Genetic survey of shallow populations of the Mediterranean red coral [Corallium rubrum (Linnaeus, 1758)]: new insights into evolutionary processes shaping nuclear diversity and implications for conservation

J.-B. Ledoux, K. Mokhtar-Jamai, C. Roby, Jean-Pierre Feral, J. Garrabou, D. Aurelle

Molecular Ecology 19(4): 675-690
DOI: 10.1111/j.1365-294X.2009.04516.x

Genetic survey of shallow populations of the Mediterranean red coral [*Corallium rubrum* (Linnaeus, 1758)]: new insights into evolutionary processes shaping nuclear diversity and implications for conservation

J.-B. LEDOUX,* K. MOKHTAR-JAMAÏ,* C. ROBY,† J.-P. FÉRAL,* J. GARRABOU‡ and D. AURELLE*

**Aix-Marseille Université, CNRS UMR 6540 DIMAR, Centre d'Océanologie de Marseille, Station Marine d'Endoume, Chemin de la Batterie des Lions, 13007 Marseille, France, †Département de Biologie, Université Laval, Québec, Canada G1V 0A6, ‡Institut de Ciències del Mar CSIC, Passeig Maritim de la Barceloneta 37-49, E-08003, Barcelona, Spain*

Abstract

Combined action from over-harvesting and recent mass mortality events potentially linked to ongoing climate changes has led to new concerns for the conservation of shallow populations (5–60 m) of *Corallium rubrum*, an octocorallian that is mainly found in the Mediterranean Sea. The present study was designed to analyse population structure and relationships at different spatial scales (from 10s of meters to 100s of kilometres) with a focus on dispersal pattern. We also performed the first analysis of the distribution of genetic diversity using a comparative approach between regional-clusters and samples. Forty populations dwelling in four distinct regions between 14 and 60 m in depth were genotyped using 10 microsatellites. Our main results indicate (i) a generalized pair-sample differentiation combined with a weak structure between regional-clusters; (ii) the occurrence of isolation by distance at the global scale, but also within two of the three analysed regional-clusters; (iii) a high level of genetic diversity over the surveyed area with a heterogeneous distribution from regional-cluster to sample levels. The evolutionary consequences of these results are discussed and their management implications are provided.

Keywords: conservation biology, *Corallium rubrum*, genetic clustering, isolation by distance, microsatellites

Introduction

Coastal marine ecosystems are important from socio-economic and ecological points of view (Costanza *et al.* 1997; Harley *et al.* 2006). These ecosystems are subjected to the direct and indirect effects of human activities, resulting in local and global pressures. In many areas, coastal ecosystems are critically endangered by severe dysfunctions, including extinction of species (Halpern *et al.* 2008; Jackson 2008). The effects of climate change are the most recent source of strong disturbance pressures in these ecosystems worldwide (Harley *et al.* 2006).

This study is focused on the Mediterranean red coral, *Corallium rubrum*, a species that is typically associated with the coralligenous assemblage, one of the richest biodiversity habitats in the Mediterranean Sea (Ballesteros 2006; Airoldi & Beck 2007). The red coral is an aposymbiotic octocorallian with a highly fragmented distribution which centers mainly around the western Mediterranean and the neighbouring Atlantic Ocean (Zibrowius *et al.* 1984). It inhabits semi-dark to dim-light rocky benthic habitats such as caves and coralligenous overhangs between 5 and 600 m in depth (Laborel & Vacelet 1961; Zibrowius *et al.* 1984; Zibrowius

pers. comm.). It is a long-lived species (more than 50 years) with a short larval phase, estimated from 4 to 12 days *in aquaria* (Vighi 1972; Weinberg 1979), displaying slow population dynamics with very low growth and mortality rates (Garrabou & Harmelin 2002; Marschal *et al.* 2004; Santangelo *et al.* 2004, 2007; Torrents 2007), and showing recruitment by pulses (Garrabou & Harmelin 2002).

The red coral species has a high socio-economical value. It has been intensively harvested since antiquity for its use in jewellery. This important fishing pressure resulted in profound changes in the species range (Santangelo & Abbiati 2001) and in the size structure of the shallow populations, preventing the colonies from reaching their potential maximum size (Garrabou & Harmelin 2002). Moreover, in 1999 and 2003, the red coral was among 30 species affected by two large-scale mass mortality events (MMEs), which were linked to positive thermal anomalies that were potentially associated with the warming trend currently detected in this area (Perez *et al.* 2000; Romano *et al.* 2000; Garrabou *et al.* 2001, 2009). MMEs have differentially impacted red coral populations. In some sites, approximately 80% of the colonies were affected, while in others only 5% of colonies showed some mortality (Garrabou *et al.* 2001). The red coral was included in several international conservation conventions to protect the species and control its harvest. However, these management measures do not take into account the distribution of genetic diversity, the connectivity between populations and the new climatic threat and its impact on shallow populations (5–60 m).

Previous genetic studies investigated effective larval dispersal as well as the spatial genetic structure of red coral populations; these studies used nuclear markers because of the lack of polymorphism in mitochondrial markers (Calderón *et al.* 2006) putatively linked to the presence of the *msh1* gene (Pont-Kingdon *et al.* 1995; McFadden *et al.* 2006; Ledoux *et al.* unpublished). Based on use of allozymes, significant differentiation was observed between samples separated by 10 km (Abbiati *et al.* 1993). Del Gaudio *et al.* (2004) used AFLP profiles to distinguish four population samples separated by tens to hundreds of kilometres. These patterns of differentiation were refined by analysis of four microsatellites among eight samples derived from two locations. This analysis revealed strong genetic structuring at distances of tens of metres (Costantini *et al.* 2007a). Finally, a large-scale study of 11 samples scattered around the northwestern Mediterranean basin, using ITS-1 sequences and five microsatellites confirmed this high level of differentiation between populations. However, a clear genetic structure could not be established at long distances despite generalized differentiation between the samples (Costantini *et al.* 2007b). Moreover, the ITS-1 signal might be affected by a lack of concerted evolution (Calderón *et al.* 2006).

The major goals of this study were to analyse the link between geographical distance and genetic structure and to gain insight into the distribution of genetic diversity. Using 10 microsatellite loci and a sampling of 40 shallow populations separated by a range from tens of metres to hundreds of kilometres and dwelling in different environments, three main objectives were addressed: (i) to thoroughly define the population structure and relationships with particular emphasis on dispersal patterns over the whole area and within each regional-cluster; (ii) to conduct the first analysis of the patterns of genetic diversity using comparative approaches; and (iii) to test for deviation from mutation-drift equilibrium impacting the genetic diversity of the studied samples. The results, discussed in light of the species' life history traits, should further our understanding of the evolutionary history of the shallow red coral populations. In particular, we show for the first time that isolation by distance (IBD) occurs at different spatial scales in this species. Moreover, we reveal contrasting patterns of genetic structure among the analysed regional-clusters. We also demonstrate a strong heterogeneity in the distribution of genetic diversity and we highlight the particular status of the Balearic cluster. These results constitute valuable data for conservation of red coral populations and could also be useful for other species with similar life-history traits belonging to the threatened Mediterranean rocky benthic community. Finally, this study is necessary for exploration of the adaptive capacity of red coral populations facing the ongoing shifts in environmental conditions (Ferrière *et al.* 2004).

Materials and methods

Sampling and DNA extraction

Corallium rubrum colonies were sampled by scuba diving at 40 sites corresponding to different habitats (overhangs, caves and interior of caves) and located at the upper limit (between 14 and 60 m in depth) of the bathymetric distribution of the species between the years 2003 and 2006. Two to 22 sites were thus sampled in four geographically distinct parts of the species range: the northwestern part of Corsica, Catalonia, the northern part of the Balearic Islands and the Liguro-Provençal region (Fig. 1; Table 1). These regions are well separated based on their insular position or due to the lack of continuous shallow rocky habitats between them (such as between Catalonia and the Liguro-Provençal region). One apical fragment measuring 2–3 cm, from around 30 colonies per site (from 21–50; mean

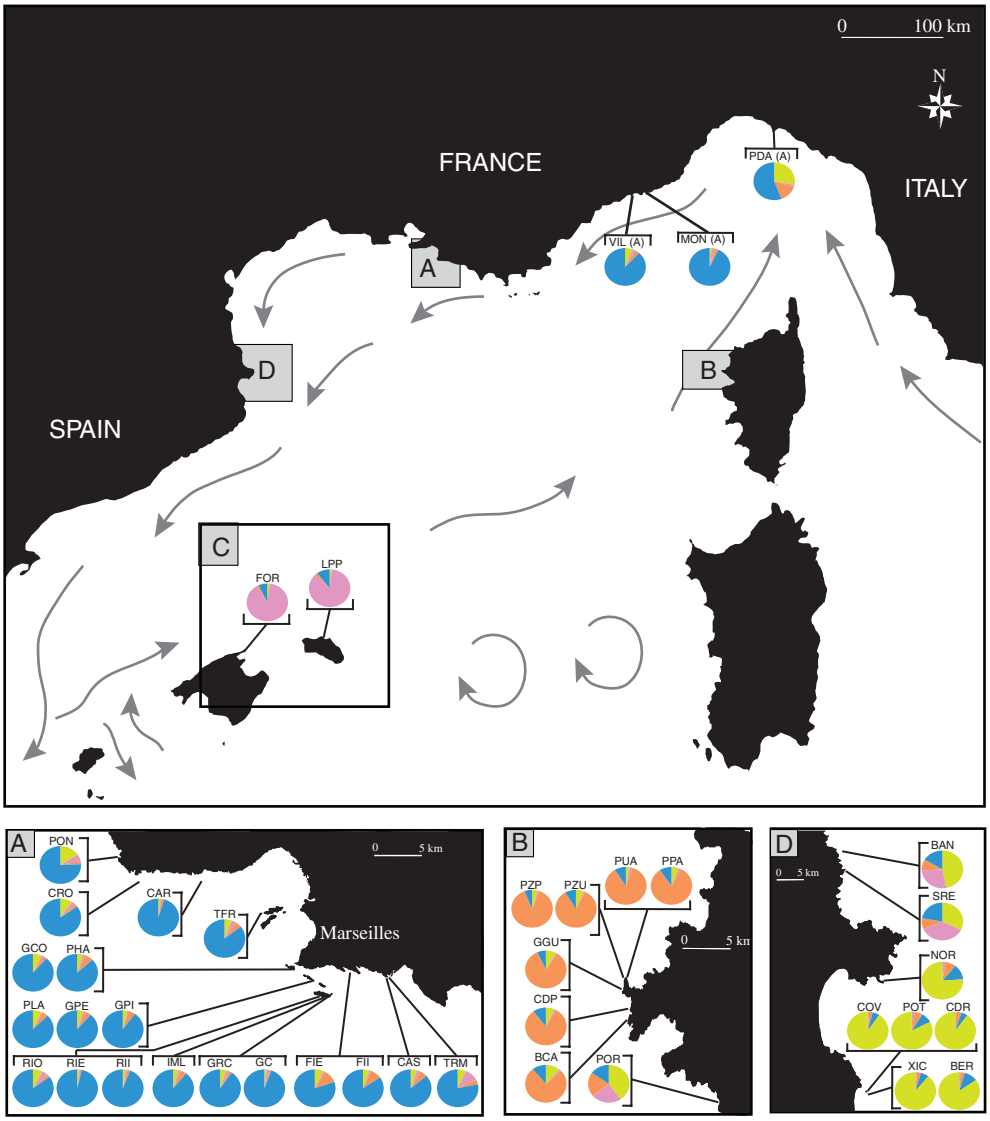


Fig. 1 Location of 40 sites distributed among the four regions (A: Liguro-Provençal, B: Corsica, C: Balearic Islands, D: Catalonia) analysed in the present study. Correspondences between sample names and locations are shown in Table 1. Grey arrows represent the predominant sea surface currents as defined in Millot (1999). Circles correspond to assignment results for each sample obtained with STRUCTURE for $K = 4$.

30.5 ± 5.2), was hand-collected randomly to cover the area of each study site. The resulting 1222 red coral fragments were preserved in 95% ethanol and stored at −80°C until DNA extraction. Total genomic DNA was extracted from 5–10 polyps using standard proteinase K/SDS digestion followed by phenol-chloroform-isoamyl alcohol (25:24:1) extraction (Sambrook *et al.* 2001).

Microsatellite analysis

All individuals were genotyped using 10 microsatellite loci (Table 2) from three different sources: Ecogenics GmbH, Bioprofiles LTD (Newcastle University) and Costantini & Abbiati (2006) (see Supporting Information). PCR amplifications were realized in a final reaction volume of 10 μL containing 1 μL of DNA, 2 μL of 5X Colourless GoTaq Flexi Buffer (Promega), 0.2 mM of dNTPs, 0.25 U of GoTaq DNA Polymerase (Promega), 3 mM of $MgCl_2$ and 1 μM of each primer. The forward primer was fluorescently labelled. The PCR reactions were done as follows for all loci, except COR9bis and COR46bis: 95 °C for 2 min followed by 25 cycles of 94 °C for 30 s and annealing temperature for 30 s without elongation; the final step was 72 °C for 20 min (C. Hubert Functional Genomic Bordeaux, pers. comm.). Amplifications of the two remaining loci were done

Table 1 *Corallium rubrum* samples collected for the study

Geographic region	Sample location	Label	Latitude (North)	Longitude (East)	Depth (m)	Habitat	*N*
Provençe-Liguria	Punta dell'Altare	PDA	44.3055°	9.21481°	35	2	29
Provençe-Liguria	Monaco	MON	43.7349°	7.43342°	35	2	21
Provençe-Liguria	Villefranche	VIL	43.6896°	7.34897°	22	1	36
Provençe-Liguria	Grotte Tremies	TRM	43.2012°	5.51339°	14	3	26
Provençe-Liguria	Castelvieil	CAS	43.1976°	5.49875°	16	2	30
Provençe-Liguria	Figuier extérieur	FIE	43.205°	5.44723°	20	2	29
Provençe-Liguria	Figuier intérieur	FII	43.205°	5.44736°	20	3	27
Provençe-Liguria	Grand Congloué	GC	43.1759°	5.40185°	40	1	30
Provençe-Liguria	Grand Congloué	GRC	43.1754°	5.40097°	52	2	29
Provençe-Liguria	Impériales du large	IML	43.1698°	5.39435°	60	2	23
Provençe-Liguria	Riou Sud	RIO	43.1728°	5.3886°	40	2	29
Provençe-Liguria	Riou Sud Extérieur	RIE	43.1728°	5.38942°	20	2	25
Provençe-Liguria	Riou Sud Intérieur	RII	43.1729°	5.38941°	20	3	33
Provençe-Liguria	Plane	PLA	43.1869°	5.39117°	38	2	29
Provençe-Liguria	Grotte à Péres entrée	GPE	43.1869°	5.39029°	20	2	50
Provençe-Liguria	Grotte à Péres intérieur	GPI	43.1868°	5.39029°	20	3	28
Provençe-Liguria	Pharillons de Maïre	PHA	43.2074°	5.3382°	39	1	32
Provençe-Liguria	Tiboulen du Frioul	TFR	43.28°	5.2876°	38	1	38
Provençe-Liguria	Grotte à corail	GCO	43.2102°	5.33273°	20	3	29
Provençe-Liguria	Carry le Rouet	CAR	43.3226°	5.162°	20	1	35
Provençe-Liguria	Carro	CRO	43.3151°	5.06052°	21	1	32
Provençe-Liguria	Ponteau / Lavera	PON	43.3643°	4.99941°	21	1	28
N-W Corsica	Grotte Palazzinu	PUA	42.3798°	8.55009°	40	2	32
N-W Corsica	Passe Palazzu	PPA	42.3799°	8.54758°	27	1	29
N-W Corsica	Palazzu (grotte à corail)	PZU	42.3802°	8.54635°	22	2	34
N-W Corsica	Palazzu	PZP	42.3802°	8.54575°	40	2	28
N-W Corsica	Garganellu	GGU	42.3726°	8.5369°	43	2	32
N-W Corsica	Cala di Ponte	CDP	42.3546°	8.55209°	26	1	30
N-W Corsica	Baja casju	BCA	42.3496°	8.55095°	30	1	28
N-W Corsica	Porto	POR	42.2725°	8.68812°	24	1	29
Catalonia	Cap de l'Abeille	BAN	42.475°	3.16192°	25	1	21
Catalonia	Sec Rederis	SRE	42.4641°	3.16658°	23	2	27
Catalonia	Cap de Norfeu	NOR	42.2443°	3.26684°	24	1	30
Catalonia	Pota del Lop	POT	42.0497°	3.2254°	35	2	35
Catalonia	Cova de la Reina	CDR	42.0461°	3.22477°	14	1	38
Catalonia	Cova de la vaca SE	COV	42.0472°	3.22583°	18	1	39
Catalonia	Carall Bernat	BER	42.0422°	3.22816°	18	1	27
Catalonia	Meda xica tunel dofi SE	XIC	42.0436°	3.22644°	35	1	34
Balearic Islands	Cap Formentor	FOR	39.9964°	3.21188°	36	1	32
Balearic Islands	Llosa des Patro Pere	LPP	40.0726°	4.1075°	34	1	29

Habitat numbers correspond respectively to 1 = overhang; 2 = cave entrance; 3 = inside cave. *N*: number of colonies sampled at each site.

following Costantini & Abbiati (2006), but with a last extension step at 72 °C for 20 min.

