Mitochondrial and nuclear markers highlight the biodiversity of *Kalotermes flavicollis* (Fabricius, 1793) (Insecta, Isoptera, Kalotermitidae) in the Mediterranean area

A. Velonà, A. Luchetti, S. Ghesini, M. Marini and B. Mantovani*

Università degli Studi di Bologna, Dipartimento di Biologia Evoluzionistica Sperimentale, via Selmi 3, 40126 Bologna, Italy

Abstract

The biodiversity of the European termite Kalotermes flavicollis is here studied through the analysis of mitochondrial (303 bp of control region and 912 bp of COI/tRNA^{Leu}/COII) and nuclear (five microsatellite and 20 Inter-SINE loci) markers on 18 colonies collected in Southern France, Corsica, Sardinia, peninsular Italy, the Balkans and Greece. Different statistical analyses (Bayesian phylogenetic analysis, parsimony network, F-statistics, PCA) were performed. Mitochondrial sequences produced an unresolved polytomy including samples from peninsular Italy, Balkans and Greece, and three main clades: southern France, Corsica-Sardinia and Portoscuso (SW Sardinia). Nuclear markers confirm these data, further highlighting a more significant divergence at the regional scale. The results obtained for the peri-Tyrrhenian area agree with major paleogeographic and paleoclimatic events that shaped the biodiversity of the local fauna. K. flavicollis biodiversity and its phylogeographic pattern are also evaluated in the light of the data available for the other native European termite taxon (genus Reticulitermes), in order to produce a more complete scenario of the Mediterranean. In the area comprised between southern France and Italy, the degree of diversity is similar; however, in the eastern area, while K. flavicollis is differentiated only at the population level, the genus Reticulitermes comprises at least six entities of specific and/or subspecific level. This discrepancy may be explained by taking into account the different evolutionary histories of the two taxa.

Keywords: European area, evolutionary histories, molecular markers, phylogeography, termites

(Accepted 22 September 2010)

Introduction

The order Isoptera arose around 150 Myr ago in the late Jurassic/Early Cretaceous period (Engel *et al.*, 2009). At present, seven families are recognized; they comprise around 280 genera for a total of over 2600 diplo-diploid species (Engel & Krishna, 2004). Isopterans are particularly interesting for their behavior, reproductive biology and ecological roles (Inward *et al.*, 2007). The typical population structure of

*Author for correspondence Fax: 0039 051 2094286 E-mail: barbara.mantovani@unibo.it these insects consists of colonies, each composed of various castes. Colony foundation is often linked to a single pair of reproducing individuals. Over time, secondary males and females may be produced within the colony, increasing the reproductive output, but this closed system clearly lacks panmixia. Colony defense is performed by soldiers, while foraging and nursing are carried out by workers. In basal termite species (the so-called 'lower termites', as for example Kalotermitidae), workers are actually nymphs (named 'pseudoergates') and are characterized by the ability to differentiate either into soldiers or reproductives (Korb, 2007, 2008).

From an ecological point of view, termites represent one of the most important degraders in tropical and temperate forest ecosystems, their diet being principally soil and wood based. Termites digest plant material (cellulose) with the help of symbiotic bacteria and protozoans in their guts.

Two genera are native in Europe: *Reticulitermes* and *Kalotermes*. They show differences in ecology, habitat preference and colony dynamics; while *Reticulitermes* species are subterranean and build colonies of thousands of individuals with different reproductive centers in the soil, *Kalotermes* spp. produce smaller colonies, generally formed by few hundred individuals that do not have a strictly subterranean reproduction (Prota, 1962). The two genera also belong to different lifetypes. *Kalotermes* species (and more generally Kalotermitidae) are classified as 'one-piece' nesters; individuals nest and feed in the same substrate, i.e. dry wood. A different lifetype is proposed for *Reticulitermes* species, in which the nesting and the feeding areas are separated, forcing the individuals to forage away from the nest (Eggleton & Tayasu, 2001).

