de Mar	5	10	9	0.002664 (0.001805)	0.8952 (0.0376)	1.723809 (1.045942)
Cap de Creus	6	11	6	0.001682 (0.001300)	0.6912 (0.1025)	1.088235 (0.751357)
Banyuls	7	10	6	0.002310 (0.001662)	0.7912 (0.0894)	1.494506 (0.957687)
Marseille	8	10	4	0.001823 (0.001404)	0.6154 (0.1358)	1.179487 (0.808195)
Corsica	9	8	5	0.002219 (0.001622)	0.8205 (0.0769)	1.435897 (0.933791)
Naples	10	12	9	0.002251 (0.001578)	0.8333 (0.0600)	1.456522 (0.915990)
Sicily	11	10	7	0.002364 (0.001662)	0.8889 (0.0361)	1.529412 (0.961385)

individual growth rates have been found to be very low (Turon *et al*. 1998). Although the real size of clones is not known we sampled individuals at least 5 m apart to minimize the chance of sampling the same clone. We tried to collect a minimum sample of 10 individuals per population to increase the probability of finding low-frequency variants. Sponge tissue was collected by SCUBA, and kept in absolute ethanol at –20 °C until processed.

DNA extraction

Fragments from individual sponges were meticulously cleaned of exogenous tissues with the aid of sterile forceps under a stereo-microscope to avoid contaminating the DNA extractions. Total genomic DNA was extracted using the DNeasy® Tissue Kit (QIAGEN).

Polymerase chain reaction amplification and sequencing

The full ITS region, including the 5.8*S* ribosomal (r)RNA gene, was amplified using the primers 9F: 5′-GTAGGTGAACCTGCGGAAGG-3′ (Carranza 1997) and 28*S*Rev: 5′-GTTAGTTTCTTTTCCTCCGCTT-3′ (Lobo Hajdu, pers. commun. 2002). Amplifications were carried out in a 50-μL volume reaction, with 1.25 units of Ampli*Taq*® DNA polymerase (Perkin–Elmer), 200 μM of dNTPs and 1 μM of each primer. The polymerase chain reaction (PCR) consisted of an initial denaturing step at 94 °C for 60 s, 35 amplification cycles (94 °C for 15 s, 45 °C for 15 s, 72 °C for 15 s), and a final step at 72 °C for 6 min. Amplifications were carried out in a GeneAmp® PCR System 9700 (Perkin–Elmer).

PCR amplified samples were purified using the GENECLEAN® III kit (BIO 101 Inc.). Cycle-sequencing with AmpliTaq® DNA polymerase, FS (Perkin–Elmer) using dye-labelled terminators (ABI PRISM™ BigDye™ v3.0 Terminator Cycle Sequencing Ready Reaction Kit) was performed in a GeneAmp® PCR System 9700 (Perkin–Elmer). The sequencing reaction was carried out in a 10-μL volume reaction: 2 μL of Terminator Ready Reaction Mix, 2 μL of HalfTerm, 10–30 ng/mL of PCR product, 5 pmol of primer and dH_2O–10 μL. The cycle-sequencing programme consisted of an initial step at 94 °C for 3 min, 25 sequencing cycles (94 °C for 10 s, 50 °C for 5 s, 60 °C for 4 min) and a rapid thermal ramp to 4 °C. The BigDye-labelled PCR products were cleaned with AGTC® Gel Filtration Cartridges (Edge BioSystems) and directly sequenced using an automated ABI PRISM® 3100 Genetic Analyser.

DNA editing

Chromatograms obtained from the automated sequencer were read and contigs assembled using the sequence editing software SEQUENCHER™ 4.0. Complete sequences were then edited in GDE (Smith *et al*. 1994), and trivial alignments (no indel events needed to be postulated) generated.

Cloning

The ITS-1 and ITS-2 regions form part of the ribosomal nuclear array, and therefore several copies of each exist per nuclear genome (Long & Dawid 1980). The multiple copies of this cluster appear to be nearly identical within a given organism due to the process of concerted evolution (Hillis & Dixon 1991). Exceptions to this rule exist among metazoans (see Carranza *et al*. 1996 for a review), but until now Porifera seemed to feature intragenomic homogeneity for these genes (Wörheide *et al*. 2002b). Our DNA amplifications resulted always in a single discrete band, but a few intraindividual polymorphisms were detected by direct sequencing of the amplified products. These polymorphisms were detected by the presence of two base calls of similar intensity for certain positions, although no length variation was detected.

Given the presence of polymorphisms, PCR products from individuals with more than one polymorphic site (16 individuals in total) were cloned into pCR®2.1-TOPO® (Invitrogen™) following the manufacturer's protocol. Positive clones were grown overnight in Luria–Bertani media; minipreps were prepared with FastPlasmid Mini (Eppendorf) and were sequenced with M13 primers. A total of 125 clones were sequenced.

In order to avoid the term 'haplotype', which represents an haploid component of a given sequence, we use the term 'sequence type' to refer to every distinct type of ITS detected in the individuals studied, as proposed by Wörheide *et al*. (2002b).

Phylogenetic analysis

We estimated the maximum likelihood (ML) model that best fits the data under the hierarchical likelihood ratio test (hLRT) criterion as implemented in MODELTEST v3.06 (Posada & Crandall 1998). Once the model was selected, the number of sequences was reduced to represent unique sequence types. Using the model estimated under the hLRT, we performed a heuristic search consisting of 1000 random addition replicates (RAS) followed by tree bisection and reconnection (TBR) branch swapping in PAUP* (Swofford 1998). Nodal support was estimated using the bootstrap approach (Felsenstein 1985) with 1000 replicates of simple addition and TBR.

Network estimation and nested analysis

The null hypotheses of no genetic differentiation among populations of the whole area studied, and between populations within the Mediterranean Sea and the Atlantic Ocean, respectively, were tested by permutation tests (10 000 replicates) on χ^2 contingency tables (Hudson *et al*. 1992) using CHIPERM v1.2 (D. Posada, unpublished software available at http://inbio.byu.edu/Faculty/kac/crandall_lab/programs.htm). A sequence type cladogram was estimated with TCS v1.12 (Clement *et al*. 2000) using the statistical parsimony procedure (Templeton *et al*. 1992; Crandall *et al*. 1994). This method estimates the unrooted tree and provides a 95% plausible set for all sequence type linkages within the unrooted tree. The resulting network was then used to construct the nested clade design following the inference procedure given in Templeton *et al*. (1987; see also Templeton & Sing 1993 and Templeton *et al*. 1995). Once the nested design was determined, an exact permutation contingency analysis of categorical variation was implemented (i) for each step level, (ii) for the associations between clades, and (iii) for geographical locations. The contingency test was performed using the software GEODIS v2.0 (Posada *et al*. 2000) on clades with more than one sequence type and more than one sample location, following the algorithm given by Templeton & Sing (1993). This software detects significant genetic and geographical associations within the sequence type cladogram and incorporates the geographical distances as clade distance (Dc) and nested clade distance (Dn). Dc measures how geographically widespread are the individuals that bear sequence types from a specific given clade. Dn measures how far the individuals bearing sequence types from a given clade are from all other individuals that bear sequence types included in the immediate higher step clade. The statistical significance of these two measures was estimated by recalculating them in 10 000 random permutations. This randomization procedure allowed testing of the null hypotheses of no geographical association within the nested clade design (Templeton *et al*. 1995). Phylogeographical interpretations of significant values for Dc and Dn were inferred using the inference key available at http://bioag.byu.edu/zoology/crandall_lab/geodis.htm