Electrophoresis of the amplification products was conducted on an ABI3130 genetic analyser (Applied Biosystems) by multiplexing PCR products (Mic13/-Mic22/Mic24/Mic27, Mic23/Mic25/Mic26 and COR9-bis/Mic20/COR46bis) and using an internal size standard for sizing (Genescan 600 LIZ, Applied Biosystems). Three reference individuals were included in each migration. Electropherograms were analysed with Genemapper version 3.0 (Applied Biosystems).

Microsatellite characteristics

MICROCHECKER v.2.2.3 (Van Oosterhout *et al.* 2004) was used to test for large allele dropout and scoring errors due to stutters. Frequencies of null alleles were estimated for each sample and locus by the expectation maximization algorithm (Dempster *et al.* 1977) implemented in FREENA (Chapuis & Estoup 2007). Repeatability tests based on extraction and genotyping replicates were performed to control the genotyping procedure. The results obtained suggested that PCR

Table 2 Primer sequences, PCR conditions and genetic characteristics of 10 polymorphic microsatellites in *Corallium rubrum*

Locus name	Primer sequence (5′-3′)	Repeat motif	Number of cycles	Ta (°C)	Size range (bp)	Number of alleles	Ho	Hs	f	r	Genbank accession number
Mic13	F: NED_CTTTGATTGGCCCTGATGTAA R: GCCAGGAAAGAATTGGGTATATTA	$(AC)_2$ A (AC) G $(AC)_3$ TA $(AC)_7$	25	59	126–141	7	0.18	0.53	0.65***	0.24	GQ169280
Mic20	F: 6-FAM_CACGTGATTGACGAAAACATTC R: TGTCGGGAAATTGTTCACTGTA	$(CA)_8$	25	59	198–286	27	0.49	0.54	0.07***	0.07	GQ169281
Mic22	F: VIC_CGAGCGAGGGAAATTAATAGG R: GATGTAATTGTCGCGCATTG	$(GT)_{16}$	25	56	153–190	18	0.3	0.64	0.49***	0.37	GQ169282
Mic23	F: VIC_GATCTCTGACTGAATGGTATTGG R: CCTGGCTACGTCCCTGACT	$(GT)_{14}$	25	56	93–143	23	0.35	0.71	0.47***	0.4	GQ169283
Mic24	F: NED_ TCGAGCACTTCCTTGGTAGC R: TGAATTCCATACACCCACTGC	$(CA)_{18}$	25	59	145–312	55	0.73	0.76	0.04***	0.02	GQ169284
Mic25	F: 6-FAM_ GCAAGGTAAAATGATGTAGTCTGG R: GATCGCACTAAATCTTAATAGTGTTCC	$(GTTT)_3(GT)_{16}$	25	59	130–208	38	0.39	0.78	0.46***	0.38	GQ169285
Mic26	F: NED_AGGGAACAATCTTTGTTGTGC R: ATGTTTGCGGACCTACACG	$(GT)_{24}$	25	59	126–200	30	0.89	0.89	−0.005	0.01	GQ169286
Mic27	F: 6-FAM_ GATCTCTTCGCGGATAGTCTG R: GACGGTGGGACGAACAGG	$(GT)_{30}$	25	59	140–536	67	0.74	0.94	0.19***	0.16	GQ169287
COR9bis	F: 6-FAM_GTTCCTTGAAGCATGTTGGATT R: CAGGAACTTCTGCCAGTACACA	$(C)_{11}(CA)_{13}$	30	53	141–200	56	0.27	0.89	0.7***	0.4	AY726758
COR46bis	F: NED_TTGGGTACAAATCAAGCTACCA R: AGACCAGCGGCATCACTTT	$(GT)_{15}$	28	53	172–243	36	0.72	0.76	0.04***	0.04	AY726761

Ho: observed heterozygosity; He: gene diversity (Nei 1967); f: Weir & Cockerham (1984) estimator of F_{IS}; *r*: null allele frequency. Presented values correspond to the mean value over all samples.

reactions and genotyping were repeatable (data not shown).

The total number of alleles, size range, observed (Ho) and unbiased heterozygosity (He Nei 1973) were calculated for each locus using FSTAT v.2.9.3 (Goudet 2001).

Hardy–Weinberg equilibrium

GENETIX v.4.05 (Belkhir *et al.* 2004) was used to test the null hypothesis of linkage equilibrium for each pair of loci in each sample with a permutation procedure ($n = 1000$). Departure from panmixia was tested for each sample using the score test for heterozygote deficiency in GENEPOP 4.0 (Rousset 2008). Significance was addressed by a Markov Chain (MC) algorithm (Guo & Thompson 1992; Raymond & Rousset 1995), with default parameters. The f estimator of F_{IS} (Weir & Cockerham 1984) was computed for each sample using GENETIX.

Interpopulation differentiations and isolation by distance

Global and pairwise differentiations between samples were quantified using Weir and Cockeram's (1984) estimator of F_{ST} (θ) in GENETIX. Genotypic differentiation between samples was tested using an exact test based on the MC algorithm (Guo & Thompson 1992) with default parameters in GENEPOP.

Pattern of isolation by distance was analysed following Rousset's method (1997). Geographic distances were measured using GOOGLE EARTH v.5.0 (http://earth.google.fr/), using Euclidian distance or taking into account the minimum distances between regions along the path of the predominant surface current as defined in Millot (1999) (Fig. 1 and Table S3, Supporting Information). Slatkin's linear F_{ST} $[F_{ST}/(1 - F_{ST})]$ was regressed on the logarithm of the distance [Ln(d)], as recommended for a two-dimensional model of IBD (Rousset 1997) using GENEPOP. The significance of the correlation between the two distance matrices was tested using a Mantel test (1967) ($n = 2000$) in GENEPOP. IBD was also assessed independently over each of the three clusters (Liguro-Provençal, Corsican and Catalonian) for which the number of available samples was sufficient.

Hierarchical analyses of molecular variance (AMOVA) were conducted in ARLEQUIN v.3.10 (Excoffier *et al.* 2005) in order to quantify genetic variation among samples (F_{ST}), among samples within a group (Fsc) and among groups (Fct). Significance of F-statistics was achieved using 1000 permutations. AMOVA tests were performed first using the groups defined by STRUCTURE v.2.2 (Pritchard *et al.* 2000) for $K = 4$ (see below) and then within each of these groups, except for the Balearic cluster because of its low number of samples.

Clustering analyses

Simulated dataset. We used computer simulations to evaluate the effect of positive and significant F_{IS} on STRUCTURE, considering genetic parameters close to those observed for *C. rubrum*. The behaviour of the model allowing for separate alpha values among clusters was also evaluated under the simulated scenarios.

The obtained results showed that the two models (admixture with correlated allele frequencies with the same alpha or with separated alphas) gave good results despite HW equilibrium not being met in all samples (see Supporting Information).

Real dataset. The underlying population structure contained in the data was analysed using the Bayesian approach implemented in STRUCTURE v.2.2 (Pritchard *et al.* 2000; Falush *et al.* 2003, 2007) that infers the number of genetic clusters K from the individual's genotypes dataset. Ten different runs of 200 000 burn-in followed by 500 000 iterations were computed for each K value, with the model allowing for admixture and correlated allele frequencies between clusters, and using the recessive allele option to deal with null alleles (Falush *et al.* 2007). Alpha, the degree of admixture between the K clusters, was evaluated separately for each cluster. Following the recommendations of Rosenberg *et al.* (2002) and Jakobsson *et al.* (2008) for cases of large datasets, we first studied the behaviour of STRUCTURE at several small values of K during a first round of analysis and then we looked for additional substructure on partitioned datasets during a second round. For the first round, the whole dataset (1222 individuals) was used. K was set to vary between 1 and 8 and only the smallest values were discussed (from $K = 2$ to $K = 4$). The clustering solution given for $K = 4$ was then used to divide the whole dataset into four partitioned datasets that were subsequently analysed in a second round of STRUCTURE with K varying between 1 and 23 (i.e. the total number of samples plus one). A K value was then selected for each partitioned dataset based on the plot of *LnP(D)* (the logarithm of the likelihood of observing the data) as a function of K (Waples & Gaggiotti 2006; Pritchard *et al.* 2007) and looking for the value that captured the major structure in the data (Pritchard *et al.* 2007). When only one clustering solution appeared for a selected K value, CLUMPP v.1.1 (Jakobsson & Rosenberg 2007) was used to average the assignment scores over the 10 runs. If several solutions were observed, the symmetric similarity coefficient (*SSC*) was used to pool runs belonging to the same mode (i.e. $SSC > 0.9$)

(Jakobsson *et al.* 2008) and the mode with the highest likelihood was selected. Finally, DISTRUCT v.1.1 (Rosenberg 2004) was used for graphical display.

Estimation and distribution of genetic diversity

Genetic diversity analyses focused on gene diversity (He), allelic richness [Ar(*g*)] and private allelic richness [Ap(*g*)]. A rarefaction method (Petit *et al.* 1998) was applied to obtain estimates of Ar(*g*) and Ap(*g*) independently of the sample size, with *g* representing the minimum number of genes observed at one locus in one of the samples (i.e. twice the number of genotypes). These computations were done for *g* equal to 22 (Mic22 for MON) or 66 (Mic22 for Catalonia), depending on whether the samples or the four regional-clusters defined by STRUCTURE were considered.

The coefficient Arst(*g*) was computed for $g = 22$ to measure the partitioning of allelic richness among samples (Petit *et al.* 1998) using ADZE (Szpiech *et al.* 2008). ADZE was also used to compute Ar(*22*) and Ap(*22*) for each sample and Ar(*66*) and Ap(*66*) for each cluster. Computations of He were done with FSTAT. Comparisons of He and Ar(*66*) values between each cluster were done using pair-group comparisons and the one-sided probability test implemented in FSTAT.

The occurrence of deviation from mutation-drift equilibrium linked to recent demographic events and affecting the genetic diversity within samples was tested using BOTTLENECK 1.2.02 (Piry *et al.* 1999). The expected heterozygosity under mutation-drift equilibrium (Heq) was computed for each locus considering a two-phase model (TPM) (Di Rienzo *et al.* 1994), with 95% of the stepwise mutation model (SMM) and variance among multiple steps equal to 12 (Piry *et al.* 1999). The null hypothesis of 'no significant difference between He and Heq' was then tested using the Wilcoxon's signed rank-test.

For all analyses, significance levels were corrected using a false discovery rate (FDR) correction for multiple tests (Benjamini & Hochberg 1995) when necessary.

Results

Microsatellite characteristics

According to MICROCHECKER, no evidence of scoring errors due to stuttering or large allele dropout was found in the whole data set. All loci were polymorphic in all populations. Total numbers of alleles ranged between 7 for Mic13 and 67 for Mic27, with a mean of 35 alleles per locus. Observed heterozygosity varied from 0.18 for Mic13 to 0.89 for Mic26. Unbiased heterozygosity ranged from 0.53 for Mic13 to 0.94 for Mic27. Mean estimated null allele frequencies by locus varied from 0.01 for Mic26 to 0.4 for Mic23 and COR9bis, with a mean value of 0.21 per locus (Table 2). No global significant LD among loci was detected considering overall samples (all $P > 0.05$ after FDR correction).

Deviations from Hardy–Weinberg equilibrium

Significant linkage disequilibrium (LD) among pairs of loci was observed in 10 samples, implying different pairs of loci and different numbers of pairs of loci (from 1 for CAR to 4 for XIC). Significant heterozygote deficiencies were observed in all samples (Table 3). Multilocus values of the F_{IS} estimator (f) ranged between 0.2 for XIC and PDA and 0.47 for GRC with a mean value equal to 0.31. Considering each locus separately, f values ranged from −0.3 for Mic20 for PDA to 1 for Mic13 for PDA, MON and TRM and for Mic25 for CDP. Departures from panmixia varied between loci and samples. Whereas COR9bis showed significant heterozygote deficiencies in all samples, the null hypothesis of panmixia was not rejected for Mic26 in any sample (Table S2). Mean estimated null allele frequency (*r*) values over all loci by sample were scattered from 0.10 for PPA to 0.35 for PON, with a mean of 0.21. Wahlund effect and/or inbreeding could partially explain the significant heterozygote deficiencies and the LD reported in 10 samples.

Population structure analysis

F_{ST}, IBD and AMOVA. Global F_{ST} was 0.097, and pairwise F_{ST} between all pairs of samples ranged from −0.001 between GPE and GPI to 0.212 between RIE and PZU. After the FDR correction, 779 among 780 pairwise comparisons appeared significant (Table S3, Supporting Information), suggesting a generalized differentiation between samples. Significant differentiations were observed even between samples belonging to the same site and depth but to different environments, and separated by less than 10 m (e.g. RIE vs. RII; FIE vs. FII but GPE vs. GPI). The non-significant comparison implied GPI and GPE that belong to the same site and depth but to different environments (interior cave vs. cave entrance).

The correlation between $F_{ST}/(1 - F_{ST})$ and Ln(*d*) was significant using Euclidian distance ($r_{Ln(d)} = 0.377$; $P < 0.001$) or taking into account major surface currents between regions ($r_{Ln(d)} = 0.382$; $P < 0.001$), confirming the occurrence of an IBD model of gene flow (Fig. 2A). At the cluster scale, correlations were also significant for the Liguro-Provençal ($r_{Ln(d)} = 0.530$; $P < 0.05$) and the Catalonian clusters ($r_{Ln(d)} = 0.864$; $P < 0.01$) but not

Table 3 Measures of genetic diversity for 40 shallow samples of *Corallium rubrum* based on 10 microsatellites loci