Hitherto, biodiversity and evolutionary researches have mainly focused on the Reticulitermes genus, which, according to recent data, consists of eight taxa of species level in Europe (Clément et al., 2001; Austin et al., 2002; Marini & Mantovani, 2002; Luchetti et al., 2004a,b, 2007; Uva et al., 2004; Velonà et al., 2010). In contrast, the genus Kalotermes has been poorly studied and is believed to comprise only one species in Europe, K. flavicollis (Fabricius, 1793) (Clément et al., 2001; Luchetti et al., 2004a). Yet, physiological, morphometric and reproductive biology data suggest that there may be some degree of geographic differentiation, e.g. between French and Sardinian colonies on one side, and colonies from the Italian peninsula on the other (Luscher, 1956; Springhetti, 1967; Rossi & Springhetti, 1983). Currently available mtDNA sequence data (910bp fragment of COI/tRNALeu/COII) suggest that there is no genetic structuring in K. flavicollis across Italy, the Balkans and Greece (Luchetti et al., 2004a). In contrast, within the same area, the genus Reticulitermes shows at least six lineages (Luchetti et al., 2007; Velonà et al., 2010).

In this paper, we present a molecular analysis of *K. flavicollis* colonies sampled from France, Italy, the Balkans and Greece. Both mitochondrial (part of the control region and the COI/tRNA^{Leu}/COII fragment) and nuclear (microsatellites and Inter-SINE loci) markers are used. In particular, five new polymorphic microsatellite loci are here characterized for *K. flavicollis* and the use of Inter-SINEs fingerprinting (Shafer & Stewart, 2007) is explored for the first time in Kalotermitidae.

SINEs (Short INterspersed Elements) are non-autonomous nuclear retrotransposons, usually present with more than 10^5 copies (Kostia *et al.*, 2000; Ohshima & Okada, 2005). The power of SINEs in phylogenetics is that their insertion is



Fig. 1. Sampling localities of *Kalotermes flavicollis* populations. Numbers refer to table 1.

random and unidirectional, so that the absence of a SINE is always an ancestral condition (Nishihara & Okada, 2008). In this analysis, we used a primer designed on the SINE *Talua*, and we amplified the region between two copies of the SINE (Luchetti, 2005; Luchetti & Mantovani, 2009) to produce a presence/absence matrix.

The aim of the present analysis is to infer genetic relationships between Mediterranean populations of *K. flavicollis* and to correlate these with the history of the basin. Results are also compared with those previously obtained for *Reticulitermes* spp. (Luchetti *et al.*, 2005, 2007; Velonà *et al.*, 2010) in order to evaluate the contribution of the historical events that have shaped the present-day biodiversity in this geographic area.

Methods

Sampling

Present work was carried out on 18 colonies collected in Greece, the Balkans, Italy and France (fig. 1, table 1). Five of these samples (Amorgos, Areopolis, Vransko, Brijoni and Porto Rose) were already analyzed for the COI/tRNA^{Leu}/COII fragment in Luchetti *et al.* (2004a) (table 1).

At each sampling site, two individuals were sequenced at two mitochondrial loci, and 8–10 specimens were genotyped for both microsatellites and I-SINE markers (with the exception of the Estagel population, for which only four individuals were available). Total DNA was extracted from the head of each termite using the CTAB protocol (Doyle & Doyle, 1987).

Mitochondrial markers

A portion of 303 bp of the control region (CR), homologous to the region surveyed by Jenkins *et al.* (1999) for *Reticulitermes* taxa, and a 912 bp fragment, spanning from the 3' end of the cytochrome oxydase sub. I (COI) to the cytochrome oxydase sub. II (COII) gene, including the tRNA^{Leu} in between, were PCR amplified with the primer pairs AT-KR (5'-GTG GCT ATA CCC ACT ATA AA-3')/TM-N-193 (5'-TGG GGT ATG AAC CAG TAG C-3'), and C1-J-2797 (5'-CCT CGA CGT TAT TCA GAT TAC C-3')/TK-N-3785 (5'-GTT TAA GAG ACC AGT ACT TG-3'), respectively.