Solving cladogram ambiguities

To solve any ambiguities before constructing the nesting, we used the empirical predictions derived from coalescent theory (Crandall & Templeton 1993; Templeton & Sing 1993; Crandall *et al*. 1994; Posada & Crandall 2001). These predictions can be summarized in three criteria (Pfenninger & Posada 2002):

1. *Frequency criterion*, as high-frequency sequence types might have been present in the population for a long time, they had more chances of originating new sequence types than did younger sequence types; so low-frequency sequence types are more likely to be connected to sequence types with high frequency.
2. *Topological criterion*, sequence types are more likely to be connected to interior sequence types than to tip sequence types.
3. *Geographic criterion*, sequence types are more likely to be connected to sequence types from the same population or region than to sequence types occurring in distant populations.

Population genetics parameters and analyses of population structure

The population genetics analyses were performed using ARLEQUIN v2.0 (Schneider *et al*. 2000). We calculated sequence type and nucleotide diversity for all populations. Sequence type frequencies per population were calculated and represented in frequency plots. We used analysis of molecular variance (AMOVA) to examine hierarchical population structure, performing 16 000 permutations to guarantee having less than 1% difference with the exact probability in 99% of cases (Guo & Thompson 1992). We used our a priori expectation of a genetic division between the Mediterranean and Atlantic populations.

Table 2 Genotyped individuals per population. Distribution of informative polymorphic sites in the sequences studied (nucleotide positions indicated on top). R = G/A, Y = T/C, W = A/T. Populations as in Table 1

Populations												ITS1		ITS2					
1	2	3	4	5	6	7	8	9	10	11	Genotype	2	81	384	440	482	492	516	569
—	—	2	1	2	1	1	2	2	—	1	I	T	Y	A	Y	C	T	G	T
1	6	7	—	2	3	2	—	—	2	2	II	A	Y	A	T	C	T	G	T
1	5	—	2	2	6	4	5	8	—	1	III	A	C	A	T	C	T	G	T
—	—	1	1	2	—	2	—	—	—	1	IV	T	T	A	C	C	T	G	T
—	—	—	3	—	—	—	1	—	—	—	V	T	T	A	Y	C	T	G	T
7	1	1	1	—	—	1	—	—	—	—	VI	A	T	A	T	C	T	G	T
1	—	—	—	—	—	—	—	—	—	—	VII	A	T	R	T	C	T	G	T
—	—	—	—	1	—	—	—	—	1	—	VIII	T	Y	A	Y	C	T	R	T
—	—	—	—	1	—	—	—	—	—	1	IX	T	Y	A	T	C	T	G	T
—	—	—	1	—	—	—	—	—	—	—	X	A	C	A	T	Y	T	G	T
—	—	—	—	—	—	—	—	—	1	—	XI	A	T	A	T	C	T	G	W
—	—	—	—	—	—	—	—	—	3	—	XII	A	C	A	T	C	T	G	W
—	—	—	—	—	—	—	—	—	3	—	XIII	A	C	A	T	C	T	R	T
—	—	—	—	—	—	—	—	—	1	—	XIV	A	Y	A	T	C	T	G	W
—	—	—	—	—	—	—	—	—	—	2	XV	T	T	A	C	C	T	G	T
—	—	—	—	—	—	—	—	—	—	1	XVI	T	C	A	T	C	T	G	T
—	—	—	—	—	—	—	—	—	—	1	XVII	T	Y	A	T	C	Y	G	T
10	12	11	9	10	10	10	8	10	11	10	Total ind								

Table 3 Sequence types (SeqT) and their frequencies per population

	ITS1		ITS2						Populations										
SeqT	2	81	384	440	482	492	516	569	1	2	3	4	5	6	7	8	9	10	11
A	T	C	A	T	C	T	G	T			0.083		0.143	0.059	0.071	0.154	0.154	0.042	0.222
C	A	C	A	T	C	T	G	T	0.167	0.611	0.292	0.200	0.190	0.529	0.429	0.615	0.385	0.375	0.167
D	T	T	A	C	C	T	G	T			0.125	0.333	0.238	0.059	0.214	0.154	0.154	0.042	0.167
E	T	T	A	T	C	T	G	T			0.083	0.267	0.095	0.059	0.071		0.154	0.042	0.111
F	A	T	A	T	C	T	G	T	0.750	0.389	0.333	0.067	0.095	0.235	0.143			0.125	0.111
G	A	T	G	T	C	T	G	T	0.083										
J	A	C	A	T	T	T	G	T				0.067							
K	A	T	A	T	C	T	G	A										0.083	
L	A	C	A	T	C	T	G	A										0.125	
M	A	C	A	T	C	T	A	T										0.125	
O	T	C	A	C	C	T	G	T			0.083	0.067	0.095	0.059	0.071	0.077	0.154		0.167
Q	T	T	A	T	C	C	G	T											0.056
R	A	T	A	T	C	T	G	T										0.042	
S	T	T	A	C	C	T	A	T					0.048						
T	T	C	A	C	C	T	A	T					0.048						
W	T	T	A	T	C	T	A	T					0.048						

Results

In total, 647 bp comprising the complete ITS-1, 5.8*S* rRNA and ITS-2 regions were sequenced for 111 individuals (plus 125 clones). The 17 genotypes found are shown in Table 2. We detected 16 different rRNA sequence types (Accession nos AY319369–AY319411) defined by eight variable sites, including two sites in the ITS-1 and six in the ITS-2 (Tables 2 and 3). The 5.8*S* region appeared invariable in all sequenced individuals. The length of the individual spacers was 219 bp for ITS-1, and 275 for ITS-2. Uncorrected p-distances ranged from 0.46 to 1.7% (average 1.2%, see Table 1). Total nucleotide composition was A = 0.12, C = 0.30, G = 0.30, T = 0.28 for ITS-1; A = 0.24, C = 0.28, G = 0.26, T = 0.22 for 5.8S; and A = 0.11, C = 0.30, G = 0.31, T = 0.28 for ITS-2. The greatest number of differences among sequence types was six nucleotides.