	Ho	He	f	Ar(22)	Ap(22)	Pw10
PDA	0.54 (0.29)	0.68 (0.23)	**0.2**	6.47 (3.56)	0.19 (0.35)	0.1
MON	0.46 (0.3)	0.62 (0.25)	**0.26**	5.7 (3.62)	0.04 (0.13)	0.14
VIL	0.5 (0.29)	0.73 (0.16)	**0.31**	6.61 (3.03)	0.27 (0.42)	**0.01**
TRM	0.47 (0.34)	0.73 (0.16)	**0.37**	7.05 (3.78)	0.05 (0.1)	**0.01**
CAS	0.55 (0.26)	0.76 (0.16)	**0.28**	8.07 (3.35)	0.04 (0.08)	**0.02**
FIE	0.48 (0.24)	0.75 (0.13)	**0.36**	7.17 (2.88)	0.16 (0.41)	0.14
FII	0.43 (0.22)	0.72 (0.16)	**0.41**	6.98 (2.72)	0.12 (0.2)	**0**
GC	0.49 (0.28)	0.71 (0.18)	**0.31**	6.91 (3.33)	0.03 (0.05)	**0.01**
GRC	0.4 (0.29)	0.75 (0.19)	**0.47**	7.79 (3.53)	0.05 (0.09)	**0.01**
IML	0.48 (0.32)	0.77 (0.19)	**0.38**	8.31 (3.97)	0.13 (0.31)	0.31
RIO	0.5 (0.35)	0.83 (0.13)	**0.41**	9.16 (4.08)	0.17 (0.3)	0.25
RIE	0.53 (0.37)	0.67 (0.23)	**0.21**	6.13 (3.06)	0 (0)	**0.02**
RII	0.52 (0.29)	0.79 (0.13)	**0.35**	7.48 (3.54)	0.13 (0.28)	0.35
PLA	0.47 (0.3)	0.75 (0.13)	**0.37**	6.83 (3.03)	0.04 (0.09)	0.12
GPE	0.47 (0.32)	0.75 (0.18)	**0.37**	7.6 (3.22)	0.12 (0.3)	**0**
GPI	0.49 (0.26)	0.75 (0.17)	**0.35**	7.23 (3)	0.02 (0.03)	0.05
PHA	0.46 (0.33)	0.67 (0.27)	**0.32**	6.85 (4)	0.19 (0.31)	**0.01**
TFR	0.49 (0.33)	0.72 (0.19)	**0.33**	7.12 (4.1)	0.11 (0.18)	0.19
GCO	0.46 (0.34)	0.73 (0.14)	**0.38**	6.81 (3.51)	0.06 (0.13)	**0.01**
CAR	0.48 (0.34)	0.66 (0.19)	**0.27**	5.15 (2.76)	0.15 (0.3)	0.1
CRO	0.52 (0.32)	0.77 (0.16)	**0.33**	7.96 (3.68)	0.11 (0.24)	**0.02**
PON	0.43 (0.3)	0.76 (0.16)	**0.44**	7.78 (3.38)	0.04 (0.06)	0.1
PUA	0.57 (0.27)	0.75 (0.17)	**0.25**	7.86 (3.71)	0.24 (0.32)	**0.01**
PPA	0.55 (0.31)	0.72 (0.23)	**0.23**	7.23 (3.69)	0.16 (0.26)	0.1
PZU	0.46 (0.22)	0.7 (0.18)	**0.34**	6.2 (3.42)	0 (0.01)	0.19
PZP	0.57 (0.25)	0.72 (0.21)	**0.21**	6.85 (3.73)	0.09 (0.18)	0.25
GGU	0.52 (0.26)	0.74 (0.18)	**0.3**	6.7 (3.29)	0.11 (0.17)	0.25
CDP	0.46 (0.29)	0.72 (0.13)	**0.37**	6.24 (3.03)	0.03 (0.06)	0.14
BCA	0.6 (0.3)	0.8 (0.13)	**0.25**	8.34 (3.46)	0.12 (0.14)	0.12
POR	0.44 (0.25)	0.68 (0.21)	**0.37**	6.37 (3.17)	0.26 (0.37)	0.19
BAN	0.58 (0.28)	0.8 (0.1)	**0.29**	7.53 (3.23)	0.02 (0.05)	0.61
SRE	0.42 (0.28)	0.72 (0.2)	**0.42**	6.94 (2.58)	0.03 (0.06)	0.12
NOR	0.57 (0.34)	0.76 (0.17)	**0.26**	7.38 (3.61)	0.03 (0.06)	0.19
POT	0.54 (0.18)	0.76 (0.16)	**0.3**	8.14 (3.49)	0.06 (0.1)	**0.01**
CDR	0.5 (0.25)	0.68 (0.21)	**0.28**	6.98 (3.46)	0.04 (0.07)	**0.01**
COV	0.55 (0.26)	0.75 (0.16)	**0.27**	8.07 (3.67)	0.09 (0.13)	**0.01**
BER	0.56 (0.23)	0.76 (0.14)	**0.27**	7.94 (3.51)	0.05 (0.14)	0.05
XIC	0.56 (0.24)	0.74 (0.17)	**0.2**	7.66 (3.25)	0.1 (0.16)	**0.01**
FOR	0.62 (0.33)	0.78 (0.19)	**0.22**	9.48 (5.43)	1 (1.94)	0.08
LPP	0.57 (0.26)	0.82 (0.12)	**0.31**	9.08 (3.61)	0.55 (0.49)	0.053
Mean Value	0.51 (0.05)	0.74 (0.04)	**0.32**	7.3 (0.91)	0.13 (0.17)	

Ho: observed heterozygosity; He: gene diversity (Nei 1967); f: Weir & Cockerham (1984) estimator of F_{IS}; Ar*(22)* and Ap*(22)*: allelic and private allelic richness, respectively, with rarefaction for a corresponding sample size of 22; Pw10: *P* values of the Wilcoxon test for differentiation between He and Heq as implemented in BOTTLENECK using the 10 loci. Standard deviations are in brackets. Values in bold are significant at the 0.05 level after FDR correction.

for the Corsican cluster ($r_{Ln(d)}$ = 0.168; P = 0.39) (Fig. 2B).

The AMOVA showed that differentiation among the four clusters and among samples within these clusters was highly significant. However, the percentage of total variation explained by grouping into clusters was around three times lower than the variation explained by the differences among samples within clusters (2.7% vs. 6.7%, respectively). Within a cluster, 7.9%, 7.8% and 4.2% of the total variation was observed among samples whereas the variation within samples represented 92.1%, 92.2% and 95.8% for the Corsican, Liguro-Provençal and Catalonian clusters, respectively. In every case, the corresponding F-statistics were significant ($P < 0.001$) (Table 4).

Clustering with STRUCTURE. During the first round of STRUCTURE, each new cluster appeared along plausible

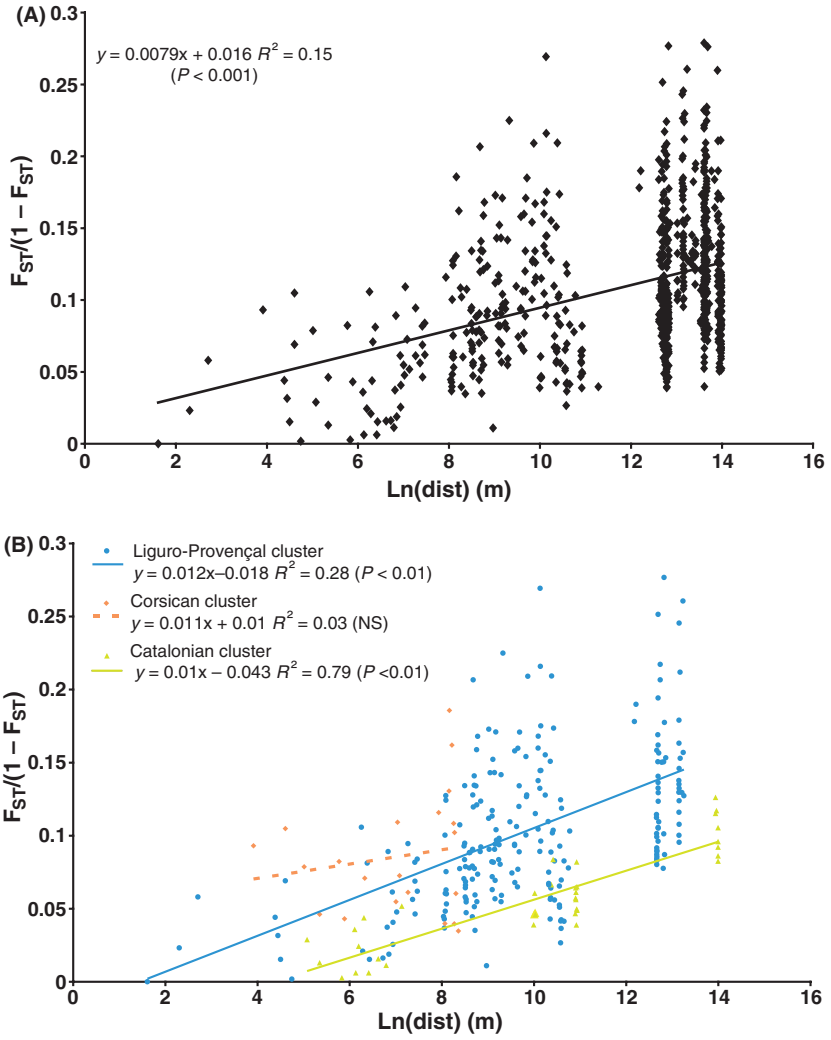


Fig. 2 Isolation by distance pattern (A) in the surveyed area, taking into account the minimum distances between regions along the path of the predominant surface current and (B) within each cluster. The two graphs show the linear regression of the genetic distance measured as $F_{ST}/(1 - F_{ST})$ over logarithms of the geographic distance (m). Solid and dashed lines represent significant and non-significant regressions, respectively.

geographical boundaries for K ranging from 2 to 4 (Fig. 3). For $K = 2$, a clear distinction was made between the 21 Liguro-Provençal samples (except PDA) and the remaining samples. For $K = 3$, the N-W Corsica/Balearic samples were separated from the Catalonian ones. POR and PDA samples, which belonged to Corsican and Liguro-Provençal regions, respectively, were assigned in the cluster with Catalonian samples. For $K = 4$, two different clustering solutions were observed. Based on likelihood criteria, the retained mode separated samples into four clusters that respected the four geographical regions, with the exception of POR (Corsica), which was pooled with the Catalonian samples. Samples with a mixed assignment among the four clusters (POR, PDA, BAN and SRE) were assigned to the cluster with the highest mean assignment scores. Each of these four clusters was then submitted to a second round of STRUCTURE (Fig. 3). For the Liguro-Provençal cluster, $K = 5$ gave the solution that captured most of the structure in the data. Following geographical distribution, CAR, PDA and MON/VIL represented three clusters. The two remaining clusters were centred on RIE and PHA and grouped quite distant samples such as TRM and TFR. For the Corsican cluster, $K = 5$ was selected. PUA, PZP and BCA belonged to one cluster and each of the remaining samples represented their own group. For the Catalonian cluster, the retained solution was $K = 4$. POR represented one cluster. BAN and SRE were grouped together. NOR was isolated in its own cluster and POT, CDR, COV, XIC and BER belonged to the fourth one. This clustering

Table 4 Results of the analysis of molecular variance (AMOVA): (a) on the whole dataset partitioning genetic variance among *Corallium rubrum* colonies, among samples within cluster and among clusters; (b) within each identified cluster partitioning genetic variance among colonies and among samples

	Grouping based on Structure results for $K = 4$			Liguro-Provençal cluster			Corsican cluster			Catalonian cluster		
Source of variation	DDL	% of variance	*P*	DDL	% of variance	*P*	DDL	% of variance	*P*	DDL	% of variance	*P*
Among clusters	3	2.7	<0.001									
Among samples within cluster	36	6.7	<0.001	21	7.8	<0.001	6	7.9	<0.001	8	4.2	<0.001
Among individuals within samples	2404	90.6	<0.001	1314	92.2		419	92.1		551	95.8	

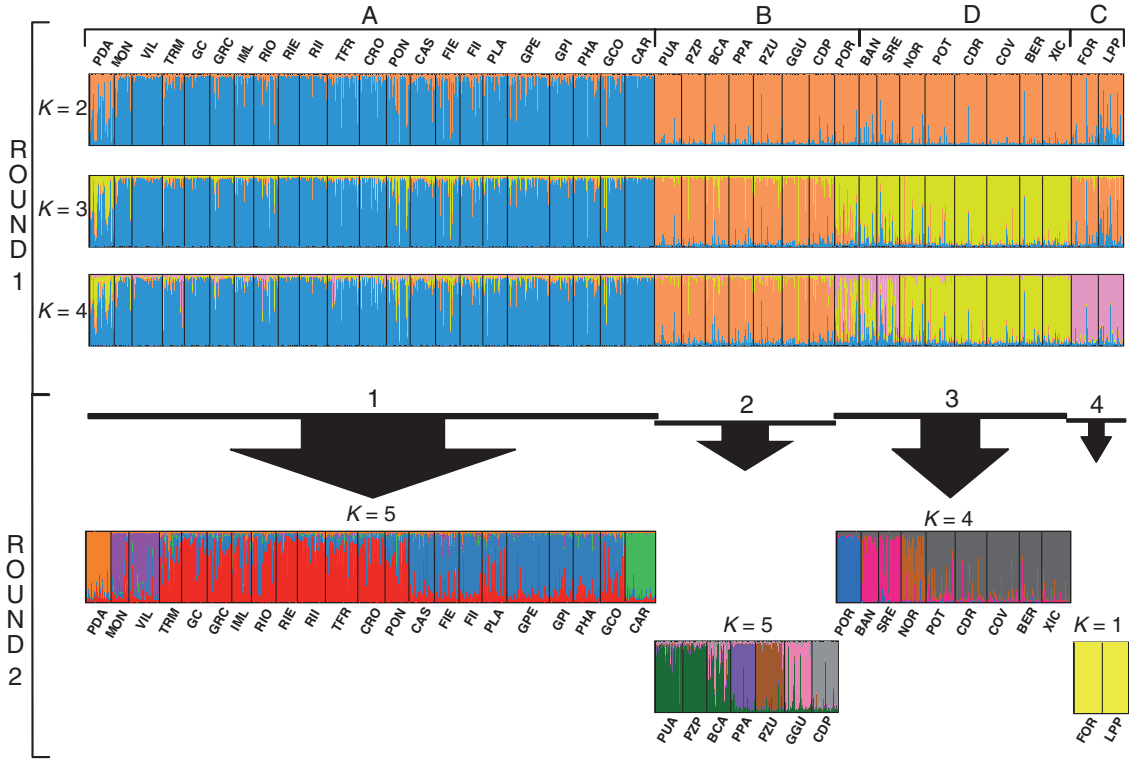


Fig. 3 Population structure as inferred by two rounds of hierarchical STRUCTURE analysis. Each individual is represented by a vertical line partitioned into *K*-coloured segments that represent the individual's membership fraction in *K* clusters. Each sample is delineated by black vertical lines and named as in Table 1. Letters correspond to the four sampled regions (A: Liguro-Provençal, B: NW Corsica, C: Balearic Islands, D: Catalonia). Numbers above the black arrows correspond to the four genetic clusters defined after the first round of STRUCTURE (1: Liguro-Provençal cluster, 2: Corsican cluster, 3: Catalonian cluster, 4: Balearic cluster).

followed the geographical partitioning of the samples. Finally, for the Balearic cluster, STRUCTURE failed to detect any structure. In every identified cluster, the mean value of alpha (admixture parameters) was relatively constant (variation of < 0.2), confirming the existence of a real structure. This value was always inferior to 1 (mean ± SD = 0.08 ± 0.05 over the two rounds), implying that levels of admixture between identified clusters are very small.

Genetic diversity analysis

There were on average 7.4 alleles per polymorphic locus and per sample, considering 22 gene copies

[Ars(22)]. At the global level, the mean Art(22) was equal to 10.2 alleles per locus, resulting in a high Arst(22) equal to 0.3. This value was larger than the global F_{ST} value (0.097). Using the rarefaction method, 47.4 alleles (16.8%) were observed in all clusters. The Balearic cluster showed the most important private allelic richness [Ap(*66*)], with 41.5 alleles (14.7%) private to this cluster, followed by the Liguro-Provençal cluster with 22.2 (7.9%) and the Catalonian and Corsican clusters with 21.2 (7.5%) and 17.7 (6.3%), respectively. Values of Ar(*66*) were equal to 7.1, 7.1, 7.5 and 9.3, whereas He values were equal to 0.74, 0.74, 0.75 and 0.81 for the Liguro-Provençal, the Corsican, the Catalonian and the Balearic cluster, respectively. Only the pairwise comparisons involving the Balearic cluster were significant, indicating a higher level of genetic diversity for this cluster.

At the sample level, gene diversity ranged from 0.62 for MON to 0.83 for RIO. Ar(*22*) and Ap(*22*) values were scattered from 5.15 for CAR to 9.48 for FOR (Table 3) and from 0 for RIE and PZU to 1 for FOR. Ar(*22*) was highly correlated with gene diversity (*R*-spearman: 0.86; dof: 38; $P < 0.05$; $R^2 = 0.74$) but not with Ap(*22*) (*R*-spearman: 0.23; dof: 38; $P = 0.1426$; $R^2 = 0.05$).

Based on the Wilcoxon signed rank-test, recent changes in the effective population size were detected. Sixteen samples (Table 3) showed significant heterozygote deficiency compared to the mutation-drift equilibrium, revealing a putative population expansion (see Discussion).

Discussion

Distribution of genetic diversity from global to sample scale with a focus on the Balearic region

Genetic diversity was evaluated for the first time at three different levels: global, regional (corresponding to the four clusters defined with STRUCTURE) and sample levels. As a whole, the levels of gene diversity were high and they were globally higher than those reported previously for the same species (Costantini *et al*. 2007b) (mean He by samples over loci: 0.77 vs. 0.50 and mean He by samples: 0.89 vs. 0.78). This discrepancy is probably due to differences in the number and identity of loci, since the two loci used in both studies show similar diversity levels. However, the present values still fall within the range of those previously reported for octocorallian species [e.g. *Corallium lauuense*, He from 0.86–0.96 (Baco & Shank 2005)] or other cnidarian species [e.g. *Acropora palmata*, He from 0.58–0.85 (Baums *et al.* 2005)].