PCR amplifications were performed in a $50\,\mu$ l reaction using *Taq* DNA polymerase PCR kit (Invitrogen), following

Localities		Haplotype			GenBank A.N.		
		CR	COI/tRNA ^{Leu} /COII	Combined	CR	COI/tRNA ^{Leu} /COII	
GREECE							
Crete							
	1 Toplou a	A1	B1	H1	FJ750486	AY954678	
	Toplou b	A2	B1	H2	GU931795		
	2 Kastelli a	A1	B2	H3		FJ750507	
	Kastelli b	A3	B1	H4	FJ750493		
Cyclades							
	3 Amorgos a	A1	B1*	H1			
	Amorgos b	A1	B1	H1			
Peloponnesus							
	4 Areopolis a	A1	B3*	H5		AY954674	
	Areopolis b	A1	B1*	H1			
BALKANS							
Croatia	5 Vranska a	A 4	B1 *	Цć	EI750400		
	Vransko h	Δ1	B1	H1	FJ750490		
	6 Brijoni a	A1	B1*	H1			
	Brijoni b	A2	B1	H2			
Clamania	211jona 2		21				
5100eniu	7 Porto Rosa a	45	B1*	Н7	EI750488		
	Porto Rose h	AJ A1	B1	H1	17750400		
	10100100500	111	DI	111			
Donincula							
1 спіпзиш	8 Viterbo a	Δ1	B4	H8		FI750513	
	Viterbold	A1	B5	H9		GU931797	
	9 Feniglia a	A1	B1	H1		00/01//	
	Feniglia b	A6	B1	H10	GU931793		
Sardinia	C						
ourunna	10 Rena Majore a	A7	B6	H11	FI750495	FI750510	
	Rena Majore b	A1	B7	H12		GU931798	
	11 Geremeas a	A8	B8	H13	FJ750491	FJ750506	
	Geremeas b	A9	B6	H14	GU931796	-	
	12 Portoscuso a	A7	B9	H15		FJ750511	
	Portoscuso b	A10	B9	H16	FJ750497		
FRANCE							
Corsica							
	13 Calvi a	A11	B10	H17	FJ750494	FJ750509	
	Calvi b	A1	B6	H18			
Continental							
	14 Marsiglia a	A12	B11	H19	FJ750483	FJ750504	
	Marsiglia b	A13	B11	H20	GU931794		
	15 Port de Bouc a	A14	B12	H21	FJ750500	FJ750515	
	Port de Bouc b	A13	B13	H22		FJ750502	
	16 Avignone a	A14	B14	H23		FJ/50503	
	Avignone b	A14	B15 B12	H24 H25		GU931799	
	17 Douidon a Boulhon b	A12	D13 B14	п23 Ц94		CI 1021800	
	18 Estagel a	A13	B13	H25		60731000	
	Estagel h	A13	B13	H22			
	Louiger 2		210				

Table 1. Sampling sites, scored haplotypes and GenBank accession numbers for each mitochondrial sequence.

Numbers before each population are the same as for fig. 1. Asterisks indicate sequences drawn from Luchetti et al. (2004a).

the kit protocol. The amplification conditions were: 30 cycles of denaturation at 94°C for 30", annealing at 48°C for 30" and extension at 72°C for 30", plus an initial denaturation at 94°C for 5' and a final extension at 72°C for 7'. The amplified products were purified with the Wizard SV Gel and PCR Clean-up System kit (Promega, Milan, Italy) and both strands were sequenced at Macrogen Inc. (Korea). GenBank accession numbers are reported in table 1.

Microsatellites markers

A microsatellite-enriched genomic library was obtained using the F.I.A.S.CO. protocol (Zane *et al.*, 2002; Velonà *et al.*, 2009).