Table 4 Individuals cloned, populations (as in Table 1), genotype, number of clones sequenced, sequence types found and uncorrected p-distances

Individual	Population	Genotype	N. clones seq.	Sequence types	p-dist (SD)
GC1	3	I	8	D, A, E, O	0.964286 (0.729210)
GD2	3	I	7	D, A, E, O	1.142857 (0.835361)
F7	4	I	10	E, D, O	0.733333 (0.592372)
CR5	5	I	5	A, D, O	1.000000 (0.796333)
CR8	5	I	9	D, A	0.444444 (0.433519)
CB1	5	VIII	11	S, T, D, W, E, O, A	1.381818 (0.918394)
CC2	6	I	9	D, A, E, O	1.110000 (0.796899)
BY10	7	I	4	D, A, E, O	1.333333 (1.024942)
M6	8	I	6	D,A,O	1.133333 (0.847218)
M12	8	I	6	D, A	1.200000 (0.883176)
CO4	9	I	6	D, A, O	1.1333333 (0.847218)
CO10	9	I	9	A, E, O	0.611111 (0.530218)
NA6	10	VIII	10	D, A, E	1.000000 (0.732584)
NA9	10	XIV	9	K, R, C	1.388889 (0.938004)
S2	11	I	6	D, A, E, O	1.200000 (0.883176)
S6	11	XVII	10	Q, A, E	1.000000 (0.732584)

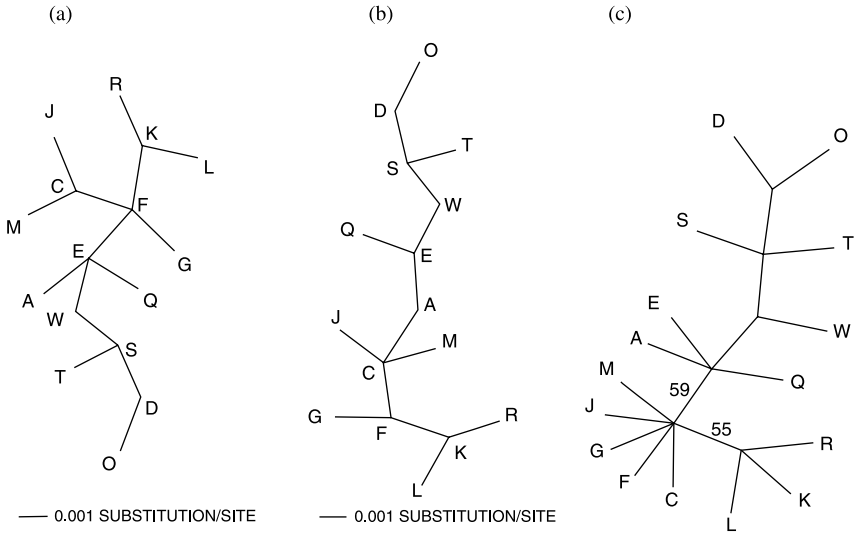

Fig. 2 Maximum likelihood (ML) estimates of sequence type phylogeny. (a, b) The two equally likely trees generated under model of sequence evolution HKY85 + θ + Γ (–ln L = 968.39652) unrooted and with branch length information. (c) Strict consensus of the two ML trees. The tree is shown unrooted, and no branch length information is provided. Bootstrap values above 50% are shown.

Intra-individual diversity

This is the first time that intragenomic variation has been confirmed for sponges; such variation is attributable to lack of homogenization in the variable sites of some individuals, a fact that has not been reported to date in other ITS studies in this group. In these individuals, the number of sequence types found per individual ranged from two to seven with uncorrected p-distances of 0.44–1.3% (average 1%, see Table 4). The number of sequence types per sampling site ranged from two to nine, with a tendency to decrease towards the western and northern ranges of the distribution (Fig. 1, Tables 1 and 3).

Patterns of phylogenetic relatedness of the sequence types

For the phylogenetic estimation of the sequence data, under the hLRT criterion, the best fit model of nucleotide substitution corresponds to a Hasegawa–Kishino–Yano 1985 (HYK85) model of nucleotide substitution with a proportion of invariable sites (θ) and a gamma distribution (Γ) of among-site rate variation (HKY85 + θ + Γ hereafter) (Hasegawa *et al*. 1985) (Ti/Tv ratio = 6.8435; θ = 0.9715; α = 0.5804). For the ML analysis under the best-fit model, the search strategy yielded optimal trees at –ln L = 968.39652 in 79.5% of replicates, and resulted on two islands of one tree each. These two trees (Fig. 2a,b) differ in

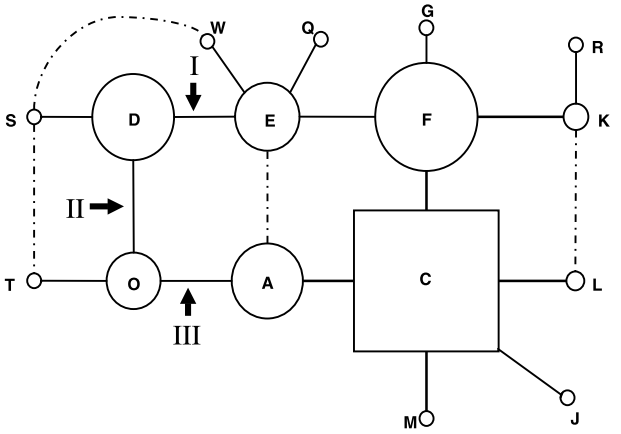

Fig. 3 Statistical-parsimony cladogram and proposed loop solutions. Lines in the statistical-parsimony cladogram represent one mutational step between sequence types; dashed lines represent connections removed to resolve loops. Arrows indicate the three favoured possibilities for solving the interior loop (I, II and III). The area of the circles is proportional to the frequency of sequence types; the square denotes the inferred ancestral sequence type.

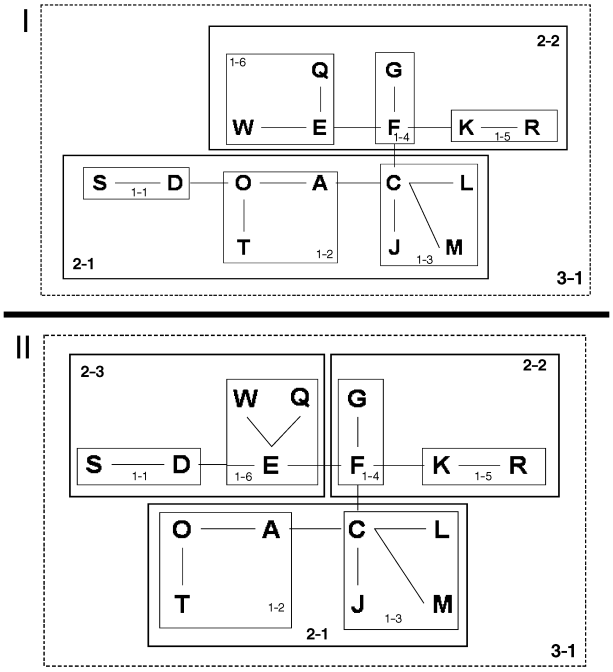

Fig. 4 Two options of the clade nesting of the parsimony network, depending on the resolution of the interior loop. Sequence type names as in Table 3. Lines represent one mutational step connecting two sequence types. Boxes show the sequence types nested together into one-, two- and three-step clades. All sequence types are nested into clade 3-1.

the branching pattern of several sequence types. The strict consensus of these two trees, with the bootstrap values plotted on its branches is shown in Fig. 2(c). Because of the low level of variation among sequences, bootstrap values were low, only two nodes being supported slightly above the 50% threshold.