This high genetic diversity is not homogenously shared, since only 16.8% of alleles were found in the four clusters. This result was refined based on the large difference observed between global Arst(*g*) and F_{ST} values, highlighting the fact that rare alleles tend to be clustered only in some samples (Comps *et al.* 2001). At the regional level, the particular status of the Balearic region previously suggested (Costantini *et al*., 2007b) was statistically confirmed with He and Ar(*66*) being significantly higher for this area. Moreover, Ap(*66*) was found to be two times higher for this cluster when compared to the other clusters. Based on these parameters, the three remaining regions (NW Corsica, Liguro-Provençal and Catalonia) were not found to be significantly different from each other. Interestingly, Rozenfeld *et al.* (2008) recently emphasized the importance of the Balearic populations of *Posidonia oceanica* as genetic suppliers for the surrounding populations. Two non-exclusive hypotheses can be formulated to explain the high genetic diversity observed for the Balearic populations of *P. oceanica* and *C. rubrum*, despite their divergent life history traits. The first hypothesis is based on contemporary oceanographic observations, by which the Balearic region is defined as a transition region between the Liguro-Provençal and the Algerian basin, each characterized by two contrasting dynamic regimes that induce highly variable hydrological conditions (García *et al.* 1994). These conditions could have led to the frequent mixing of different gene pools from the south and the north of the islands, contributing to the high level of observed genetic diversity (Rozenfeld *et al.* 2008). The second hypothesis is based on the observation that colonization processes are usually followed by an important decrease in genetic variability away from refugia (Hewitt 2000). Despite the absence of data concerning the past distribution of *C. rubrum* populations, we may hypothesize that the significantly higher genetic diversity observed for the Balearic cluster could result from the geographic proximity between the Balearic Islands and some putative glacial refugia, when compared to other regions under study.

Heterogeneity in patterns of diversity is also reported at the sample level. This heterogeneity is not linked to depth or to the habitats of the samples, suggesting that these factors do not influence the levels of genetic diversity (data not shown). Although the allelic and private allelic richness are not correlated, the highest values were observed for two samples coming from the Balearic region. The independence between these two parameters underscores their complementarities and the necessity to estimate both of them. Moreover, focusing more precisely on these two parameters is essential to ensure that the widest range of existing alleles is available for future evolution (Bonin *et al.* 2007; also see below). Finally, the bottleneck analysis detected signs of population expansion (i.e. a population recovering from

a past bottleneck) for 16 samples. However, the Wahlund effect could mimic population expansion (Cornuet & Luikart 1996; Dupont *et al.* 2009). This confounding effect could therefore partially explain the observed results, since five of these 16 samples (XIC, RIE, GCO, PUA and CRO) showed significant LD putatively linked to the Wahlund effect (Slatkin 2008). The absence of significant signal of population reduction also needs to be treated with caution. Indeed, inertia in the loss of genetic diversity inducing inaccurate inference on recent demographic events has been observed in the case of long-lived species (Kuo & Janzen 2004; Lippé *et al.* 2006) or when IBD structure occurred (Leblois *et al.* 2006).

Population relationships

The influence of geographical proximity on genetic relationships was demonstrated by the IBD observed over the surveyed area. This result refutes the hypothesis of an island model of dispersal previously proposed (Costantini *et al.* 2007b) and suggests the occurrence of migration-drift equilibrium among neighbouring populations (Hutchison & Templeton 1999). However, despite this clinal genetic variation that could be problematic for STRUCTURE (Guillot & Santos 2009; also see Rosenberg *et al.* 2005; Chen *et al.* 2007), genetic discontinuities were revealed, putatively explaining the relatively low R^2 value (0.15) of the regression model. Indeed, the first round of clustering revealed a structure pattern that followed the geographical partitioning of the samples and potentially also the distribution of suitable habitats (e.g. bathymetric barrier between Corsican and Balearic clusters or absence of rocky habitats between Catalonian and Liguro-Provençal clusters). The AMOVA results confirmed that the differentiation among-clusters was significant, but it was low compared to the differentiation among-samples within clusters. Two non-exclusive hypotheses could be proposed to explain this result. Microsatellites are susceptible to size homoplasy (SH) that could have consequences on inferred relationships among populations (Viard *et al.* 1998). Effective population size, mutation rates and models and divergence time between-populations, all affect SH, but most of these parameters are unknown. However, the impact of SH on assignment methods is negligible when the markers used show a high level of heterozygosity (Estoup *et al.* 2002), as observed with our dataset.

The weak regional genetic structure may also be explained by a putative inter-regional gene flow preventing divergence by genetic drift. However, this is unlikely considering the IBD revealed over the surveyed area. Alternatively, this structure could be due to the retention of an ancestral polymorphism that could be enhanced by longevity (Lippé *et al.* 2006). A shared ancestral polymorphism implies a recent split between regional-clusters and/or large effective population sizes. High heterozygosity values and large amounts of null alleles could be interpreted as cues for a large effective population size (Chapuis & Estoup 2007; Hellberg 2007), but analyses of temporal samples are needed to confirm this parameter. Regarding the recent divergence hypothesis, the sea level was around 120 m lower than it is today during the last glacial maximum (LGM, 21 000 BP) (Lambeck & Bard 2000). The recolonization of new habitats was mostly done gradually because of the observed IBD. *C. rubrum* is estimated to reach sexual maturity at approximately 10 years of age (Torrents *et al.* 2005). However, at least 20 years are needed to reach high reproductive potential (Garrabou & Harmelin 2002; Santangelo *et al.* 2003). These parameters therefore imply that the recolonization occurred within a maximum of 2000 generations, highlighting the relatively recent origin of these shallow populations, and potentially explaining the low genetic divergence observed between regional-clusters.

The within cluster analysis emphasized the different structure patterns among the three studied regions. The subdivision of samples was indeed stronger for the Corsican than for the Liguro-Provençal and the Catalonian clusters, with, respectively, 7.9%, 7.8% and 4.2% of the total genetic variation observed among samples. This is in agreement with the stronger mean assignment of individuals for the Corsican and the Liguro-Provençal clusters than for the Catalonian cluster during the first round of clustering analysis (data not shown). Based on the IBD results, these differences between the three clusters could be explained in terms of their concordance with the regional drift-gene flow equilibrium (Hutchison & Templeton 1999). The absence of significant IBD in the Corsican cluster ($r = 0.168$; $P = 0.39$) and the wide range of the scatterplot observed for the regression model suggest a stronger impact of genetic drift in this cluster compared to the other ones, where IBD is observed putatively due to regional drift-gene flow equilibrium. As previously shown, despite these clinal variations, STRUCTURE results suggest a putative barrier to gene flow between the north and the south of the Cap de Creus in the Catalonian region. Regarding the Liguro-Provençal cluster, the eastern (PDA, MON/VIL) and CAR samples were isolated from the remaining ones, indicating putative barriers to gene flow.

Evolutionary hypothesis and consequences for conservation and management plans

The results presented above allow us to suggest an evolutionary scenario to explain the observed pattern of

nuclear genetic diversity. Due to restricted dispersal abilities and IBD, we argue that recolonization from LGM refugia probably occurred mainly following a gradual process. Regarding the generation time of the species, this process is likely to have been recent, potentially inducing the retention of ancestral polymorphism and the weak regional structure observed. At the same time, the differential reproductive success (Costantini *et al.* 2007a), combined with the localized dispersal, could induce small-scale genetic structure coupled with high heterogeneity in the distribution of nuclear genetic diversity. Small-scale studies with known spatial locations of individuals, designed to help define breeding units (e.g. Underwood *et al.* 2007) and specific samplings of cohorts to evaluate the effective size of red coral populations (e.g. Calderón *et al.* 2009) are needed to test these different hypotheses. Analyses of populations from the Algerian basin and from deep habitats, such as below sea level during the LGM, should also help to refine this evolutionary history.

The results of the present study provide information that could significantly enhance the management of shallow populations of *C. rubrum.* Because of their slow dynamics, the recovery of red coral populations from environmental disturbances should be measured in decades (Garrabou *et al.* 2001). As a complement, the strong genetic differentiation between nearby samples implies that this recovery should be mainly due to self-recruitment. The genetic diversity results obtained at different spatial scales correspond to a baseline for monitoring genetic consequences of actual global changes (Schwartz *et al.* 2007) on shallow populations of *C. rubrum*. They also represent a new step in the conservation of the red coral, since they allow the targeting of particular samples that need to be protected as putative genetic pool. In fact, despite the controversy about the role of genetic factors in species extinction (e.g. Lande 1988), genetic diversity is one of the three levels of biodiversity that need attention in regards to conservation (McNeely *et al.* 1990). Genetic diversity is indeed required for populations to evolve in response to environmental changes (Frankham 2005). Conservation of the putative adaptive diversity and evolutionary potential require preservation of the natural network of genetic connections existing between populations (Crandall *et al.* 2000). The combined action of climate changes and harvesting could induce selective genetic changes (Allendorf *et al.* 2008) and the extinction of local populations in the near future (Garrabou *et al.* 2009), resulting in the loss of genetic diversity and the isolation of populations by disrupting the IBD pattern. Protection and management plans for shallow red coral populations should therefore concomitantly address measures at local and global scales.

The first step should be the development of a marine reserve network based on the restricted dispersal abilities of the species (Palumbi 2004) and focused on areas with high density of populations or high genetic diversity, such as, west of the Liguro-Provençal region, including RIO IML and CAS or Balearic Islands, including LPP and FOR. Restoration programmes could also be planned to help the recovery of local populations (Baums 2008; Linares *et al.* 2008). Bearing in mind the restricted dispersal ability of red coral, these restoration programs must consider putatively strong local adaptation (Lenormand 2002) to environmental parameters, such as depth or temperature (Hughes *et al.* 2003; Torrents *et al.* 2008). Resistance to increasing temperatures could be taken into account to choose source populations for restoration projects. Finally, scenarios available for the Mediterranean basin for the 21st century predict a dramatic increase in the frequency of heat waves (Diffenbaugh *et al.* 2007; IPCC 2007), which will particularly affect shallow populations of *Corallium rubrum*. Considering the ongoing warming, global conservation actions must be focused on the reduction of harvesting pressure. The necessity to promote large areas of harvesting exclusion on shallow habitats (0–60 m) must therefore be seriously considered through international concerted actions.

Acknowledgements

We are grateful to J. M. Dominici (Réserve Naturelle de Scandola, Parc Régional de Corse), R. Graille, B. de Ligondes, C. Marschal, O. Torrents, and F. Zuberer (Centre d'Océanologie de Marseille) for the assistance during sampling. We are also grateful to J. Coll, C. Linares, N. Teixido and D. Diaz for providing samples from the Medes and the Balearic Islands. We thank C. Hubert, A. Chenuil and C. Rocher for helpful discussion concerning applications and analyses of microsatellite loci. The comments made by Stephen R. Palumbi, Cécile Fauvelot and three anonymous reviewers greatly improved the original manuscript. This study was funded by the 'Provence–Alpes–Côte d'Azur' region (J-B Ledoux Doctoral Fellowship) and by the IFB Foundation (Institut Français de la Biodiversité), the environmental department of TOTAL-FINA-ELF group, an 'Action Concertée Initiative' (ACI) program 'Jeunes chercheurs 2003' (number JC5360) from the French Research Ministry and the MEDCHANGE project funded by the Agence Nationale pour la Recherche (ANR). It is also a contribution to the responsive mode project GBIRM (European Network of Excellence MARBEF).

References

Abbiati M, Santangelo G, Novelli S (1993) Genetic variation within and between two Tyrrhenian populations of the Mediterranean alcyonarian *Corallium rubrum*. *Marine Ecological Progress Series*, **95**, 245–250.

Airoldi L, Beck M (2007) Loss, status and trends for coastal marine habitats of Europe. *Oceanography and Marine Biology: an annual review*, **45**, 345–405.

Allendorf FW, England PR, Luikart G, Ritchie PA, Ryman N (2008) Genetic effects of harvest on wild animal populations. *Trends in Ecology and Evolution*, **23**, 327–337.

Baco AR, Shank TM (2005) Population Genetic Structure of the Hawaiian Precious Coral Corallium Lauuense (Octocorallia: Coralliidae) Using MicrosatellitesIn: *Cold-Water Corals and Ecosystems* (eds Freiwald A, Roberts JM), pp. 663–678. Springer-Verlag, Berlin.

Ballesteros E (2006) Mediterranean coralligenous assemblages: a synthesis of present knowledge. *Oceanography and Marine Biology: An Annual Review*, **44**, 123–195.

Baums IB (2008) A restoration genetics guide for coral reef conservation. *Molecular Ecology*, **17**, 2796–2811.

Baums IB, Miller MW, Hellberg ME (2005) Regionally isolated populations of an imperiled Caribbean coral, *Acropora palmata*. *Molecular Ecology*, **14**, 1377–1390.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate – a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, **57**, 289–300.

Bonin A, Nicole F, Pompanon F, Miaud C, Taberlet P (2007) Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology*, **21**, 697–708.

Calderón I, Garrabou J, Aurelle D (2006) Evaluation of the utility of COI and ITS markers as tools for population genetic studies of temperate gorgonians. *Journal of Experimental Marine Biology and Ecology*, **336**, 184–197.

Calderón I, Palacin C, Turon X (2009) Microsatellite markers reveal shallow genetic differentiation between cohorts of the common sea urchin Paracentrotus lividus (Lamarck) in northwest Mediterranean. *Molecular Ecology*, **18**, 3036–3049.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology Evolution*, **24**, 621–631.

Chen C, Durand E, Forbes F, Francois O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes*, **7**, 747–756.

Comps B, Gomory D, Letouzey J, Thiebaut B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics*, **157**, 389–397.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Costantini F, Abbiati M (2006) Development of microsatellite markers for the Mediterranean gorgonian coral *Corallium rubrum*. *Molecular Ecology Notes*, **6**, 521–523.

Costantini F, Fauvelot C, Abbiati M (2007a) Fine-scale genetic structuring in *Corallium rubrum*: evidence of inbreeding and limited effective larval dispersal. *Marine Ecology Progress Series*, **340**, 110–119.

Costantini F, Fauvelot C, Abbiati M (2007b) Genetic structuring of the temperate gorgonian coral (*Corallium rubrum*) across the western Mediterranean Sea revealed by microsatellites and nuclear sequences. *Molecular Ecology*, **16**, 5168–5182.

Costanza R, d'Arge R, Groot Rd, *et al*. (1997) The value of the world's ecosystem services and natural capital. *Nature*, **387**, 253–260.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, **15**, 290–295.

Del Gaudio D, Fortunato G, Borriello M, *et al.* (2004) Genetic typing of *Corallium rubrum*. *Marine Biotechnology*, **6**, 511–515.

Dempster A, Laird N, Rubin D (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B*, **39**, 1–38.

Di Rienzo A, Peterson A, Garcza J, *et al.* (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 3166–3170.

Diffenbaugh N, Pal J, Giorgi F, Gao X (2007) Heat stress intensification in the Mediterranean climate change hotspot. *Geophysical Research Letters*, **34**, 1–6.

Dupont L, Viard F, Dowell MJ, Wood C, Bishop JDD (2009) Fine- and regional-scale genetic structure of the exotic ascidian *Styela clava* (Tunicata) in southwest England, 50 years after its introduction. *Molecular Ecology*, **18**, 442–453.

Estoup A, Jarne P, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, **11**, 1591–1604.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Ferrière R, Dieckmann U, Couvet D (2004) *Evolutionary Conservation Biology*. Cambridge University Press, Cambridge, UK.

Frankham R (2005) Genetics and extinction. *Biological Conservation*, **126**, 131–140.

García MJL, Millot C, Font J, García-Ladona E (1994) Surface circulation variability in the Balearic Basin. *Journal of Geophysical Research Ocean*, **99**, 3285–3296.

Garrabou J, Coma R, Bensoussan N, *et al*. (2009) Mass mortality in NW Mediterranean rocky benthic communities: effects of the 2003 heat wave. *Global Change Biology*, **15**, 1090–1103.

Garrabou J, Harmelin JG (2002) A 20-year study on life-history traits of a harvested long-lived temperate coral in the NW Mediterranean: insights into conservation and management needs. *Journal of Animal Ecology*, **71**, 966–978.

Garrabou J, Perez T, Sartoretto S, Harmelin JG (2001) Mass mortality event in red coral *Corallium rubrum* populations in the Provence region (France, NW Mediterranean). *Marine Ecological Progress Series*, **217**, 263–272.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Guillot G, Santos F (2009) A computer program to simulate multilocus genotype data with spatially autocorrelated allele frequencies. *Molecular Ecology Resources*, **9**, 1112–1120.

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

Halpern BS, Walbridge S, Selkoe KA, *et al.* (2008) A global map of human impact on marine ecosystems. *Science*, **319**, 948–952.

Harley CDG, Hughes AR, Hultgren KM, *et al.* (2006) The impacts of climate change in coastal marine systems. *Ecology Letters*, **9**, 228–241.

Hellberg ME (2007) Footprints on water: the genetic wake of dispersal among reefs. *Coral Reefs*, **26**, 463–473.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hughes TP, Baird AH, Bellwood DR, *et al.* (2003) Climate change, human impacts and the resilience of coral reefs. *Science*, **301**, 929–933.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898–1914.

IPCC (2007) Climate change 2007: the physical science basis. Contribution of working group I to the Fourth Assesment. Report of the intergovernmental panel on climate change (eds Solomon SD, Qin M, Manning Z, *et al.*), 996 pp. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Jackson JBC (2008) Ecological extinction and evolution in the brave new ocean. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 11458–11465.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jakobsson M, Scholz SW, Scheet P, *et al.* (2008) Genotype, haplotype and copy-number variation in worldwide human populations. *Nature*, **451**, 998–1002.