Microsatellite loci were identified using the TANDEM REPEAT FINDER software (Benson, 1999). PCR primers were then designed using the PRIMER3 online software (Rozen &

Locus	Primer Sequence	Motif	T _m	GenBank A. N.	
C22	F: 5'-TCCCAAAAATCCTTGCGTAT-3'	(TG) ₁₀	54°C	FJ750516	
C24	F: 5'-TAGCATCGGGAATGGACTTT-3'	$(AC)_2AG(AC)_6$	54°C	FJ750517	
D17	R: 5'-GAGGTACACTTGGCCGTCTT-3' F: 5'-GAGTTGGAAGACCCAAGCTG-3'	(AC) _c	54°C	FI750518	
2	R: 5'-GGCCCCTTTATGAATGTTTTC-3'			T J , 00010	
D30	F: 5'-CTGTGGGAAAAAGGATGACG-3' R: 5'-AGCTGACTGGATCTGCCACT-3'	$(CA)_2 GA(CA)_6$	56°C	FJ750519	
D52	F: 5'-TGAGAGCATGGAGTGGTGAG-3' R: 5'- CCGGAGGCAACAGACTAAAT-3'	(TG) ₁₀	54°C	FJ750520	

Table 2. List of primers for the five microsatellite loci here isolated.

Motifs, annealing temperatures and Genbank accession numbers of sequenced alleles are also given.

Skaletsky, 2000). Each primer pair was optimized for PCR amplification by testing over a range of annealing temperatures. The amplification conditions were 94° C for 5' for the initial denaturation, 35 cycles at 94° C for 30", annealing for 30" (see in table 2 the temperature utilized for each locus) and 72° C for 30", plus a final extension at 72° C for 7'.

The 10µl PCR reactions included 8 ng of genomic DNA, 10µM of each primer, 1.5 mM MgCl₂, 200µM dNTPs, 10mM of buffer $10 \times$ (Invitrogen kit), 1µl BSA 0.2% and 1U of *Taq* polymerase (Invitrogen). Genotyping was performed in a Beckman CEQ8000, using 5'-labelled forward primers (Sigma).

Inter-SINE markers

This method is based on the PCR amplification of sequences comprised between two copies of a SINE retrotransposon. The SINE used here was isolated in the subterranean termite species Reticulitermes lucifugus, and its presence was also confirmed within the Kalotermitidae family, in both Cryptotermes secundus and K. flavicollis (Luchetti, 2005; Luchetti & Mantovani, 2009). Two primers were tested, TaF (5' - AGT GGC CGT GCG GTC TAA G - 3') and TaR (5'- CGT CGC AGA GAC CTC TAC CTG - 3'). The first anneals at the 5' end of the SINE, priming upstream, the second anneals within the SINE, priming downstream. Only the *Ta*F primer generated a repeatable fingerprinting pattern, using the following amplification conditions: 94°C for 5' for the initial denaturation, 35 cycles at 94°C for 30", 42°C for 30" and 72°C for 30", plus a final extension at 72°C for 7'. The 10µl PCR reactions included 8ng of genomic DNA, 10µM of primer, 1.5 mM MgCl₂, 200 µM of dNTPs, 10 mM of buffer 10× and 1U of Taq polymerase (Invitrogen). PCR products were resolved in 2% agarose gels in a TAE buffer 1 × and bands were used to create a presence (1)/absence (0) matrix.

Statistical analyses

Sequences were aligned with CLUSTALW algorithm implemented in MEGA4 (Tamura *et al.*, 2007); this software was also used to calculate the number of variable sites.

The Shimodaira-Hasegawa test (Shimodaira & Hasegawa, 1999) was performed with PAUP* 4.0b (Swofford, 2001) to determine if the two isolated mitochondrial fragments could be analyzed in a combined dataset; significance of P was computed after 1000 replicates. The test resulted significant

(P < 0.05), impeding the use of the combined dataset for the phylogenetic analysis. So, through the software MODELTEST 3.06 (Posada & Crandall, 1998), we determined the best substitution model for both the control region (TrN+ Γ , Γ =1.6368) and the COI/tRNA^{Leu}/COII (TIM+I, I=0.7794).