The statistical parsimony analysis (Fig. 3) revealed a network with a central position for the most frequent sequence type C. Five of the 16 sequence types are one mutational step from sequence type C. The most parsimonious cladogram also revealed five closed loops among sequence types, three exterior and two interior. To resolve the ambiguous loops in the cladogram, we suggest breaking the connections indicated by dotted lines in Fig. 3. In the loop formed by connecting sequence types C–L–K–F, the connection between sequence types L and K seems less likely according to the topology and frequency criteria, even though they are found in the same population. The ambiguous loop connecting sequence types S–T–O–D can be solved in a similar way, and the same two criteria easily solve the loop connecting D–E–S–W. For the remaining two interior loops, the A–C–E–F can be solved by breaking the connection between E and A appealing to the frequency criterion because sequence type A is more likely linked to sequence type C than to sequence type E. But the remaining interior loop (D–E–F–C–A–O) could be broken at three different places (I, II, III; Fig. 3) with similar probabilities according to the different criteria. So we decided to explore these three alternative ways. The three resulting cladograms led to different nesting designs and, thus, potentially to different inferences about population history. We explored the inferences drawn from all three solutions of the loop, and despite the differences in the resulting nesting designs, the interpretation of the population history was essentially identical for options I and III and slightly different for option II. Therefore, the two nesting solutions corresponding to the removal of connection I and II are presented (Fig. 4).

Population history

The results of the nested clade analysis (NCA) for options I and II (Fig. 4) can be followed in Tables 5 and 6. The two options analysed yielded significant associations between sequence type clades and geographical distribution at all levels (Table 6). In both options the inference at the one-step clade level was the same. Within clade 1-3 a restricted gene flow with isolation-by-distance was found. For the two-step clades different inferences were found between the two alternative nestings. In option I, restricted gene flow with isolation-by-distance was found for both clades 2-1 and 2-2, and the overall inference was a continuous range expansion of the species. For option II, restricted gene flow with isolation-by-distance was also inferred for clade 2-2, whereas for clade 2-1, restricted gene flow with some long-distance colonization was detected. The overall inference for option II was restricted gene flow with isolation-by-distance.

Table 5 Above, nested contingency analysis for the sequence type–geography association for the nested design given in Fig. 4-(I). Below, the same for the nested design given in Fig. 4-(II). Columns show increasing nesting levels, from sequence types to 3-step clades. Dc and Dn distances are given for each clade. Detection of significant differences between the observed and expected distances under a situation of random geographical distribution of sequence types is indicated with a superscript capital S (significantly small) or L (significantly large). I-T are the interior-tip distances. Tip clades are shown in bold

Sequence types	Dc	Dn	One-step clades	Dc	Dn	Two-step clades	Dc	Dn	Three-step clades
Option I									
S	0	376							
D	588	579	**1-1**	**575^{S}**	**763^{S}**				
I-T	588	202							
O	676	630							
T	0	473	1-2	655^{S}	801^{S}				
A	674	637				**2-1**	**975^{S}**	**1045^{S}**	
I-T	675	161							
C	1213	1208							
J	0	883							
L	0^{S}	1139	1-3	1213^{L}	1080^{L}				
M	0^{S}	1139	I-T	476^{L}	236^{L}				
I-T	1213^{L}	105							3-1
F	1415	1410							
G	0	1254	1-4	1415	1415^{L}				
I-T	1415	156							
K			**1-5**	**0^{S}**	**1509**				
R						2-2	1358^{L}	1242^{L}	
E	659	653							
W	0	479	1-6	660^{S}	1132^{S}				
Q	0	777	I-T	854^{L}	226^{L}				
I-T	659	25							
Option II									
O	676	630							
T	0	473	1-2	655^{S}	858^{S}				
A	674	637							
I-T	675	161				**2-1**	**1066**	**1080**	
C	1213	1208							
J	0	883	1-3	1213^{L}	1128^{L}				
L	0^{S}	1139	I-T	557^{L}	270^{L}				
M	0^{S}	1139							
I-T	1213^{L}	105							
F	1415	1410							
G	0	1254	1-4	1415^{S}	1449^{S}				
I-T	1415	156				2-2	1483^{L}	1349^{L}	3-1
K			**1-5**	**0^{S}**	**1798^{L}**				
R			I-T	1415^{L}	48^{S}				
E	659	653							
W	0	479	1-6	660	600				
Q	0	777							
I-T	659	25				**2-3**	**604^{S}**	**878^{S}**	
						I-T	552^{L}	358^{L}	
S	0	376	1-1	575	571				
D	588	579	I-T	–84	–28				
I-T	588	202							

Patterns of recent population structure

We found large sequence type diversity (mean 0.736191 ± 0.154313) and low values of nucleotide diversity (mean 0.001944 ± 0.00065) per population (Table 1). A significant genetic differentiation between localities, based on a χ^2 test (Hudson *et al*. 1992) was found between Mediterranean and Atlantic populations ($P < 0.005$), among Mediterranean

Table 6 χ^2 test of geographical association of clades and biological inference from the NCA analysis of the two nested clade options favoured (I and II). Probability *P* is the probability of obtaining a χ^2-statistic larger or equal to the observed statistic based on 10 000 resamples. Abbreviations for the inferences are: CRE, contiguous range expansion; IbD, isolation-by-distance; LDC, long-distance colonization; RGF, restricted gene flow

Clades nested with	χ^2-statistic	*P*	Chain of inference	Inference
Option I				
1-1	3.4667	1	No significant clade distances	—
1-2	6.6037	0.9924	No significant clade distances	—
1-3	43.2955	0.0972	1-2-3-4 NO	RGF with IbD
1-4	2.9763	1	No significant clade distances	—
1-6	9.8222	0.7770	No significant clade distances	—
2-1	40.9016	0.0021	1-2-3-4 NO	RGF with IbD-
2-2	47.4882	0.0014	1-2-3-4 NO	RGF with IbD
Entire Cladogram	25.6757	0.0037	1-2-11 YES-12 NO	CRE
Option II				
1-1	3.4667	1	No significant clade distances	—
1-2	6.6037	0.9925	No significant clade distances	—
1-3	43.2955	0.0965	1-2-3-4 NO	RGF with IbD
1-4	2.9763	1	No significant clade distances	—
1-6	9.8222	0.7812	No significant clade distances	—
2-1	24.4714	0.0039	1-2-3-4 NO	RGF with IbD
2-2	19.3846	0.0412	1-2-3-5-6-7 YES	RGF with LDC
2-3	2.5386	0.9681	No significant clade distances	—
Entire Cladogram	67.4164	0.0000	1-2-3-4 NO	RGF with IbD

Table 7 Analysis of molecular variance (AMOVA) among populations of *Crambe crambe*. Groups correspond to Mediterranean Sea and Atlantic Ocean. The significance tests were based on 16 000 permutations

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among groups	1	2.744	0.04235 Va	9.78*
Among populations within groups	9	5.596	0.01426 Vb	3.29*
Within populations	178	67.015	0.37649 Vc	86.93*
Total		188	75.354	0.43311

*Significant values at $P < 0.05$.

populations ($P < 0.05$), and among Atlantic populations ($P < 0.05$). Pooling all the populations together the significance was also high ($P < 0.0001$).