Kuo C, Janzen F (2004) Genetic effects of a persistent bottleneck on a natural population of ornate box turtles (*Terrapene ornata*). *Conservation Genetics*, **5**, 425–437.

Laborel J, Vacelet J (1961) Répartition bionomique du *Corallium rubrum* LMCK dans les grottes et falaises sous-marines. *Rapports et Procès-Verbaux des Réunions de la Commission Internationale pour l'Exploration Scientifique de la Mer Méditerranée*, **16**, 464–469.

Lambeck K, Bard E (2000) Sea-level change along the French Mediterranean coast for the past 30 000 years. *Earth and Planetary Science Letters*, **175**, 203–222.

Lande R (1988) Genetics and demography in biological conservation. *Science*, **241**, 1455–1460.

Leblois R, Estoup A, Streiff R (2006) Genetics of recent habitat contraction and reduction in population size: does isolation by distance matter? *Molecular Ecology*, **15**, 3601–3615.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, **17**, 183–189.

Linares C, Coma R, Zabala M (2008) Restoration of threatened red gorgonian populations: an experimental and modelling approach. *Biological Conservation*, **141**, 427–437.

Lippé C, Dumont P, Bernatchez L (2006) High genetic diversity and no inbreeding in the endangered copper redhorse, Moxostoma hubbsi (Catostomidae, Pisces): the positive sides of a long generation time. *Molecular Ecology*, **15**, 1769–1780.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer research*, **27**, 209–220.

Marschal C, Garrabou J, Harmelin JG, Pichon M (2004) A new method for measuring growth and age in the precious red coral *Corallium rubrum* (L.). *Coral Reefs*, **23**, 423–432.

McFadden CS, France SC, Sánchez JA, Alderslade P (2006) A molecular phylogenetic analysis of the Octocorallia (Cnidaria: Anthozoa) based on mitochondrial protein-coding sequences. *Molecular Phylogenetics and Evolution*, **41**, 513–527.

McNeely J, Miller K, Reid W, Mittermeier R, Werner T (1990) Strategies for conserving biodiversity. *Environment*, **32**, 16–20.

Millot C (1999) Circulation in the Western Mediterranean Sea. *Journal of Marine Systems*, **20**, 423–442.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences the United States of America*, **70**, 3321–3323.

Palumbi SR (2004) Marine reserves and ocean neighborhoods: the spatial scale of marine populations and their management. *Annual Review of Environmental Resources*, **29**, 31–68.

Perez T, Garrabou J, Sartoretto S, *et al.* (2000) Mortalité massive d'invertébrés marins: un événement sans précédent en Méditerranée nord-occidentale – mass mortality of marine invertebrates on the Provence littoral (France): an unprecedented event in the NW Mediterranean. *Comptes Rendus de l'Académie des Sciences de Paris, Série III*, **323**, 853–865.

Petit RJ, Mousadik AE, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Pont-Kingdon GA, Okada NA, Macfarlane JL, *et al.* (1995) A coral mitochondrial mutS gene. *Nature*, **375**, 109–111.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Pritchard JK, Wen X, Falush D (2007) *Documentation for the STRUCTURE software, Version 2*. Chicago. Available at http://pritch.bds.uchicago.edu.

Raymond M, Rousset F (1995) GENEPOP (ver. 1.2): a population genetics software for exact test and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Romano J-C, Bensoussan N, Younes WAN, Arlhac D (2000) Anomalie thermique dans les eaux du golfe de Marseille durant l'été 1999. Une explication partielle de la mortalité d'invertébrés fixés ? *Comptes Rendus de l'Académie des Sciences Série III*, **323**, 415–427.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rosenberg NA, Mahajan S, Ramachandran S, *et al.* (2005) Clines, clusters, and the effect of study design on the inference of human population structure. *Plos Genetics*, **1**, 660–671.

Rosenberg NA, Pritchard JK, Weber JL, *et al.* (2002) Genetic structure of human populations. *Science*, **298**, 2381–2385.

Rousset F (1997) Genetic differentiation and estimation of gene flow from FStatistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Rozenfeld AF, Arnaud-Haond S, Herna'ndez-García E, *et al.* (2008) Network analysis identifies weak and strong links in a metapopulation system. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 18824–18829.

Sambrook J, Fritsch E, Maniatis T (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory Press, New York.

Santangelo G, Abbiati M (2001) Red coral: conservation and management of an over-exploited Mediterranean species. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **11**, 253–259.

Santangelo G, Bramanti L, Iannelli M (2007) Population dynamics and conservation biology of the over-exploited Mediterranean red coral. *Journal of Theoretical Biology*, **244**, 416–423.

Santangelo G, Carletti E, Maggi E, Bramanti L (2003) Reproduction and population sexual structure of the overexploited Mediterranean red coral *Corallium rubrum*. *Marine Ecological Progress Series*, **248**, 99–108.

Santangelo G, Maggi E, Bramanti L, Bongiorni L (2004) Demography of the over-exploited Mediterranean red Coral (*Corallium rubrum* L. 1758). *Scientia Marina*, **68**(Suppl. 1), 199–204.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, **22**, 25–33.

Slatkin M (2008) Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nature Review Genetics*, **9**, 477–485.

Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **27**, 2498–2504.

Torrents O (2007) *Biologie des populations du corail rouge Corallium rubrum (L. 1758) de Méditerranée nord-occidentale*, PhD Dissertation. Université de la Méditerranée, France.

Torrents O, Garrabou J, Marschal C, Harmelin JG (2005) Age and size at first reproduction in the commercially exploited red coral *Corallium rubrum* (L.) in the Marseilles area (France, NW Mediterranean). *Biological Conservation*, **121**, 391–397.

Torrents O, Tambutté E, Caminiti N, Garrabou J (2008) Upper thermal thresholds of shallow vs. deep populations of the precious. *Journal of Experimental Marine Biology and Ecology*, **357**, 7–19.

Underwood JN, Smith LD, Oppen MJHV, Gilmour JP (2007) Multiple scales of genetic connectivity in a brooding coral on isolated reefs following catastrophic bleaching. *Molecular Ecology*, **16**, 771–784.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **6**, 255–256.

Viard F, Franck P, Dubois M-P, Estoup A, Jarne P (1998) Variation of microsatellite size homoplasy across electromorphs, loci, and populations in three invertebrate species. *Journal of Molecular Evolution*, **47**, 42–51.

Vighi M (1972) Etude sur la reproduction du *Corallium rubrum* (L.). *Vie et milieu*, **23**, 21–32.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419–1439.

Weinberg S (1979) The light-dependent behaviour of planula larvae of *Eunicella singularis* and *Corallium rubrum* and its implication for octocorallian ecology. *Bijdragen tot de Dierkunde*, **49**, 16–30.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Zibrowius H, Monteiro-Marques V, Grasshoff M (1984) La répartition du *Corallium rubrum* dans l'Atlantique (Cnidaria: Anthozoa: Gorgonaria). *Téthys*, **11**, 163–170.

This article is a part of the PhD thesis of J. -B. Ledoux. J. -B. Ledoux's current research focuses on the application of genetic and evolutionary approaches to unravel the population structure and local adaptation processes in *Corallium rubrum*. K. Mokhtar-Jamaï is a PhD student interested in population genetics and phylogeography of gorgonians. C. Roby is a PhD student interested in marine ecological and evolutionary process such as hybridization in marine organisms. J. -P. Féral is a researcher interested in ecology and evolution of marine biodiversity. J. Garrabou is a researcher involved in marine population and community studies focused in the conservation of long-lived invertebrate species. D. Aurelle is an associate professor interested in evolutionary biology and population genetics.

Supporting information

Additional supporting information may be found in the online version of this article.

Document S1 Development of microsatellite loci.

Table S1 Impact of heterozygote deficiency on STRUCTURE outputs

Table S2 Multilocus and monolocus values of F_{IS} estimator (f) computed for each samples

Tables S3 a and b Pairwise F_{ST} values

Fig. S1 STRUCTURE plots.

Genetic survey of shallow populations of the Mediterranean red coral (*Corallium rubrum* (Linnaeus, 1758)): new insights into evolutionary processes shaping nuclear diversity and implications for conservation.

J-B Ledoux, K. Mokhtar-Jamaï, C. Roby, J-P Féral, J. Garrabou, D. Aurelle

SUPPLEMENTARY MATERIAL

The following issues are addressed:

- **Development of microsatellite loci.**
- **Impact of heterozygote deficiency on STRUCTURE outputs (Table S1).**
- **Multilocus and monolocus values of Fis estimator (f) computed for each samples: Table S2.**
- **Pairwise Fst values: Table S3 a and b.**
- **STRUCTURE plots: Figure S1.**

Development of microsatellite loci:

Six loci were isolated by Ecogenics GmbH (Zurich, Switzerland) from an enriched library. Size-selected genomic DNA was ligated into SAULA/SAULB-linker and enriched by magnetic bead selection with biotin-labeled $(GT)_{13}$ and $(GATA)_8$ oligonucleotide repeats (Gautschi *et al.* 2000a; Gautschi *et al.* 2000b). Of 948 recombinant colonies screened, 150 gave a positive signal after hybridization. Plasmids from 115 positive clones were sequenced and primers were designed for 22 microsatellite inserts. Of these, twelve were tested for polymorphisms and migration patterns. Finally six were retained: Mic22, Mic23, Mic24, Mic25, Mic26 and Mic27. A second enriched library for the common repeat CA was

purchased from Bioprofiles LTD (Newcastle University). Plasmids from 40 positive clones were sequenced. Specific primers were designed for 12 microsatellites using Primer3 v 0.4.0 (Rozen & Skaletsky 2000 at http://frodo.wi.mit.edu/). After tests of electrophoresis patterns and polymorphism levels of the 12 loci, only two were retained: Mic13, Mic20. Finally, we tested the five loci from Costantini & Abbiati (2006). Due to the high level of null alleles of these microsatellites (Costantini *et al.* 2007a, b), we designed new primer pairs using Primer3 v 0.4.0 for the 5 loci. After amplification tests, two new primer pairs were retained for population analysis: COR9bis and COR46bis, corresponding respectively to loci COR9 and COR46 of Costantini *et al.* (2007 a, b). Whereas no significant change was observed between results obtained using the two primer pairs for the COR9 locus (r around 0.4 in both cases), the use of COR46bis resulted in an important decrease in null allele frequency (0.05 vs. 0.4). Ten new microsatellite loci (Mic13, Mic20, Mic22, Mic23, Mic24, Mic25, Mic26, Mic27, COR9bis and COR46bis) were therefore isolated and used for population genetics analyses of *Corallium rubrum*.

Impact of heterozygote deficiency on STRUCTURE outputs:

Simulated dataset:

EASYPOP version 2.0.1 (Balloux 2001) was used to simulate datasets showing variable levels of heterozygosity deficiency. Four populations were simulated under hierarchical island models (2 archipelagos of 2 populations) with a combination of different levels of differentiation and inbreeding (Table S2) leading to five different scenarios. For each scenario, five replicates were used. In every case, the number of individuals in each population was set to 1000, and ten loci were simulated with the uniform Kam model of evolution and a mutation rate equal to 0.0005 (Estoup & Angers 1998). As in our dataset, the

total number of alleles was high (350), with maximal variability for the initial population. Five thousand generations were simulated to reach equilibrium between drift, gene flow and mutation. Finally, 30 individuals were sampled in each of the 4 populations. Levels of gene flow between populations that would result in differentiation between populations on the same order of magnitude as those observed in the real dataset were chosen (pairwise Fst from 0.004 to 0.3). A combination of inbreeding and null alleles was used to simulate the origin of heterozygote deficiencies. Three different levels of selfing were used (0.1, 0.25 and 0.5) to simulate reproduction between relatives. Null alleles were simulated following Carlsson (2008). Two alleles by locus were randomly chosen and deleted from the dataset. Consequently, heterozygotes carrying the null allele became homozygotes for the alternative allele, and homozygotes for the null allele became null homozygotes. Null allele frequencies and Fis values ranged from 0.02 to 0.24 and from 0.02 to 0.5, respectively, as in our dataset. STRUCTURE was run under admixture with or without separated alpha values between clusters and with the correlated allele frequency models. For each simulated dataset, a first round was conducted with five runs of 100000 burn-in followed by 250000 iterations for each K value. K was set to vary from 1 to 5. In cases where the solution of K=4 was not found, a second round of STRUCTURE was conducted on a partitioned dataset based on the most likely solution of the preceding round.

Results:

Table S2 summarizes results of the simulation study. Two models (admixture with correlated allele frequencies with the same alpha or with separated alphas) gave good results despite HW equilibrium not being met for all samples. In some cases and for both models, it was necessary to use a partitioned dataset obtained after the first round of STRUCTURE to reach a K value of 4. The first round distinguished between the two archipelagos, whereas the

second round differentiated between the two populations inside each archipelago. For the 5th scenario, implying the smallest Fst values (mean value over the 5 replicates Fst = 0.06), none of the 2 models was able to find the solution K=4. Only the two archipelagos were identified in every case (mean Fst values between archipelago = 0.08), whereas intra-archipelago clusters were not defined (mean Fst value between population inside archipelagos = 0.01). Therefore, considering these results, a model with admixture allowing for separated alpha values between clusters and for correlated allele frequencies among them was chosen.

REFERENCES:

Balloux F (2001) EASYPOP (version 1.7),A computer program for the simulation of population genetics. *Journal of Heredity* **92**, 301-302.

Carlsson J (2008) Effects of Microsatellite Null Alleles on Assignment Testing. *Journal of Heredity* **99**, 616-623.

Costantini F, Abbiati M (2006) Development of microsatellite markers for the Mediterranean gorgonian coral Corallium rubrum. *Molecular Ecology Notes* **6**, 521-523

Costantini F, Fauvelot C, Abbiati M (2007a) Fine-scale genetic structuring in Corallium rubrum: evidence of inbreeding and limited effective larval dispersal. *Marine Ecology Progress Series* **340**, 110-119.

Costantini F, Fauvelot C, Abbiati M (2007b) Genetic structuring of the temperate gorgonian coral (Corallium rubrum) across the western Mediterranean Sea revealed by microsatellites and nuclear sequences. *Molecular Ecology* **16**, 5168-5182.

Estoup A, Angers B (1998) Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. *pp. 55-86 in Advances in Molecular Ecology, edited by G. Carvalho. IOS Press, Amsterdam.*

Gautschi B, Tenzer I, Müller JP, Schmid B (2000a) Isolation and characterization of microsatellite loci in the bearded vulture (Gypaetus barbatus) and crossamplification in three Old World vulture species. *Molecular Ecology* **9**, 2193-2195.

Gautschi B, Widmer A, Koella J (2000b) Isolation and characterization of microsatellite loci in the Dice Snake (Natrix tessellata). *Molecular Ecology* **9**, 2191-2193.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.*

TABLE AND FIGURE LEGENDS:

Table S1: STRUCTURE efficiency in cases of heterozygote deficiency: characteristics of the five simulated scenario and associated results. For each scenario, five replicate datasets were simulated using EASYPOP (I to V). r: null allele frequency; f: Weir and Cockerham (1984) estimator of Fis. Final clustering results are shown for the two models: admixture with correlated allele frequency between clusters with the same or separated alphas among clusters. Each column indicates the number K of clusters obtained for a first round of analyses. The values in brackets indicate the number of clusters identified when a second round of analyses was performed.

Table S2: Multilocus and monolocus values of Fis estimator (f) computed for each samples. Values in bold are considered significant at a 0.05 level after FDR correction.

Table S3a: Pairwise Fst values, among which 789 are significant after FDR correction for multiple tests at 0.05 (bold = non-significant value) are shown above the diagonal. Pairwise Euclidian distances are shown below the diagonal (b). A-Provenço-ligurian region; B-Corsican region; C-Balearic Islands; D-Catalonian region.

Table S3b: Pairwise Fst values among which 789 are significant after FDR correction for multiple tests at 0.05 (bold = non-significant value) are shown above the diagonal. Pairwise geographical distance taking into account major surface current between regions are shown below the diagonal. A-Provenço-ligurian region; B-Corsican region; C-Balearic Islands; D-Catalonian region.

Figure S1: Logarithm of the likelihood values as a function of the number of cluster K. Likelihood values are the result of 10 runs for each K value. A) Results from the first round (with whole dataset) of STRUCTURE are shown. B) Results from the second round of STRUCTURE using a partitioned dataset defined previously with K=4.