The Bayesian phylogenetic analysis was carried out using MRBAYES 3.01 (Huelsenbeck & Ronquist, 2001), partitioning the total dataset and setting the ML parameters (lset) as follows: for the control region nst=6, rates=gamma and for the COI/tRNA^{Leu}/COII nst=6, rates=propinv. The Markov Chain Monte Carlo process was set so that two runs of four chains ran simultaneously for 10⁶ generations, when the two runs converged onto stationary distribution (standard deviation of split frequencies<0.01); trees were sampled every 100 generations, for a total of 10,001 trees. Burnin was set to 25% of sampled tree, so that a consensus tree was calculated on 7501 trees. A parsimony network was inferred using TCS 1.21 (Clement *et al.*, 2000), with gaps considered as 5th state characters. A pairwise divergence matrix was calculated using PAUP* (Swofford, 2001).

For microsatellite markers, the presence of null alleles was tested with MICROCHECKER 2.2.1 (Van Oosterhout *et al.*, 2004). The allele number, allelic richness and Weir and Cockerham *F*-statistics were calculated with FSTAT 2.9.3.2 (Goudet, 2002). Expected and observed heterozygosities were computed with GENETIX 4.05 (Belkhir *et al.*, 2003). The software GENEPOP 1.2 (Raymond & Rousset, 1995) was utilized for linkage disequilibrium analysis and to calculate the probability to fit to the Hardy-Weinberg equilibrium (HWE).

For the analysis of Inter-SINE (I-SINE) markers, we calculated the Shannon Index using FAMD 1.108b (Schlüter & Harris, 2006) and the genetic diversity using HICKORY 1.1 (Holsinger *et al.*, 2002) as measures of variability within populations. HICKORY uses a Bayesian approach without assuming Hardy-Weinberg equilibrium. The analysis was performed first by testing all the models implemented in the software. To select the best model, we used the Deviance Information Criterion (DIC: Spiegelhalter *et al.*, 2002), an analogue of the Akaike Information Criterion. The model with the lowest DIC value was chosen as the best one. The same analysis evaluated inbreeding (f) and differentiation (θ) coefficients.

For both microsatellite and I-SINE markers, the multivariate dataset was reduced into a two-dimensional space through a principal components analysis (PCA) performed with PAST 1.76 (Hammer *et al.*, 2001).



Fig. 2. Bayesian analysis performed on mitochondrial markers. Numbers indicate bayesian posterior probability values. Capital letters indicate the three lineages identified (A–B–C) and the unresolved polytomy (D).

Results

Mitochondrial DNA

The CR amplified fragment showed 15 variable sites out of 303, leading to 14 haplotypes. In the COI/tRNA^{Leu}/ COII fragment, 16 out of the COI 178 bp were variable (six at the first, three at the second and seven at the third codon position), while the COII sequence showed 32 variable sites out of 662 bp (six at the first, five at the second and 21 at the third codon position); the tRNA^{Leu}, 72 bp long, showed only one variable site across its length. The total number of haplotypes for the COI/tRNA^{Leu}/COII fragment is 16.

The alignment (CR+COI/tRNA^{Leu}/COII) showed 64 variable sites in 1215 bp, leading to 26 haplotypes. Only the two individuals of the Amorgos sample shared the same haplotype H1, which is also the most common combined haplotype present in six other populations, spanning from Greece to Croatia and peninsular Italy (table 1).

The Bayesian phylogenetic analysis (fig. 2) showed three distinct clades: the first (A) includes all the specimens from continental France, the second clade (B) embodies the specimens from Corsica and Sardinia, except Portoscuso (Sardinia), which constitutes the third clade (C). All the specimens from the area spanning from Crete to the peninsular Italy constitute an unresolved polytomic group (D).

The haplotype parsimony network, built only on the more variable COI/tRNA^{Leu}/COII fragment, partially agrees with the phylogenetic inference (fig. 3). The analysis splits the haplotypes into two networks; the first includes all the continental French samples (B11–B16), and the second includes all the other colonies. The polytomy seen in the phylogenetic tree (D) is a star-like network connected, albeit with a certain number of missing haplotypes, both to the Portoscuso lineage (B9) and to the much more differentiated Corsica-Sardinian clade (B6–B8, B10).