Incorporating both sequence divergence and sequence type frequencies per populations, the AMOVA detected significant structure between Mediterranean and Atlantic groups, among populations within groups, and within populations, the latter with the highest percentage of the differentiation found (Table 7).

Discussion

The level of intraspecific variation (1.2%) detected in the ITSs of *Crambe crambe* is six times higher than that found in the mitochondrial gene cytochrome *c* oxidase subunit I (COI), as determined by a study at a similar geographical scale with the same sponge individuals (Duran *et al*. 2003). Other sponge COI data included in Shearer *et al*. (2002) also reveal low variability in this gene. The amount of variation in the ITS region of *C. crambe* is in the range of that found in the sponge *Leucetta chagonensis* from the Pacific Ocean (0.1–1.6%, Wörheide *et al*. 2002b), and it clearly differs from that of *Astrosclera willeyana* from the Indo-Pacific, which shows length variation in the ITS region (Wörheide *et al*. 2002a). The GC content found in *C. crambe* (59%) is slightly higher than that from other sponge species (Wörheide *et al*. 2002a, 2002b). When the intraspecific variation of ITS in this sponge is compared with the range of variation found in other marine species it is found to be relatively low. Examples can be found in the algae (1.7%, Connell 2000), corallimorpharians (2.5%, Chen *et al*. 1996) and scleractinian corals (3–29%, Odorico & Miller 1997; Diekmann *et al*. 2001; Rodriguez-Lanetty & Hoegh-Guldberg 2002).

Owing to the multicopy nature of the ribosomal array, we have found intra-individual polymorphism in the ITS regions, which had not been reported previously in sponges. Although cloning could generate sequencing errors, the fact that we have found the same alleles many times in the individuals studied suggests that the alleles found are real. Furthermore, all the variation found after sequencing the clones was already accounted for in the PCR-based sequencing of the same individuals, and new polymorphic sites were not detected from the clones. Thus,

the simplest explanation of the observed sequence variation implies that concerted evolution might not have yet homogenized the ITS copies.

Inspection of sequence type frequencies and their distribution among populations (Fig. 1) suggests a population differentiation that can be explained by different phylogeographical patterns. First, following coalescent theory predictions (Crandall 1996), sequence type C is likely to be the most ancestral type because it is present in all populations studied, is the most frequent and has the highest number of mutational connections. Second, the Atlantic populations have only three sequence types, one of them unique (G), suggesting a potential invasion from the Mediterranean Sea with a founder effect event. To our knowledge, no other Atlantic populations of *C. crambe* have been reported, and the limited Atlantic distribution of the species may also be partially responsible for the pattern found in this phylogeographical study. Finally, the distribution of sequence types in the Mediterranean suggests a subdivision with at least three different zones, the first in the western part of the range (Cabo de Gata, Balearic Islands, Tossa de Mar, Cap de Creus and Banyuls sur Mer), a second in the central area (Marseille and Corsica) and a third in the eastern part of the range (Naples), this later population being the most diverse genetically, with nine sequence types present. Although Sicily is also rich in sequence types and geographically located in the central part of the range, its sequence type composition and frequency fit better with the western group, as it has more sequence types in common with this group than with Naples. All these patterns might have been caused by large or medium-scale hydrodynamic processes that strongly influence larval dispersal and constitute potential barriers to gene flow among populations. Water circulation within the different basins in the western Mediterranean (Hopkins 1985) could limit larval transport and favour genetic differentiation among populations in agreement with the patterns found. A similar result suggesting a strong influence of Mediterranean currents in larval dispersal has been found in a phylogeographical study of the polychaete *Sabella spallanzanii* (Patti & Gambi 2001).

Inferring intraspecific sequence evolution

The ML estimate of sequence type phylogeny and the statistical parsimony (SP) cladograms exhibited somewhat compatible topologies. However, ML failed to resolve most of the relationships among sequences and among higher level clades because of the low sequence divergence observed. This low degree of nucleotide variation between sequence types results in a poorly resolved consensus tree with low support for almost all branches. Intraspecific phylogenies generally resolve lineages that have been separated for long periods, where accumulation of genetic divergence translates into a signal that may correspond well with geographical separation. In our case, we find little phylogenetic signal even among geographically structured sequence types. This can be taken as suggestive of a recent origin for most sequence types. The SP cladogram, although it includes five loops that might be due to recombination or homoplasy events, is better resolved. This result highlights how network approaches may be more effective than phylogenetic approaches at detecting intraspecific evolution, as previously suggested by Posada & Crandall (2001).

Three hypotheses might explain the shallow divergence among sequence types. First, *C. crambe* might be a relatively young species that has recently spread across its range and, accordingly, it has had no time to generate large ITS sequence divergence. An alternative hypothesis regarding the high levels of sequence type diversity relative to nucleotide diversity is that *C. crambe* is an old species that has experienced changes in its population demography, i.e. a strong recent bottleneck that has reduced its former genetic diversity followed by a new expansion and accumulation of new mutations. Lastly, these results could be the consequence of low mutation rates at the locus studied, or even at the genomic level.

Inferring population history

Working with data containing intra-individual variation for phylogeographical purposes adds analytical difficulties to this study. The fact is that there are no examples in the literature supporting or rejecting specific methods to analyse this kind of data. We have chosen the NCA because it works with allele frequency data per population, and we do believe that our sampling design and subsequent cloning of polymorphic individuals are giving us a proper idea of the allele variants found in each population studied. Nevertheless we realize that our results have to be handled with care and more specific analytical tools will have to be developed in this field. Working with other molecular markers (i.e. mtDNA or microsatellites) in the same sample set seems a good option to test the results found.

The population history of *C. crambe*, as inferred by the NCA, involved historical events as well as recurrent gene flow. For the one step-level clades, both nesting designs gave the same results. Restricted gene flow with isolation-by-distance was inferred for clade 1-3, which includes the most frequent and widespread sequence type (sequence type C) and the unique sequence types M and L from Naples and J from the Balearic Islands. Because restricted gene flow implies only limited movement by individuals during any given generation, it takes time for a newly arisen sequence type to spread geographically. Keeping in mind that the ancestral sequence type (probably C) is

expected to be frequent near its site of geographical origin and that most mutational derivates of the ancestral sequence type will also occur near the ancestral site of origin (Templeton 1998), it seems that the western Mediterranean is the region where the oldest populations of the sponge are found nowadays. This area has thus acted as the centre of radiation to other zones, specifically to the Macaronesian archipelagos.