Table S1:

Simulated Datasets									Results	
Scenario	Selfing rate	Migration rate between archipelagos	Migration rate within archipelagos	Mean r	Mean f	Mean Fst within archipelagos	Mean Fst between archipelagos	Replicats	Admixture correlated same alpha	Admixture correlated separated alpha
I	0.1	0.00001	0.001	0.16	0.21	0.1	0.22	1	2 (4)	2 (4)
								2	2 (4)	2 (4)
								3	2 (4)	2 (4)
								4	4	4
								5	2 (4)	2 (4)
II			0.0005	0.14	0.21	0.15	0.25	1	4	4
								2	4	4
								3	4	4
								4	2 (4)	4
								5	4	4
III	0.5	0.00001	0.001	0.17	0.42	0.13	0.28	1	4	4
								2	2 (4)	2 (4)
								3	4	2 (4)
								4	4	2 (4)
								5	4	4
IV			0.0005	0.18	0.43	0.18	0.29	1	4	4
								2	4	4
								3	4	4
								4	4	4
								5	4	4
V	0.25	0.0006	0.01	0.13	0.34	0.02	0.09	1	2	2
								2	2	2
								3	2	2
								4	2	2
								5	2	2

Table S2:

Samples Names	Multilocus f values	Monolocus f values									
		Mic13	Mic20	Mic22	Mic23	Mic24	Mic25	Mic26	Mic27	COR9bis	COR46bis
PDA	**0.2**	**1**	-0.3	-0.13	**0.45**	-0.07	**0.4**	-0.02	**0.13**	**0.48**	-0.1
MON	**0.26**	**1**	0.12	**0.66**	**0.41**	0.17	0.13	0.17	-0.02	**0.57**	0.04
VIL	**0.31**	**0.75**	0.02	**0.49**	**0.56**	0.05	**0.66**	0.02	0.07	**0.65**	0.11
TRM	**0.37**	**1**	0.13	**0.54**	**0.73**	-0.06	**0.58**	0.1	0.04	**0.87**	0.08
CAS	**0.28**	**0.62**	0.06	**0.31**	**0.26**	-0.13	**0.57**	0.03	**0.19**	**0.79**	0.2
FIE	**0.36**	**0.43**	-0.09	**0.38**	**0.83**	0.07	**0.6**	0.05	**0.47**	**0.71**	-0.02
FII	**0.41**	**0.79**	0.23	**0.59**	**0.55**	0	**0.53**	0.03	**0.49**	**0.69**	0.19
GC	**0.31**	**0.92**	0	**0.55**	**0.72**	-0.03	**0.32**	0.02	0.11	**0.67**	0.1
GRC	**0.47**	**0.94**	0.53	**0.94**	**0.51**	0.11	**0.57**	0.08	**0.41**	**0.72**	0.02
IML	**0.38**	**0.91**	0.1	**0.66**	**0.82**	0.11	**0.51**	0.01	0	**0.73**	0.05
RIO	**0.41**	**0.84**	0	**0.8**	**0.93**	0.08	**0.6**	-0.05	**0.12**	**0.76**	-0.01
RIE	**0.21**	**0.92**	-0.28	**0.66**	**0.81**	-0.24	**0.51**	-0.05	**0.21**	**0.48**	-0.12
RII	**0.35**	**0.72**	-0.11	**0.76**	**0.71**	0.09	**0.5**	-0.04	**0.16**	**0.67**	-0.01
PLA	**0.37**	**0.82**	0.13	**0.51**	**0.72**	-0.05	**0.9**	0	**0.27**	**0.65**	0.01
GPE	**0.37**	**0.54**	-0.1	**0.71**	**0.83**	0.05	**0.8**	-0.03	0.07	**0.77**	0.06
GPI	**0.35**	**0.91**	0.12	**0.38**	**0.31**	0.13	**0.71**	-0.08	**0.47**	**0.68**	0.06
PHA	**0.32**	**0.88**	0.24	**0.41**	**0.8**	0.09	**0.72**	-0.02	**0.14**	**0.5**	0.11
TFR	**0.33**	**0.66**	0	**0.68**	**0.53**	-0.05	**0.8**	-0.02	0.04	**0.83**	0.06
GCO	**0.38**	**0.74**	0.06	**0.64**	**0.89**	-0.03	**0.62**	-0.04	**0.13**	**0.91**	0.17
CAR	**0.27**	**0.93**	0.16	**0.77**	**0.85**	0.05	**0.2**	-0.04	0.08	**0.49**	-0.06
CRO	**0.33**	**0.57**	-0.05	**0.88**	**0.54**	**0.32**	**0.54**	-0.04	0.01	**0.7**	0.13
PON	**0.44**	**0.69**	**0.46**	**0.81**	**0.61**	-0.1	**0.82**	-0.05	**0.55**	**0.62**	**0.29**
PUA	**0.25**	**0.59**	-0.06	0.16	0.11	0.04	**0.57**	0.11	0.08	**0.89**	-0.06
PPA	**0.23**	**0.93**	0.09	0.23	-0.03	0.04	**0.32**	-0.02	**0.16**	**0.68**	0.05
PZU	**0.35**	**0.48**	0.2	0.1	**0.26**	0.09	**0.48**	0.05	**0.66**	**0.88**	0
PZP	**0.21**	**0.57**	-0.06	0.16	0.09	0.21	**0.46**	-0.06	-0.02	**0.58**	0.03
GGU	**0.3**	**0.87**	-0.19	0.16	**0.59**	0.07	**0.4**	0.02	0.05	**0.71**	0.06
CDP	**0.38**	**0.19**	**0.63**	**0.67**	**0.27**	0	**1**	0.05	**0.17**	**0.68**	0.31
BCA	**0.25**	**0.59**	-0.28	0.11	**0.39**	-0.02	**0.63**	0.03	0.02	**0.92**	0.04
POR	**0.37**	**0.58**	-0.07	**0.93**	0.09	0.07	**0.58**	0.03	**0.25**	**0.79**	-0.01
BAN	**0.28**	**0.56**	0.08	**0.66**	**0.5**	0.1	**0.44**	-0.09	0	**0.71**	-0.06
SRE	**0.42**	**0.9**	0.03	**0.77**	**0.64**	0.04	**0.68**	-0.09	**0.62**	**0.56**	0.1
NOR	**0.26**	**0.77**	0.01	**0.79**	0.11	0.03	**0.39**	-0.12	-0.02	**0.8**	0.07
POT	**0.3**	**0.62**	0.18	**0.35**	**0.45**	-0.07	**0.24**	0.14	**0.38**	**0.59**	0.07
CDR	**0.28**	**0.64**	0.26	**0.2**	**0.66**	0.13	0.1	-0.04	**0.34**	**0.66**	-0.05
COV	**0.27**	**0.37**	0.11	**0.37**	**0.58**	-0.04	0.05	-0.05	**0.59**	**0.74**	-0.05
BER	**0.27**	**0.12**	**0.68**	**0.46**	**0.44**	-0.06	**0.2**	-0.04	**0.39**	**0.6**	-0.06
XIC	**0.2**	**0.31**	0.06	**0.35**	**0.4**	-0.04	**0.28**	-0.06	**0.33**	**0.57**	-0.2
FOR	**0.22**	**0.07**	0.21	**0.86**	0.08	0.13	0.14	-0.03	0.03	**0.85**	-0.09
LPP	**0.31**	**0.3**	0.03	**0.64**	0.17	**0.19**	**0.62**	0.02	**0.14**	**0.86**	0.14

		A																					B								D								C	
		VIL	MON	PDA	GC	GRC	GPE	GPI	PLA	RIO	RIE	RII	FII	FIE	PHA	IML	TFR	CAS	CRO	CAR	PON	TRM	GCO	PZU	POR	CDP	PZP	BCA	PPA	PUA	GGU	NOR	XIC	POT	COV	CDR	BER	SRED	BAN	FOR	LPP
A	VIL		0.113	0.155	0.088	0.078	0.08	0.077	0.073	0.068	0.126	0.081	0.076	0.082	0.124	0.07	0.121	0.1	0.087	0.131	0.073	0.098	0.091	0.165	0.181	0.148	0.147	0.118	0.141	0.119	0.135	0.115	0.103	0.107	0.112	0.155	0.099	0.112	0.108	0.13	0.114
	MON	8.51		0.147	0.116	0.111	0.096	0.1	0.088	0.083	0.161	0.102	0.145	0.108	0.154	0.074	0.073	0.128	0.131	0.193	0.118	0.131	0.151	0.196	0.169	0.168	0.163	0.14	0.148	0.145	0.149	0.143	0.117	0.108	0.116	0.152	0.11	0.105	0.12	0.145	0.147
	PDA	166.38	158.53		0.151	0.12	0.128	0.116	0.107	0.087	0.193	0.098	0.139	0.09	0.175	0.101	0.112	0.115	0.128	0.196	0.113	0.116	0.127	0.132	0.151	0.16	0.092	0.099	0.129	0.092	0.098	0.119	0.09	0.087	0.087	0.138	0.077	0.134	0.093	0.134	0.148
	GC	198.16	206.67	357.72		0.015	0.076	0.063	0.051	0.044	0.061	0.037	0.084	0.077	0.136	0.026	0.109	0.101	0.071	0.171	0.059	0.093	0.108	0.189	0.147	0.162	0.161	0.119	0.157	0.126	0.148	0.126	0.111	0.096	0.109	0.138	0.094	0.134	0.114	0.135	0.115
	GRC	198.24	206.75	357.83	0.09		0.08	0.059	0.043	0.029	0.083	0.021	0.07	0.065	0.12	0.014	0.104	0.086	0.059	0.146	0.036	0.089	0.08	0.163	0.138	0.135	0.138	0.096	0.143	0.105	0.127	0.106	0.095	0.081	0.093	0.123	0.069	0.123	0.098	0.131	0.109
	GPE	199.58	208.09	359.11	1.66	1.75		**-0.004**	0.039	0.041	0.107	0.068	0.053	0.051	0.062	0.054	0.08	0.069	0.068	0.135	0.059	0.081	0.081	0.125	0.092	0.118	0.097	0.081	0.084	0.075	0.09	0.06	0.058	0.043	0.048	0.069	0.045	0.068	0.074	0.099	0.096
	GPI	199.59	208.1	359.12	1.67	1.76	0.005		0.029	0.034	0.105	0.055	0.053	0.044	0.079	0.045	0.079	0.06	0.061	0.118	0.047	0.078	0.07	0.122	0.101	0.113	0.095	0.075	0.078	0.071	0.089	0.063	0.059	0.043	0.048	0.074	0.038	0.073	0.073	0.103	0.1
	PLA	199.5	208.01	359.03	1.58	1.67	0.08	0.09		0.038	0.088	0.043	0.057	0.055	0.103	0.047	0.064	0.072	0.055	0.138	0.042	0.064	0.068	0.123	0.119	0.116	0.109	0.081	0.099	0.093	0.091	0.077	0.066	0.049	0.049	0.083	0.04	0.082	0.064	0.103	0.089
	RIO	197.53	206.04	358.61	1.12	1.03	3.15	3.16	3.1		0.056	0.005	0.051	0.036	0.07	0.017	0.056	0.055	0.041	0.111	0.026	0.052	0.046	0.102	0.086	0.091	0.073	0.039	0.072	0.052	0.069	0.048	0.046	0.045	0.049	0.075	0.037	0.073	0.053	0.075	0.06
	RIE	197.43	205.94	358.51	1.02	0.93	3.25	3.26	3.2	0.1		0.048	0.118	0.117	0.153	0.082	0.118	0.12	0.108	0.199	0.096	0.109	0.134	0.216	0.157	0.181	0.18	0.134	0.173	0.154	0.162	0.142	0.14	0.122	0.134	0.164	0.126	0.149	0.125	0.14	0.114
	RII	197.45	205.96	358.53	1.04	0.95	3.27	3.27	3.21	0.12	0.015		0.063	0.049	0.105	0.018	0.067	0.083	0.053	0.118	0.04	0.053	0.062	0.121	0.107	0.114	0.101	0.064	0.101	0.076	0.089	0.077	0.073	0.07	0.073	0.101	0.057	0.102	0.074	0.096	0.078
	FII	196.78	205.29	357.03	4.9	4.99	5.06	5.07	4.98	6.13	6.03	6.05		0.021	0.099	0.058	0.126	0.098	0.076	0.114	0.074	0.105	0.063	0.134	0.146	0.159	0.124	0.097	0.115	0.1	0.105	0.08	0.092	0.081	0.087	0.114	0.063	0.098	0.103	0.149	0.127
	FIE	196.77	205.28	357.02	4.89	4.98	5.05	5.06	4.97	6.12	6.02	6.04	0.01		0.097	0.042	0.085	0.074	0.055	0.134	0.044	0.077	0.081	0.099	0.106	0.119	0.081	0.057	0.082	0.05	0.079	0.061	0.058	0.051	0.059	0.085	0.045	0.069	0.076	0.109	0.1
	PHA	205.28	213.79	365.24	6.35	6.26	4.82	4.83	4.9	5.79	5.89	5.91	8.89	8.88		0.097	0.133	0.101	0.087	0.168	0.087	0.122	0.071	0.166	0.123	0.168	0.146	0.118	0.109	0.107	0.137	0.082	0.095	0.083	0.092	0.094	0.099	0.085	0.122	0.123	0.123
	IML	197.06	205.57	358.11	0.91	0.83	3.3	3.3	3.22	0.62	0.52	0.54	5.79	5.78	6.36		0.075	0.074	0.051	0.125	0.036	0.071	0.075	0.145	0.101	0.115	0.117	0.074	0.112	0.079	0.096	0.082	0.075	0.057	0.074	0.094	0.056	0.087	0.075	0.107	0.091
	TFR	214.8	223.31	375.51	15.97	15.88	14.45	14.46	14.53	15.41	15.51	15.53	18.51	18.5	9.62	15.98		0.101	0.103	0.17	0.092	0.084	0.114	0.141	0.124	0.149	0.101	0.081	0.093	0.094	0.084	0.099	0.077	0.08	0.083	0.11	0.072	0.09	0.074	0.106	0.107
	CAS	192.37	200.88	352.1	8.23	8.34	8.9	8.91	8.82	9.45	9.35	9.37	4.4	4.39	13.07	9.14	22.69		0.08	0.169	0.072	0.084	0.071	0.147	0.147	0.152	0.118	0.096	0.123	0.096	0.113	0.089	0.087	0.063	0.083	0.126	0.076	0.104	0.074	0.109	0.099
	CRO	229.9	238.41	390.4	31.87	31.77	30.33	30.34	30.41	31.3	31.4	31.42	34.4	34.39	25.51	31.67	18.82	38.58		0.146	0.011	0.095	0.079	0.118	0.123	0.13	0.114	0.085	0.105	0.073	0.115	0.075	0.068	0.063	0.066	0.093	0.067	0.086	0.086	0.094	0.082
	CAR	224.08	232.59	384.61	25.29	25.5	24.06	24.07	24.14	25.03	25.13	25.15	28.13	28.12	19.24	25.6	11.23	32.31	8.26		0.145	0.134	0.102	0.185	0.206	0.183	0.176	0.147	0.146	0.162	0.148	0.135	0.16	0.156	0.157	0.179	0.13	0.156	0.151	0.183	0.164
	PON	238.12	246.63	398.23	39.76	39.67	38.23	38.24	38.31	39.2	39.3	39.32	42.3	42.29	33.41	39.77	29.98	46.48	7.86	16.12		0.084	0.084	0.125	0.126	0.109	0.102	0.07	0.108	0.07	0.107	0.066	0.056	0.054	0.056	0.092	0.054	0.079	0.079	0.094	0.089
	TRM	193.04	201.55	351.28	9.66	9.75	10.33	10.34	10.25	10.88	10.78	10.8	5.83	5.82	14.5	10.57	24.12	1.43	40.01	33.74	47.91		0.084	0.123	0.104	0.139	0.113	0.089	0.088	0.087	0.084	0.076	0.073	0.067	0.074	0.11	0.073	0.087	0.076	0.075	0.049
	GCO	205.88	214.39	365.74	6.94	6.85	5.41	5.42	5.49	6.38	6.48	6.5	9.48	9.47	0.59	6.95	9.03	13.66	24.98	18.74	32.81	15.09		0.131	0.149	0.16	0.124	0.096	0.107	0.106	0.1	0.074	0.091	0.082	0.089	0.12	0.069	0.116	0.105	0.131	0.11
B	PZU	174.9	175.42	221.12	272.57	272.48	273.98	273.99	273.9	273.36	273.26	273.28	271.05	271.04	278.89	272.13	288.66	266.54	304.81	298.03	311.98	265.94	279.39		0.141	0.16	0.086	0.093	0.096	0.077	0.101	0.095	0.091	0.105	0.082	0.131	0.099	0.13	0.112	0.121	0.117
	POR	192.84	193.47	241.96	287.04	287.08	288.41	288.42	288.33	287.9	287.8	287.82	285.31	285.3	294.02	287.19	303.68	280.3	319.21	313.19	327	279.7	294.52	19.78		0.154	0.117	0.095	0.107	0.099	0.1	0.1	0.092	0.074	0.077	0.082	0.084	0.096	0.109	0.11	0.096
	CDP	177.9	178.2	224.75	273.78	273.69	275.22	275.23	275.14	274.59	274.49	274.51	272.49	272.48	280.28	273.55	290.03	266.53	305.59	299.41	313.36	265.93	280.78	3.51	16.53		0.118	0.067	0.14	0.101	0.107	0.119	0.121	0.104	0.104	0.147	0.105	0.13	0.113	0.138	0.129
	PZP	174.95	175.47	221.04	272.52	272.43	273.93	273.94	273.85	273.31	273.21	273.23	271	270.99	278.84	272.08	288.61	266.49	304.76	297.98	311.93	265.89	279.34	0.05	19.73	3.49		0.038	0.073	0.041	0.053	0.076	0.067	0.079	0.072	0.102	0.07	0.113	0.1	0.102	0.126
	BCA	178.12	178.5	225.12	273.83	273.74	275.25	275.26	275.17	274.84	274.74	274.76	272.49	272.48	280.37	273.65	290.12	266.65	273.88	299.5	313.45	266.05	280.87	3.93	15.97	0.56	3.91		0.057	0.034	0.037	0.053	0.063	0.059	0.062	0.089	0.054	0.084	0.07	0.095	0.088
	PPA	175.06	175.57	221.03	272.67	272.58	274.08	274.09	274	273.46	273.36	273.38	272.15	272.14	279	272.23	288.77	266.65	304.91	298.13	312.08	266.05	279.5	0.1	19.92	3.71	0.15	4.05		0.044	0.066	0.048	0.073	0.077	0.077	0.083	0.075	0.067	0.086	0.091	0.101
	PUA	175.18	175.68	220.99	272.89	272.8	274.3	274.31	274.22	273.68	273.58	273.6	271.37	271.36	279.2	272.43	288.96	266.84	305.13	298.33	312.28	266.24	279.7	0.32	20.09	3.85	0.36	4.27	0.21		0.059	0.05	0.061	0.061	0.064	0.079	0.065	0.068	0.07	0.083	0.092
	GGU	175.25	175.82	222.04	272.16	272.07	273.54	273.55	273.46	272.98	272.88	272.9	270.63	270.62	278.44	271.69	288.2	266.08	304.35	297.57	311.52	265.48	278.94	1.14	19.55	2.79	1.1	3.19	1.19	1.43		0.06	0.062	0.063	0.06	0.093	0.05	0.086	0.084	0.107	0.101
D	NOR	383.13	391.64	541.59	202.99	202.9	203.25	203.26	203.33	201.87	201.97	201.99	207.87	207.86	200.31	202.12	201.22	210.9	188.65	195.68	188.42	212.21	200.17	434.93	446.6	435.43	434.88	435.33	435.03	435.25	434.17		0.044	0.036	0.043	0.057	0.043	0.064	0.072	0.088	0.089
	XIC	393.58	402.09	552.64	218.14	218.05	218.65	218.66	218.73	217.02	217.12	217.14	223.05	223.04	215.9	217.23	217.53	225.98	206.05	212.63	206.2	227.19	215.72	440.31	451.52	440.64	440.26	440.51	440.41	440.63	439.49	22.58		0.015	0.007	0.035	0.013	0.053	0.059	0.072	0.087
	POT	393.42	401.93	552.46	217.82	217.73	218.31	218.32	218.39	216.7	216.8	216.82	222.72	222.71	215.56	216.9	217.16	225.65	205.63	212.33	205.88	226.86	215.38	440.32	451.54	440.6	440.27	440.47	440.42	440.64	439.38	21.86	0.75		0.003	0.023	0.012	0.043	0.038	0.074	0.078
	COV	393.48	401.99	552.54	217.95	217.86	218.45	218.46	218.53	216.83	216.93	216.95	222.85	222.84	215.7	217.03	217.31	225.78	205.8	212.5	206.05	226.99	215.52	440.35	451.53	440.65	440.3	440.52	440.45	440.67	439.5	22.18	0.46	0.34		0.029	0.006	0.053	0.048	0.074	0.085
	CDR	393.61	402.12	552.67	218.09	218	218.59	218.6	218.67	216.97	217.07	217.09	222.99	222.98	215.84	217.17	217.45	225.93	205.95	212.65	206.19	227.13	215.66	440.42	451.63	440.76	440.37	440.62	440.52	440.74	439.6	22.35	0.45	0.49	0.16		0.04	0.062	0.076	0.102	0.113
	BER	393.5	402.01	552.58	218.12	218.03	218.62	218.63	218.7	217	217.1	217.12	223.02	223.01	215.88	217.2	217.5	225.95	206.07	212.73	206.3	227.15	215.7	440.19	451.38	440.52	440.14	440.39	440.29	440.51	439.37	22.71	0.21	0.89	0.61	0.55		0.058	0.046	0.092	0.092
	SRED	385.08	393.59	542.86	198.9	198.81	198.94	198.95	199.02	197.78	197.88	197.9	203.88	203.87	195.56	198.09	195.3	207.07	181.05	188.59	179.82	208.32	195.21	442.26	454.72	442.94	442.21	442.84	442.36	442.58	441.55	32.06	54.64	53.92	54.24	54.41	54.77		0.046	0.076	0.079
	BAN	385.52	394.03	543.03	198.78	198.69	198.8	198.81	198.88	197.66	197.76	197.78	203.74	203.73	195.39	197.97	195.11	206.99	180.77	188.33	179.48	208.21	195.12	442.65	455.12	443.27	442.6	443.25	442.75	442.97	441.94	33.31	55.89	55.17	55.49	55.56	56.02	1.25		0.076	0.07
C	FOR	534.39	542.72	689.6	397.19	396.89	399.1	399.11	399	396.38	396.48	396.5	401.38	401.37	397.49	395.9	402.87	402.68	398.75	402.8	401.82	403.56	397.56	519.04	524.97	518.09	519.41	517.74	519.05	519.17	518.95	249.24	227.64	228.32	228.03	228.19	227.43	279.24	280.46		0.034
	LPP	482.9	491.31	632.12	360.34	360.66	362.6	362.61	362.5	360.1	360.2	360.22	364.96	364.95	362.54	359.91	369.11	365.52	368.37	371.08	372.64	366.28	362.71	451.28	456.51	450.15	451.45	449.76	451.29	451.4	450.17	250.31	230.53	231.13	230.88	230.85	230.32	279.88	281.05	79.5	