The genetic diversity within each of the four identified groups is up to 0.7% (average = 0.2%) and does not overlap the range of between-group distances (fig. 4). The highest values of divergence are found in the comparison between the continental France lineage and the Corsica-Sardinia clade but Portoscuso (from 2.5% to 3.2%). Portoscuso haplotypes are less differentiated from the peninsular Italy, the Balkans and the Greece populations (0.5%; square in fig. 4) than from the Corsica/Sardinian (1.3%; circle in fig. 4) and the continental France ones (1.8%; diamond in fig. 4; see also table S1 in supplementary materials).

Microsatellites markers

The enriched library screening produced a total of 70 positive clones, but only nine of them allowed primer



Fig. 3. Parsimony network built on COI/tRNA^{Leu}/COII sequences. Haplotype designation is as in table 1; the small lines across branch connections represent possible missing/ideal haplotypes. Dotted branch connection is not supported by parsimony 95% limit. Circle magnitudes are proportional to scored haplotype frequency.



Fig. 4. Histogram representing the frequency of pairwise sequences divergence among analyzed samples. Each bin represents the 0.1% of sequence divergence. Within group variability – white bars. Inter-group variability – light grey: peninsular Italy/Balkans/Greece polytomy (D) vs Corsica/Sardinia (clade B); dark grey: peninsular Italy/Balkans/Greece polytomy (D) vs Continental France (clade A); black: continental France (clade A) vs Corsica/Sardinia (clade B).

Indicators represent average divergence values \pm standard deviation (horizontal bars) between Portoscuso (clade C) and \Box , peninsular Italy/Balkans/Greece polytomy; \bigcirc , Corsica/Sardinia; \Diamond , Continental France.

design, and five were found to be polymorphic after population PCR assay. In particular, locus C22 was polymorphic in all populations, while C24 in only half of them (table 3). The linkage-disequilibrium analysis did not show any significant association between loci (see table S2 in supplementary materials).