At the two-step clade level, two different inferences were made depending on the nested clade design. For nesting option I, restricted gene flow with isolation-by-distance was inferred for all clades. Isolation-by-distance with some long-distance colonization was found in nesting option II for clade 2-2 (sequence types F, G and K, R) suggesting the potential recent invasion of the Atlantic Ocean, possibly linked to one or more sporadic long-dispersal events. It is also evident that for sequence type F there is a break in its distribution in Corsica and Marseille but it appears again in Naples. Atlantic populations have high frequencies of the two most frequent sequence types (C and F) suggesting a founder effect. Only one new sequence type (G) unique to Madeira, is found in the Atlantic, indicating that colonization may have been recent and that there has not been enough time to accumulate more changes at detectable frequencies. The observation that Atlantic populations show less intragenomic variation might also reflect recent founder event. The fact that the abundance of this sponge in the Atlantic populations studied is high (S. Duran pers. obs. 2003) suggests that the sponge has found conditions for spreading in this new area. Even if Atlantic waters are different from the Mediterranean waters in both physical and biological conditions it is known that the Macaronesian region (Canaries, Madeira and Azores archipelagos) has a strong Mediterranean component in its faunal composition (Wirtz & Martins 1993; Wirtz 1998), of which our results provide further evidence.

Crambe crambe is a sponge with high levels of bioactive metabolites; these substances help the sponge avoid both predation and competition and have powerful antimicrobial and antiviral properties (Becerro *et al*. 1994). In the Mediterranean, it is one of the most efficient sponges in terms of space competition and lacks known predators. These characteristics may confer on this sponge invasive capabilities and high potential for colonizing new areas where physical conditions are within its tolerated range.

Colonization of the Canaries and Madeira by larvae arriving via oceanic currents seems unlikely if we take into account that larvae of *C. crambe* remain in the water column for just a few hours or a couple of days at most (Uriz *et al*. 1998). The Canary Current, which runs southwards from the Iberian Peninsula, reaches a maximal velocity in the order of 30 cm/s (Batten *et al*. 2000; Zhou *et al*. 2000), which would imply a dispersal of only 50 km for a larva passively drifting for 2 days. This dispersal range is far too short to cover the ~1300 km that separate the Macaronesian archipelagos from the Mediterranean Sea. Besides, *C. crambe* has not apparently been able to enter the Atlantic Iberian waters further away than Cape San Vicente (M.J. Uriz, pers. commun. 2002) and has not been reported from the Atlantic shores of North Africa, neither is it present in the Azores (J. Xavier, pers. commun. 2002). This discontinuous distribution, together with the low genetic diversity of ITSs in the Atlantic populations, indicates that a human-mediated invasion is likely in this case (although rafting on natural debris cannot be discarded). The high differentiation between sequence type frequencies in the two Atlantic populations (Table 3 or Fig. 1) is indicative of low gene flow between them, and strongly suggests two independent colonizations of these archipelagos. Introduction of *C. crambe* to the Atlantic may have happened by the transport of larvae in ballast water, or via fouling on ship hulls or rafting in debris. Transport of marine invertebrates via ballast water has been documented for many species with a planktonic phase in their life cycle, in many marine habitats as well as trophic groups (Carlton & Geller 1993).

Regarding other organisms, there are studies reporting a strong genetic discontinuity between each side of the Strait of Gibraltar in fish and invertebrates (Borsa *et al*. 1997; Chikhi *et al*. 1997; Hawkins *et al*. 2000) but it is difficult to generalize about the patterns of colonization and their directions (to or from the Mediterranean), as different groups have different biogeographical histories and are influenced by different life-history traits, such as the dispersal potential. Other biogeograhic studies of sponge communities have reported western Mediterranean sponge assemblages originating from Atlantic assemblages (Maldonado & Uriz 1996; Carballo *et al*. 1997). Our results, however, suggest that this particular species has a colonization history from the Mediterranean to the Atlantic. Although this is consistent with other observations of a similarity between the Mediterranean and the Macaronesian faunas, the influence of Mediterranean fauna has never been evaluated from a historical perspective. In this respect, the Macaronesian archipelagos have undergone intense maritime trading with Spain and Portugal for historical reasons. The case of *C. crambe* may hence constitute the first evidence of recent introduction of Mediterranean fauna in the Macaronesia archipelagos due to human transport, and not due to a biogeographical relationship of both zones.

Despite the high sequence type diversity in some populations, the amount of sequence divergence remains low between all sequence types, so populations of *C. crambe* might be relatively young. If we take into account that the genus *Crambe* is represented by five species, three of them found in the Mediterranean Sea (Maldonado & Uriz 1996), it seems plausible that origin and radiation of the genus

occurred in the Mediterranean Sea after the Messinian crisis 5–6 Mya (Duggen *et al*. 2003) from an unknown phylogenetic lineage (Maldonado 1985) and that it required man-mediated dispersal to traverse the unique environmental barrier that poses the Strait of Gibraltar (see Gofas 1998).

The contiguous range expansion suggested by option I is consistent with the loss of sequence type variation to the north, and to the west of the geographical area of the distribution. Also, those sequence types found in the ancestral population(s) that were the source of the range expansion (C and F) became geographically widespread. Some of the sequence types found in the expanding populations (D in the Mediterranean) became more frequent than some of the older sequence types from which they have originated (E or O). The inference for option II of the NCA for the entire cladogram is isolation-by-distance with restricted gene flow. This option would be in agreement with the dispersal features of the species.

Even though the two inferences are slightly different, both options detect the isolation-by-distance model of dispersal and invasion of this sponge from the Mediterranean to the Atlantic and both results are strongly supported by a microsatellite study performed with the same sample set (Duran *et al*. 2004). Thus, even if our results contrast with those of previous studies that describe the colonization history of marine invertebrates species between the Mediterranean Sea and the Atlantic Ocean, we strongly believe that the colonization pattern suggested by our data could be more general in marine invertebrates than previously recognized.

Acknowledgements

We are indebted to David Posada for long theoretical discussions and for his advice on the use of his software. Marta Pascual made useful comments on the manuscript. We are grateful to M. Zabala, C. Palacín, S. López-Legentil, P. Wirtz, E. Cebrian, N. Boury-Esnault, G. Villani and E. Ballesteros for providing samples from Corsica, Cabo de Gata, Banyuls-sur-Mer, Madeira, Gran Canaria, Marseille, Naples and Sicily, respectively. This research was supported by a Predoctoral fellowship from the Government of Catalonia to SD. Research was funded by project REN2001-2312 of the Spanish Ministry of Science and Technology. Funding for laboratory work was in part provided by NASA Fundamental Biology Program to GG. We declare that all the experiments done comply with the current Spanish and USA laws.