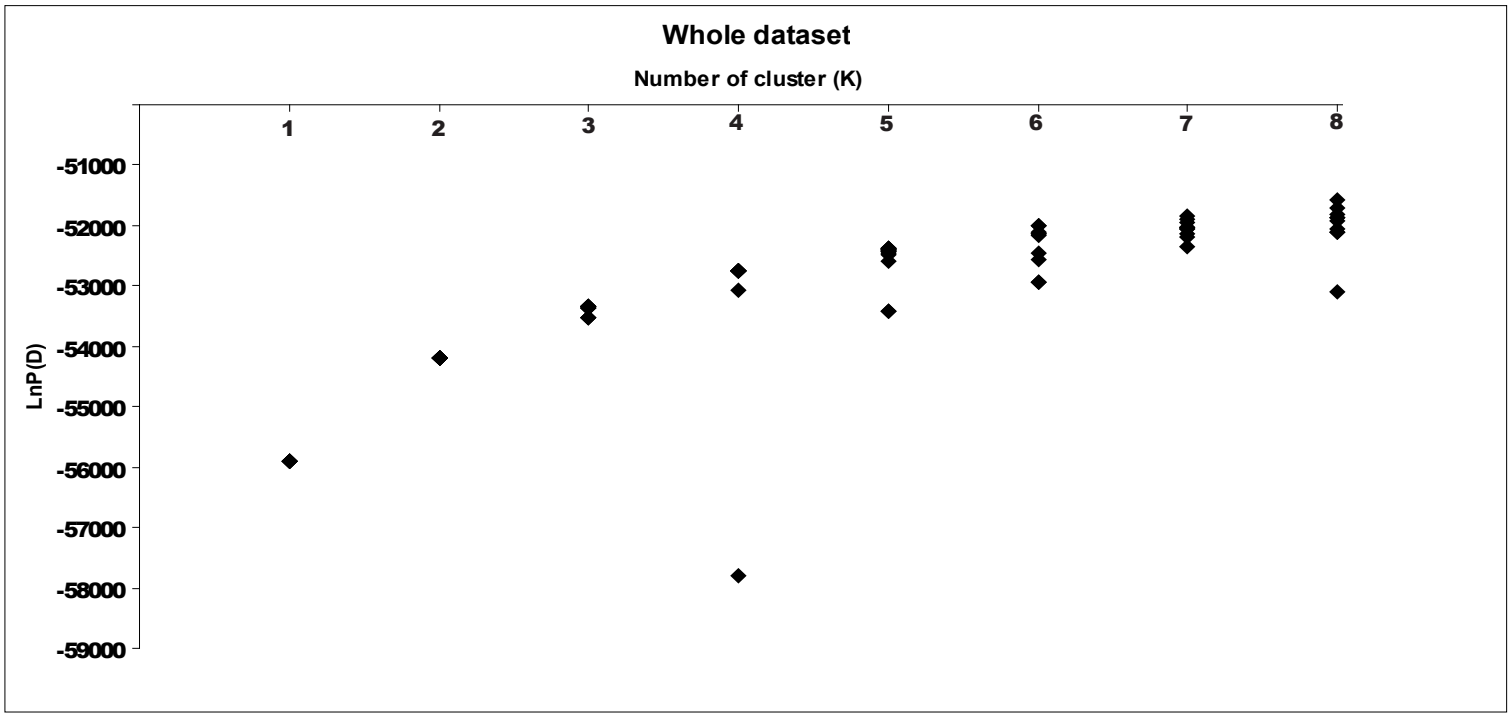
Table S3a

		A																					B								D								C	
		VIL	MON	PDA	GC	GRC	GPE	GPI	PLA	RIO	RIE	RII	FII	FIE	PHA	IML	TFR	CAS	CRO	CAR	PON	TRM	GCO	PZU	POR	CDP	PZP	BCA	PPA	PUA	GGU	NOR	XIC	POT	COV	CDR	BER	SRED	BAN	FOR	LPP
A	VIL		0.113	0.155	0.088	0.078	0.08	0.077	0.073	0.068	0.126	0.081	0.076	0.082	0.124	0.07	0.121	0.1	0.087	0.131	0.073	0.098	0.091	0.165	0.181	0.148	0.147	0.118	0.141	0.119	0.135	0.115	0.103	0.107	0.112	0.155	0.099	0.112	0.108	0.13	0.114
	MON	8.51		0.147	0.116	0.111	0.096	0.1	0.088	0.083	0.161	0.102	0.145	0.108	0.154	0.074	0.073	0.128	0.131	0.193	0.118	0.131	0.151	0.196	0.169	0.168	0.163	0.14	0.148	0.145	0.149	0.143	0.117	0.108	0.116	0.152	0.11	0.105	0.12	0.145	0.147
	PDA	201.63	194.9		0.151	0.12	0.128	0.116	0.107	0.087	0.193	0.098	0.139	0.09	0.175	0.101	0.112	0.115	0.128	0.196	0.113	0.116	0.127	0.132	0.151	0.16	0.092	0.099	0.129	0.092	0.098	0.119	0.09	0.087	0.087	0.138	0.077	0.134	0.093	0.134	0.148
	GC	315.77	325.22	512.58		0.015	0.076	0.063	0.051	0.044	0.061	0.037	0.084	0.077	0.136	0.026	0.109	0.101	0.071	0.171	0.059	0.093	0.108	0.189	0.147	0.162	0.161	0.119	0.157	0.126	0.148	0.126	0.111	0.096	0.109	0.138	0.094	0.134	0.114	0.135	0.115
	GRC	315.86	325.31	512.67	0.09		0.08	0.059	0.043	0.029	0.083	0.021	0.07	0.065	0.12	0.014	0.104	0.086	0.059	0.146	0.036	0.089	0.08	0.163	0.138	0.135	0.138	0.096	0.143	0.105	0.127	0.106	0.095	0.081	0.093	0.123	0.069	0.123	0.098	0.131	0.109
	GPE	317.35	326.7	514.1	1.66	1.75		-0.004	0.039	0.041	0.107	0.068	0.053	0.051	0.062	0.054	0.08	0.069	0.068	0.135	0.059	0.081	0.081	0.125	0.092	0.118	0.097	0.081	0.084	0.075	0.09	0.06	0.058	0.043	0.048	0.069	0.045	0.068	0.074	0.099	0.096
	GPI	317.36	326.71	514.11	1.67	1.76	0.005		0.029	0.034	0.105	0.055	0.053	0.044	0.079	0.045	0.079	0.06	0.061	0.118	0.047	0.078	0.07	0.122	0.101	0.113	0.095	0.075	0.078	0.071	0.089	0.063	0.059	0.043	0.048	0.074	0.038	0.073	0.073	0.103	0.1
	PLA	317.27	326.62	514.02	1.58	1.67	0.08	0.09		0.038	0.088	0.043	0.057	0.055	0.103	0.047	0.064	0.072	0.055	0.138	0.042	0.064	0.068	0.123	0.119	0.116	0.109	0.081	0.099	0.093	0.091	0.077	0.066	0.049	0.049	0.083	0.04	0.082	0.064	0.103	0.089
	RIO	316.55	325.87	513.37	1.12	1.03	3.15	3.16	3.1		0.056	0.005	0.051	0.036	0.07	0.017	0.056	0.055	0.041	0.111	0.026	0.052	0.046	0.102	0.086	0.091	0.073	0.039	0.072	0.052	0.069	0.048	0.046	0.045	0.049	0.075	0.037	0.073	0.053	0.075	0.06
	RIE	316.65	325.97	513.47	1.02	0.93	3.25	3.26	3.2	0.1		0.048	0.118	0.117	0.153	0.082	0.118	0.12	0.108	0.199	0.096	0.109	0.134	0.216	0.157	0.181	0.18	0.134	0.173	0.154	0.162	0.142	0.14	0.122	0.134	0.164	0.126	0.149	0.125	0.14	0.114
	RII	316.67	325.99	513.49	1.04	0.95	3.27	3.27	3.21	0.12	0.015		0.063	0.049	0.105	0.018	0.067	0.083	0.053	0.118	0.04	0.053	0.062	0.121	0.107	0.114	0.101	0.064	0.101	0.076	0.089	0.077	0.073	0.07	0.073	0.101	0.057	0.102	0.074	0.096	0.078
	FII	316.18	325.24	512.97	4.9	4.99	5.06	5.07	4.98	6.13	6.03	6.05		0.021	0.099	0.058	0.126	0.098	0.076	0.114	0.074	0.105	0.063	0.134	0.146	0.159	0.124	0.097	0.115	0.1	0.105	0.08	0.092	0.081	0.087	0.114	0.063	0.098	0.103	0.149	0.127
	FIE	316.17	325.23	512.96	4.89	4.98	5.05	5.06	4.97	6.12	6.02	6.04	0.01		0.097	0.042	0.085	0.074	0.055	0.134	0.044	0.077	0.081	0.099	0.106	0.119	0.081	0.057	0.082	0.05	0.079	0.061	0.058	0.051	0.059	0.085	0.045	0.069	0.076	0.109	0.1
	PHA	329.94	340.46	524.5	6.35	6.26	4.82	4.83	4.9	5.79	5.89	5.91	8.89	8.88		0.097	0.133	0.101	0.087	0.168	0.087	0.122	0.071	0.166	0.123	0.168	0.146	0.118	0.109	0.107	0.137	0.082	0.095	0.083	0.092	0.094	0.099	0.085	0.122	0.123	0.123
	IML	316.04	325.26	512.77	0.91	0.83	3.3	3.3	3.22	0.62	0.52	0.54	5.79	5.78	6.36		0.075	0.074	0.051	0.125	0.036	0.071	0.075	0.145	0.101	0.115	0.117	0.074	0.112	0.079	0.096	0.082	0.075	0.057	0.074	0.094	0.056	0.087	0.075	0.107	0.091
	TFR	352.12	364.24	545.99	15.97	15.88	14.45	14.46	14.53	15.41	15.51	15.53	18.51	18.5	9.62	15.98		0.101	0.103	0.17	0.092	0.084	0.114	0.141	0.124	0.149	0.101	0.081	0.093	0.094	0.084	0.099	0.077	0.08	0.083	0.11	0.072	0.09	0.074	0.106	0.107
	CAS	314.15	323.31	511.19	8.23	8.34	8.9	8.91	8.82	9.45	9.35	9.37	4.4	4.39	13.07	9.14	22.69		0.08	0.169	0.072	0.084	0.071	0.147	0.147	0.152	0.118	0.096	0.123	0.096	0.113	0.089	0.087	0.063	0.083	0.126	0.076	0.104	0.074	0.109	0.099
	CRO	368.06	373.78	561.84	31.87	31.77	30.33	30.34	30.41	31.3	31.4	31.42	34.4	34.39	25.51	31.67	18.82	38.58		0.146	0.011	0.095	0.079	0.118	0.123	0.13	0.114	0.085	0.105	0.073	0.115	0.075	0.068	0.063	0.066	0.093	0.067	0.086	0.086	0.094	0.082
	CAR	367.02	370.61	558.84	25.29	25.5	24.06	24.07	24.14	25.03	25.13	25.15	28.13	28.12	19.24	25.6	11.23	32.31	8.26		0.145	0.134	0.102	0.185	0.206	0.183	0.176	0.147	0.146	0.162	0.148	0.135	0.16	0.156	0.157	0.179	0.13	0.156	0.151	0.183	0.164
	PON	374.85	384.16	569.19	39.76	39.67	38.23	38.24	38.31	39.2	39.3	39.32	42.3	42.29	33.41	39.77	29.98	46.48	7.86	16.12		0.084	0.084	0.125	0.126	0.109	0.102	0.07	0.108	0.07	0.107	0.066	0.056	0.054	0.056	0.092	0.054	0.079	0.079	0.094	0.089
	TRM	314.49	323.78	511.6	9.66	9.75	10.33	10.34	10.25	10.88	10.78	10.8	5.83	5.82	14.5	10.57	24.12	1.43	40.01	33.74	47.91		0.084	0.123	0.104	0.139	0.113	0.089	0.088	0.087	0.084	0.076	0.073	0.067	0.074	0.11	0.073	0.087	0.076	0.075	0.049
	GCO	330.53	341.05	525.09	6.94	6.85	5.41	5.42	5.49	6.38	6.48	6.5	9.48	9.47	0.59	6.95	9.03	13.66	24.98	18.74	32.81	15.09		0.131	0.149	0.16	0.124	0.096	0.107	0.106	0.1	0.074	0.091	0.082	0.089	0.12	0.069	0.116	0.105	0.131	0.11