References

Althoff DM, Pellmyr O (2002) Examining genetic structure in a bogus yucca moth: a sequential approach to phylogeography. *Evolution*, **56**, 1632–1643.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Batten ML, Martinez JR, Bryan DW, Buch EJ (2000) A modelling study of the coastal eastern boundary current system off Iberia and Morocco. *Journal of Geophysical Research*, **105**, 14173–14195.

Becerro MA (1994) Chemically mediated bioactivity of the encrusting sponge *Crambe crambe* and its ecological implications. PhD Thesis, Universitat de Barcelona.

Becerro MA, Lopez NI, Turon X, Uriz MJ (1994) Antimicrobial activity and surface bacterial film in marine sponges. *Journal of Experimental Marine Biology and Ecology*, **179**, 195–205.

Becerro MA, Uriz MJ, Turon X (1997) Chemically mediated interactions in benthic organisms: the chemical ecology of *Crambe crambe* (Porifera, Poecilosclerida). *Hydrobiologia*, **356**, 77–89.

Berlinck RGS, Braeckman JC, Daloze D *et al.* (1992) Crambines C1 and C2: two further ichthyotoxic guanidine alkaloids from the sponge *Crambe crambe*. *Journal of Natural Products*, **55**, 528–532.

Bernatchez L (2001) The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. *Evolution*, **55**, 351–379.

Borchiellini C, Chombard C, Lafay B, Boury-Esnault N (2000) Molecular systematics of sponges (Porifera). *Hydrobiologia*, **420**, 15–27.

Borojevic R (1970) Différentiation cellulaire dans l'embriogénèse et la morphogénèse chez les spongiaires. In: *The Biology of the Porifera* (ed. Fry WG), pp. 467–490. Academic Press, London.

Borsa P, Naciri M, Bahiri L *et al*. (1997) Infraspecific zoogeography of the Mediterranean: population genetic analysis on sixteen Atlanto-Mediterranean species (fishes and invertebrates). *Vie Milieu*, **47**, 295–305.

Boury-Esnault N (1971) Spongiaires de la zone rocheuse littorale de Banyuls-sur-Mer. I. Ecologie et répartition. *Vie Milieu*, **22**, 159–192.

Burton M (1936) *IX Sponges*. *The Fishery Ground: Near Alexandria*, p. 17. Ministry of Commerce and Industry, Egypt.

Carballo JL, Naranjo S, García-Gómez JC (1997) Where does the Mediterranean Sea begin? Zoogeographical affinities of the littoral sponges of the Strait of Gibraltar. *Journal of Biogeography*, **24**, 223–232.

Carlton JT, Geller JB (1993) Ecological roulette: the global transport of nonindigenous marine organisms. *Science*, **261**, 78–82.

Carranza S (1997) Taxonomia molecular mitjançant la seqüenciació del DNA ribosòmic 18*S*. Aplicació a l'origen i filogènia dels Platihelmints. PhD Thesis, Universitat de Barcelona.

Carranza S, Giribet G, Ribera C, Baguñà J, Riutort M (1996) Evidence that two types of 18*S* rDNA coexist in the genome of *Dugesia* (*Schmidtea*) *mediterranea* (Platyhelminthes, Turbellaria, Tricladida). *Molecular Biology and Evolution*, **13**, 824–832.

Chen CA, Willis BL, Miller DJ (1996) Systematic relationships between tropical Corallimorpharians (Cnidaria: Anthozoa: Corallimorpharia): utility of the 5.8*S* and internal transcribed spacers (ITS) regions of the rRNA transcription unit. *Bulletin of Marine Science*, **59**, 196–208.

Chikhi L, Agnese JF, Bonhomme F (1997) Strong differences of mitochondrial DNA between Mediterranean sea and Eastern Atlantic populations of *Sardinella aurita*. *Canadian Royal Academy of Sciences*, **III** (320), 289–297.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Connell LB (2000) Nuclear ITS region of the alga *Heterosigma akashiwo* (Chlorophyta: Raphidophyceae) is identical in isolates from Atlantic and Pacific basins. *Marine Biology*, **136**, 953–960.

Crandall KA (1996) Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type I sequences. *Molecular Biology and Evolution*, **13**, 115–131.

Crandall K, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics*, **134**, 959–969.

Crandall KA, Templeton AR, Sing CF (1994) Intraspecific cladogram estimation: problems and solutions. In: *Models in Phylogeny Reconstruction* (eds Scotland RW, Siebert DJ, Williams DM), pp. 273–297. Clarendon Press, Oxford.

Diekmann OE, Bak RPM, Stam WT, Olsen JL (2001) Molecular genetic evidence for probable reticulate speciation in the coral genus *Madracis* from the Caribbean fringing reef slope. *Marine Biology*, **139**, 221–233.

Duggen S, Hoernle K, Van Den Bogaard P, Rupke L, Phipps Morgan J (2003) Deep roots of the Messinian salinity crisis. *Nature*, **422**, 602–606.

Duran S, Pascual M, Estoup A, Turon X (2002) Polymorphic microsatellite loci in the sponge *Crambe crambe* (Porifera: Poecilosclerida) and their variation in two distant populations. *Molecular Ecology Notes*, **2**, 478–480.

Duran S, Pasaial M, Estoup A, Turon X (2004) Strong population structure in the sponge crambe crambe (Poecilosclevida) as revealed by ncicrosakllite marters. *Molecular Ecology*, in press.

Duran S, Pascual M, Turon X (2003) Low levels of genetic variation in mtDNA sequences over the biogeographic range of the sponge *Crambe crambe* (Poecilosclerida). *Marine Biology*, in press.

Emerson BC, Paradis E, Thebaud C (2001) Revealing demographic histories of species using DNA sequence data. *Trends in Ecology and Evolution*, **16**, 707–716.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.

Galera J, Turon X, Uriz MJ, Becerro MA (2000) Microstructure variation in sponges sharing growth form: the encrusting demosponges *Dysidea avara* and *Crambe crambe*. *Acta Zoologica*, **81**, 93–107.

Gofas S (1998) Marine molluscs with a very restricted range in the Strait of Gibraltar. *Diversity and Distributions*, **4**, 255–266.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hawkins SJ, Corte-Real HBSM, Pannacciulli FG, Weber LC, Bishop JDD (2000) Thoughts on the ecology and evolution of the intertidal biota of the Azores and other Atlantic islands. *Hydrobiologia*, **440**, 3–17.

Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarterly Review of Biology*, **66**, 411–453.

Hopkins TS (1985) Physics of the sea. In: *Western Mediterranean* (ed. Margalef R), pp. 100–125. Pergamon Press, Barcelona.

Hudson RR, Boos DD, Kaplan NL (1992) A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution*, **9**, 138–151.

Jares-Erijman EA, Sakai R, Rinehart KL (1991) Crambescidins: new antiviral and cytotoxic compounds from the sponge *Crambe crambe*. *Journal of Organic Chemistry*, **56**, 5712–5715.

Long EO, Dawid IB (1980) Repeated genes in eukaryotes. *Annual Review of Biochemistry*, **49**, 727–764.

Lopez JV, Peterson CL, Willoughby R *et al*. (2002) Characterization of genetic markers for in vitro cell line identification of the marine sponge, *Axinella corrugata*. *Journal of Heredity*, **93**, 27–36.

Maldonado A (1985) Evolution of the Mediterranean Basins and a detailed reconstruction of the Cenozoic paleoceanography. In: *Western Mediterranean* (ed. Margalef R), pp. 17–60. Pergamon Press, Oxford.

Maldonado A, Uriz MJ (1996) Skeletal morphology of two controversial Poecilosclerid genera (Porifera, Demospongiae): *Discorhabdella* and *Crambe*. *Helgoländer Meeresuntersuchungen*, **50**, 369–390.

Odorico DM, Miller DJ (1997) Variation in the ribosomal internal transcribed spacers and the 5.8*S* rRNA among five species of *Acropora* (Cnidaria; Scleractinia): patterns of variation consistent with reticulate evolution. *Molecular Biology and Evolution*, **14**, 465–473.

van Oppen MJH, Wörheide G, Takabayashi M (2002) Nuclear markers in evolutionary and population genetic studies of scleractinian corals and sponges. In: *Proceedings of the 9th International Coral Reef Symposium* (eds Moosa KM, Soemodihardjo S, Soegiarto A, Romimohtarto K, Nontji A, SoeKarno, Suharsono) Vo. 1, pp. 131–138. 23–27 October 2000, Bali, Indonesia.

Patti FP, Gambi MC (2001) Phylogeography of the invasive polychaete *Sabella spallanzanii* (Sabellidae) based on the nucleotide sequence of internal transcribed spacer 2 (ITS2) of nuclear rDNA. *Marine Ecology Progress Series*, **215**, 169–177.

Pfenninger M, Posada D (2002) Phylogeographic history of the snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution*, **56**, 1776–1788.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution*, **16**, 37–45.

Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.

Pulitzer-Finali G (1983) A collection of Mediterranean demosponges (Porifera) with, in appendix, a list of the Demospongiae hitherto recorded from the Mediterranean Sea. *Annual Museo Civico Storia Naturale Genova*, **84**, 445–621.

Rodriguez-Lanetty M, Hoegh-Guldberg O (2002) The phylogeography and connectivity of the latitudinally widespread scleractinian coral *Plesiastrea versipora* in the western Pacific. *Molecular Ecology*, **11**, 1177–1189.

Saritas MU (1972) A preliminary study of the siliceous sponge (Porifera) fauna of Engeceli Limani in the gulf of Izmir (Aegean Sea). *Scientific Reports of the Faculty of Science, Ege University*, **143**, 17–18.

Schmidt O (1862) *Die Songien Des Adriatischen Meeres*, Leipzig.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN, Version 2.000: A Software for Population Genetics Data Analysis*. University of Geneva, Switzerland.

Shearer TL, van Oppen MJH, Romano SL, Wörheide G (2002) Slow mitochondrial DNA sequence evolution in Anthozoa (Cnidaria). *Molecular Ecology*, **11**, 2475–2487.

Smith SW, Overbeek R, Woese CR, Gilbert W, Gillevet PM (1994) The Genetic Data Environment: an expandable GUI for multiple

sequence analysis. *Computer Applications in the Biosciences*, **10**, 671–675.

Solé-Cava AM, Boury-Esnault N (1999) Patterns of intra- and interspecific genetic divergence in marine sponges. *Memoirs of the Queensland Museum*, **44**, 591–601.

Swofford DL (1998) *PAUP*: Phylogenetic Analysis Using Parsimony and Other Methods*. Sinauer Associates, Sunderland, MA.

Tarjuelo I, Posada D, Crandall K, Pascual M, Turon X (2001) Cryptic species of *Clavelina* (Ascidiacea) in two different habitats: harbours and rocky littoral zones in the northwestern Mediterranean. *Marine Biology*, **139**, 455–462.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, **117**, 343–351.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767–782.

Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, **134**, 659–669.

Turon X, Tarjuelo I, Uriz MJ (1998) Growth dynamics and mortality of the encrusting sponge *Crambe crambe* (Poecilosclerida) in contrasting habitats: correlation with population structure and investment in defence. *Functional Ecology*, **12**, 631–639.

Uriz MJ, Maldonado M, Turon X, Martí R (1998) How do reproductive output, larval behaviour, and recruitment contribute to adult spatial patterns in Mediterranean encrusting sponges? *Marine Ecology Progress Series*, **167**, 137–148.

Uriz MJ, Rosell D, Martín D (1992) The sponge population of the Cabrera Archipelago (Balearic Islands): characteristics, distribution and abundance of the most representative species. *Marine Ecology PSZNI*, **13**, 101–117.

Vacelet J (1999) Planktonic armoured propagules of the excavating sponge *Alectona* (Porifera: Demospongiae) are larvae: evidence from *Alectona wallichii* and *A. mesatlantica* sp. nov. *Memoirs of the Queensland Museum*, **44**, 627–642.

Wirtz P (1998) Twelve invertebrate and eight fish species new to the marine fauna of Madeira, and a discussion of the zoogeography of the area. *Helgoländer Meeresuntersuchungen*, **52**, 197–207.

Wirtz P, Martins HR (1993) Notes on some rare and little known marine invertebrates from the Azores, with a discussion of the zoogeography of the region Arquipélago. *Life and Marine Sciences*, **11A**, 55–63.

Wörheide G, Degnan BM, Hooper JN, Reitner J (2002a) Phylogeography and taxonomy of the indo-Pacific reef cave dwelling coralline demosponge *Astrosclera 'willeyana'*: new data from nuclear internal transcribed spacers sequences. In: *Proceedings of the 9th International Coral Reef Symposium* (eds Moosa KM, Soemodihardjo S, Soegiarto A, Romimohtarto K, Nontji A, Soekarno, Suharsoho), vol. 1, pp. 339–346. 23–27 October 2000, Bali, Indonesia.

Wörheide G, Hooper JN, Degnan BM (2002b) Phylogeography of western Pacific *Leucetta 'chagosensis'* (Porifera: Calcarea) from ribosomal DNA sequences: implications for population history and conservation of the Great Barrier Reef World Heritage Area (Australia). *Molecular Ecology*, **11**, 1753–1768.

Zhou M, Paduan JD, Niiler PP (2000) Surface currents in the Canary Basin from drifter observations. *Journal of Geophysical Research*, **105**, 21893–21911.

This study was a part of Sandra Duran's PhD research project. Her research interests focus on the phylogeography and population genetics of marine invertebrates, and the relation between population structure, reproductive strategies and dispersal. Research at the laboratory of Gonzalo Giribet focuses on the study of biodiversity and its origins using the historical information contained in morphology, anatomy and molecular sequence data and the phylogenetic patters derived from those sources of data. Xavier Turon and his group work on biology of benthic invertebrates, including chemical ecology, reproductive biology, larval biology and population genetics.