B	PZU	508.8	502.59	316.78	811.85	811.76	813.06	813.07	812.98	812.25	812.15	812.17	811.73	811.72	837.1	812.78	851.09	809.85	860.08	857.26	870.14	810.43	837.69		0.141	0.16	0.086	0.093	0.096	0.077	0.101	0.095	0.091	0.105	0.082	0.131	0.099	0.13	0.112	0.121	0.117
	POR	528.57	522.36	336.55	832.26	832.17	833.47	833.47	833.39	832.66	832.56	832.57	832.14	832.13	857.51	833.19	871.5	830.26	880.49	877.67	890.55	830.84	858.1	19.78		0.154	0.117	0.095	0.107	0.099	0.1	0.1	0.092	0.074	0.077	0.082	0.084	0.096	0.109	0.11	0.096
	CDP	512.33	506.12	320.31	816.02	815.93	817.23	817.23	817.15	816.42	816.32	816.33	815.9	815.89	841.27	816.95	855.26	814.02	864.25	861.43	874.31	814.6	841.86	3.51	16.53		0.118	0.067	0.14	0.101	0.107	0.119	0.121	0.104	0.104	0.147	0.105	0.13	0.113	0.138	0.129
	PZP	508.84	502.63	316.82	811.81	811.72	813.02	813.03	812.94	812.21	812.11	812.13	811.69	811.68	837.06	812.74	851.05	809.81	860.04	857.22	870.1	810.39	837.65	0.05	19.73	3.49		0.038	0.073	0.041	0.053	0.076	0.067	0.079	0.072	0.102	0.07	0.113	0.1	0.102	0.126
	BCA	512.75	506.54	320.73	816.44	816.35	817.65	817.65	817.57	816.84	816.74	816.75	816.32	816.31	841.69	817.37	855.68	814.44	864.67	861.85	874.73	815.02	842.28	3.93	15.97	0.56	3.91		0.057	0.034	0.037	0.053	0.063	0.059	0.062	0.089	0.054	0.084	0.07	0.095	0.088
	PPA	508.69	502.48	316.67	811.96	811.87	813.17	813.18	813.09	812.36	812.26	812.28	811.84	811.83	837.21	812.89	851.2	809.96	860.19	857.37	870.25	810.54	837.8	0.1	19.92	3.71	0.15	4.05		0.044	0.066	0.048	0.073	0.077	0.077	0.083	0.075	0.067	0.086	0.091	0.101
	PUA	508.48	502.27	316.46	812.17	812.08	813.38	813.39	813.3	812.57	812.47	812.49	812.05	812.04	837.42	813.1	851.41	810.17	860.4	857.58	870.46	810.75	838.01	0.32	20.09	3.85	0.36	4.27	0.21		0.059	0.05	0.061	0.061	0.064	0.079	0.065	0.068	0.07	0.083	0.092
	GGU	509.91	503.7	317.89	813.6	813.51	814.81	814.82	814.73	814	813.9	813.92	813.48	813.47	838.85	814.53	852.84	811.6	861.83	859.01	871.89	812.18	839.44	1.14	19.55	2.79	1.1	3.19	1.19	1.43		0.06	0.062	0.063	0.06	0.093	0.05	0.086	0.084	0.107	0.101
D	NOR	632.62	640.28	824.82	337.25	337.17	334.98	334.99	335.06	335.95	336.05	336.07	357.5	357.49	330.16	336.34	335.44	351.23	325.9	333.19	323.7	347.44	330.75	1133.24	1153.01	1136.77	1133.28	1137.19	1133.13	1132.92	1134.35		0.044	0.036	0.043	0.057	0.043	0.064	0.072	0.088	0.089
	XIC	666.2	674.01	859.63	357.95	357.87	356.14	356.15	356.22	357.11	357.21	357.23	373.21	373.2	351.32	357.04	369.18	370.08	347.97	355.18	345.52	373.94	350.73	1174.22	1193.99	1177.75	1174.26	1178.17	1174.11	1173.9	1175.33	22.58		0.015	0.007	0.035	0.013	0.053	0.059	0.072	0.087
	POT	665.45	673.26	858.88	357.2	357.12	355.39	355.4	355.47	356.36	356.46	356.48	372.46	372.45	350.57	356.29	368.43	369.33	347.22	354.43	344.77	373.19	349.98	1173.75	1193.52	1177.28	1173.79	1177.7	1173.64	1173.43	1174.86	21.86	0.75		0.003	0.023	0.012	0.043	0.038	0.074	0.078
	COV	665.79	673.6	859.22	357.54	357.46	355.73	355.74	355.81	356.7	356.8	356.82	372.8	372.79	350.91	356.63	368.77	369.67	347.56	354.77	345.11	373.53	350.32	1173.94	1193.71	1177.47	1173.98	1177.89	1173.83	1173.62	1175.05	22.18	0.46	0.34		0.029	0.006	0.053	0.048	0.074	0.085
	CDR	665.94	673.75	859.37	357.69	357.61	355.88	355.89	355.96	356.85	356.95	356.97	372.95	372.94	351.06	356.78	368.92	369.82	347.71	354.92	345.26	373.68	350.47	1174.1	1193.87	1177.63	1174.14	1178.05	1173.99	1173.78	1175.21	22.35	0.45	0.49	0.16		0.04	0.062	0.076	0.102	0.113
	BER	666.34	674.15	859.77	358.09	358.01	356.28	356.29	356.36	357.25	357.35	357.37	373.35	373.34	351.46	357.18	369.32	370.22	348.11	355.32	345.66	374.08	350.87	1174.27	1194.04	1177.8	1174.31	1178.22	1174.16	1173.95	1175.38	22.71	0.21	0.89	0.61	0.55		0.058	0.046	0.092	0.092
	SRED	612.6	620.21	826.71	337.43	337.64	336.75	336.76	336.83	337.72	337.82	337.84	317.39	317.4	306.42	312.14	317.41	318.31	301.94	299.97	300.82	319.76	305.83	1110.18	1129.95	1113.71	1110.22	1114.13	1110.07	1109.86	1111.29	32.06	54.64	53.92	54.24	54.41	54.77		0.046	0.076	0.079
	BAN	612.52	620.33	826.29	351.78	351.69	330.27	330.28	330.35	331.24	331.34	331.36	317.17	317.27	306.21	312.02	317.38	318.28	301.69	299.81	300.68	319.54	305.62	1110.42	1130.19	1113.95	1110.46	1114.37	1110.31	1110.1	1111.53	33.31	55.89	55.17	55.49	55.56	56.02	1.25		0.076	0.07
C	FOR	1442.57	1452.49	1640.58	1101.11	1099.81	1106.77	1106.78	1106.69	1100.53	1100.63	1100.65	1105.5	1105.49	1090.36	1100.32	1095.93	1108.47	1086.65	1089.1	1091.6	1109.17	1090.47	519.04	524.97	518.09	519.41	517.74	519.05	519.17	518.95	784.64	753.99	754.67	754.38	754.27	753.81	830.24	831.35		0.034
	LPP	1522.07	1531.99	1720.08	1180.61	1179.31	1186.27	1186.28	1186.19	1180.03	1180.13	1180.15	1185	1184.99	1169.86	1179.82	1175.43	1187.97	1166.15	1168.6	1171.1	1188.67	1169.97	451.28	456.51	450.15	451.45	449.76	451.29	451.4	450.17	864.14	833.49	834.17	833.88	833.77	833.31	909.74	910.85	79.5	

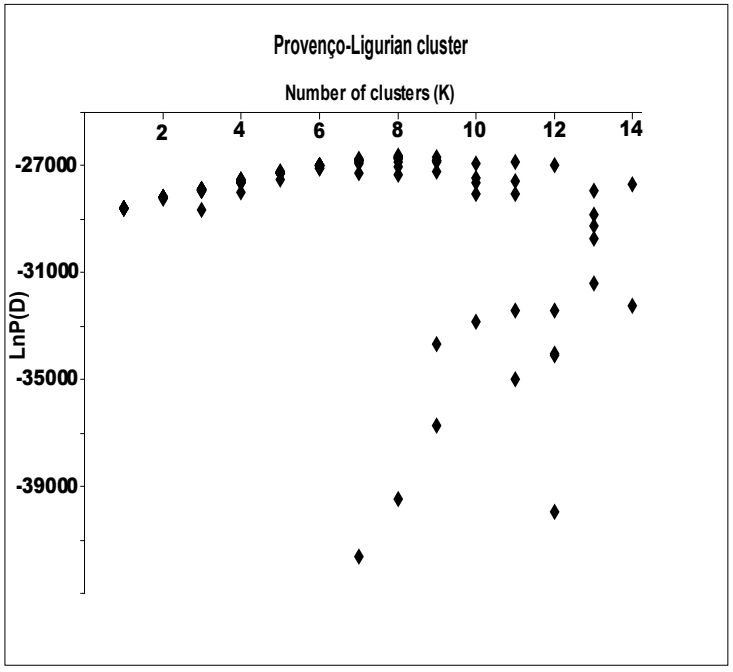
Table S3b

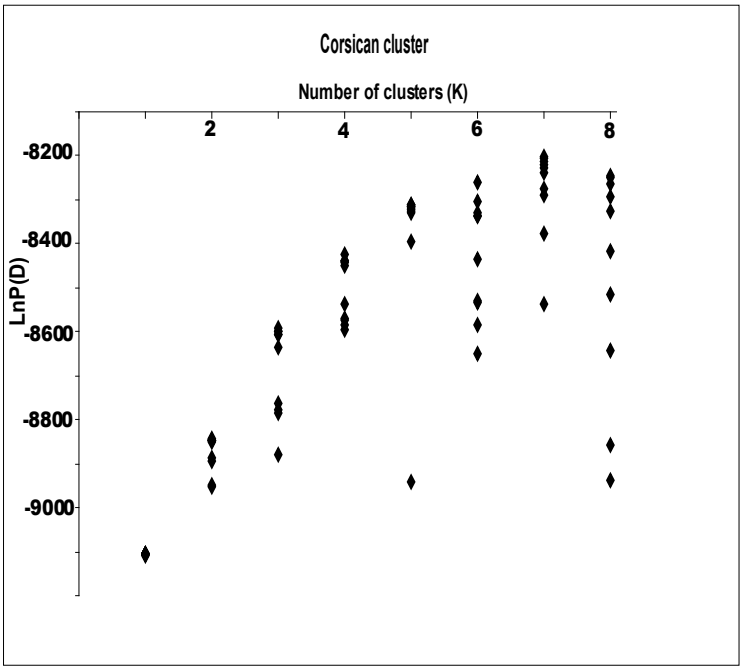
Figure S1

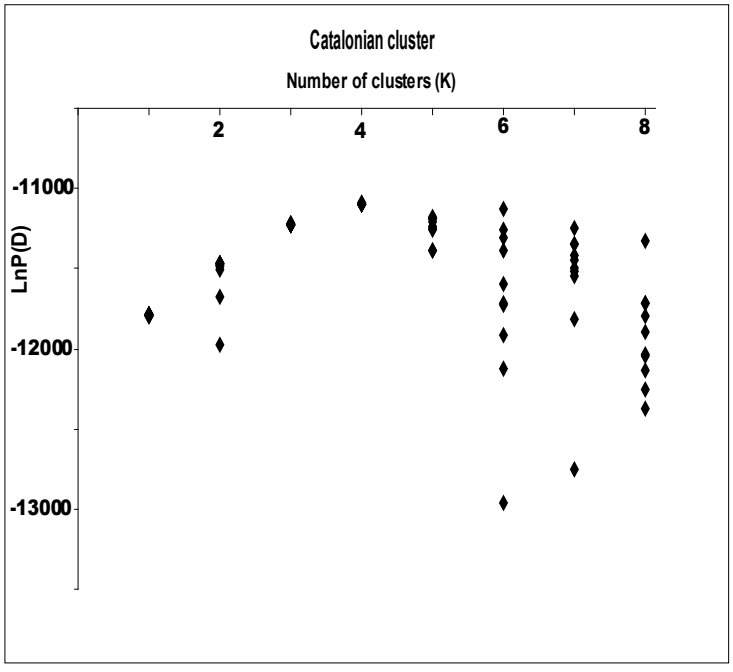
A

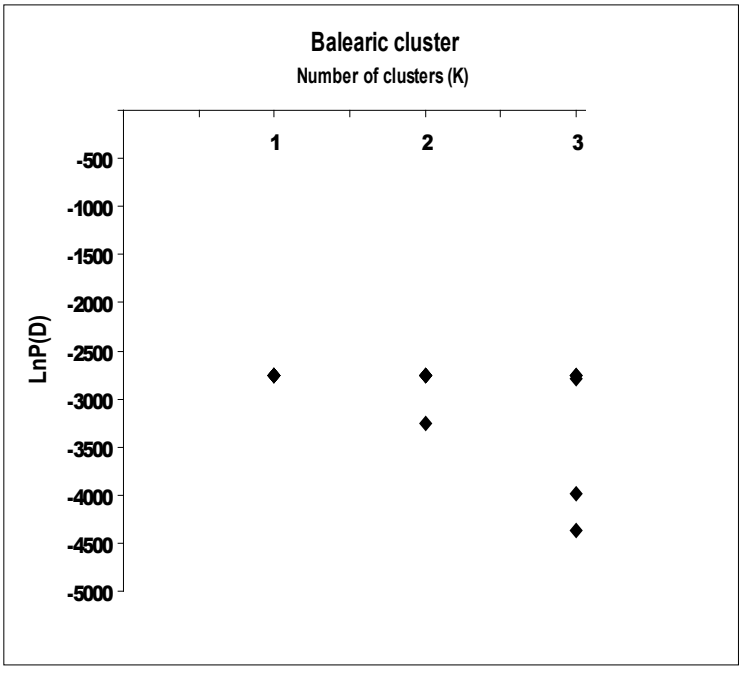

Whole dataset
Number of cluster (K)
LnP(D)
-51000
-52000
-53000
-54000
-55000
-56000
-57000
-58000
-59000
1
2
3
4
5
6
7
8

B


Provençe-Ligurian cluster
Number of clusters (K)
LnP(D)
-27000
-31000
-35000
-39000
2
4
6
8
10
12
14


Corsican cluster
Number of clusters (K)
LnP(D)
-8200
-8400
-8600
-8800
-9000
2
4
6
8


Catalonian cluster
Number of clusters (K)
LnP(D)
-11000
-12000
-13000
2
4
6
8


Balearic cluster
Number of clusters (K)
LnP(D)
-500
-1000
-1500
-2000
-2500
-3000
-3500
-4000
-4500
-5000
1
2
3