	$\begin{array}{c} \hline \\ C22 \\ \hline \\ A/A_C \\ H_O/H_E \end{array}$	$\begin{array}{c} \underline{C24} \\ \underline{CC} \\ \underline{A}_{C} \\ \underline{A}_{E} \\ \underline{H}_{O} \\ \underline{H}_{E} \end{array}$	D17 A/A _C H _O /H _E	D30 A/A _C H _O /H _E	D52 A/A _C H _O /H _E	ALL LOCI N/A _C H _O /H _E	S.I. ±S.D.	Hs
								±S.D.
1 Toplou	4/2.624 0.333/0.653*	1/1.000	2/1.222 0.111/0.105	1/1.000	2/1.883 0.400/0.480	10/1.546 0.169/0.247	1.160 0.01	0.318 0.02
2 Kastelli	3/1.776 0.222/0.364	3/2.145 0.286/0.541	1/1.000	2/1.368 0.200/0.180	4/2.602 0.500/0.675*	10/1.778 0.242/0.352	0.974 0.01	0.223 0.02
3 Amorgos	2/1.908 0/0.494**\$	2/1.924 0/0.480*	1/1.000	2/1.857 0/0.444*	4/2.493 0.300/0.650*§	10/1.836 0.060/0.414***	0.793 0.03	0.174 0.03
4 Areopolis	4/2.460 0/0.617***\$	1/1.000	1/1.000	2/1.833 0/0.444**\$	2/1.624 0.400/0.320	10/1.583 0.080/0.276***	0.280 0.02	0.113 0.03
5 Vransko	3/2.444 0.6/0.645	2/1.883 0/0.480**\$	1/1.000	3/1.627 0.333/0.290	3/2.409 0.600/0.635	10/1.873 0.307/0.410	0.850 0.02	0.172 0.02
6 Brijoni	2/1.222 0.111/0.105	1/1.000	2/1.906 0.900/0.495*	2/1.876 0/0.469**\$	4/2.210 0.600/0.530	10/1.643 0.322/0.320	0.816 0.02	0.191 0.02
7 Porto Rose	2/1.908 0.667/0.494	3/2.072 0.200/0.535*§	3/2.182 0.833/0.542	1/1.000	4/2.697 1.000/0.680	10/1.972 0.540/0.450	0.300 0.001	0.122 0.03
8 Viterbo	2/1.368 0.200/0.180	1/1.000	1/1.000	2/1.509 0.300/0.255	2/1.789 0.571/0.408	10/1.333 0.214/0.169	0.835 0.03	0.182 0.02
9 Feniglia	3/1.915 0.400/0.455	2/1.833 0/0.480**\$	2/1.845 0.300/0.455	3/1.990 0.200/0.460*	8/3.431 0.800/0.855	10/2.213 0.340/0.539**	0.783	0.187
10 Rena Majore	2/1.368 0.200/0.180	1/1.000	2/1.881 0.778/0.475	2/1.368 0.200/0.180	3/2.351	10/1.594 0.436/0.289	0.544	0.110
11 Geremeas	2/1.673 0 444/0 346	3/2.557 0/0.625*	2/2.000	5/2.134 0.300/0.485*	3/2.000	10/2.073 0 199/0 472*	0.790	0.157
12 Portoscuso	2/1.222 0.111/0.105	2/1.505	2/1.876 0 250/0 469	3/2.170	2/1.576 0 333/0 279	10/1.670 0 239/0 325	0.600	0.127
13 Calvi	3/2.147 0.500/0.540	1/1.000	2/1.450 0/0.219	3/2.089 0.300/0.545	6/3.104 0.750/0.781	10/1.958 0.310/0.417	0.288	0.095
14 Marsiglia	2/1.765 0.556/0.401	2/1.923 0.750/0.500	2/1.405 0/0.197	2/1.200 0.100/0.095	2/1.816 0.625/0.430	10/1.622 0.406/0.324	0.854	0.184
15 Port de Bouc	2/1.912 0.875/0.492	1/1.000	2/1.929 0.750/0.469	1/1.000	2/1.786 0.500/0.375	8/1.524 0.425/0.267	0.682	0.146
16 Avignone	3/2.442 0.889/0.642	1/1.000	2/1.286 0.143/0.134	1/1.000	1/1.000	9/1.300 0.206/0.155	0.477	0.104
17 Boulbon	2/1.368	4/2.590 0/0.656***8	4/2.580	2/1.368	5/2.809 0 125/0 711***8	10/2.143	0.575	0.151
18 Estagel	2/2.000 1.000/0.500	1/1.000	1/1.000	1/1.000	1/1.000	4/1.200 0.200/0.100	0.000	0.115 0.03

Table 3. Number of alleles (A), allelic richness (A_C), number of individuals (N), observed (H_O)/expected (H_E) heterozygosity per locus and per population over all loci are given for the five microsatellite loci.

§, possible presence of null alleles; *, P<0.05; **, P<0.01; ***, P<0.001. Shannon Index (S.I.), genetic diversity (H_S) and their standard deviations are computed on I-SINE marker.

Five populations (Feniglia, Geremeas, Portoscuso, Marsiglia and Boulbon) were polymorphic at all loci, while Estagel and Avignone samples were polymorphic only at one (C22) or two (C22, D17) loci, respectively (table 3). Allelic richness values per population ranged from 1.000 to 3.431; the range was 1.200–2.213 over all loci per population. On the whole, the number of alleles per locus ranged from eight (C24 and D17) to 13 (D52), with an average allele number per locus of 9.8 (not shown). Considering the total number of loci analyzed over all populations, 29% of the polymorphic loci (20 out of 68) showed significant deviations from Hardy-Weinberg equilibrium (HWE); null alleles may explain 55% (11 out of 20) of these deviations. Considering all loci, five populations (Amorgos, Areopolis, Feniglia, Geremeas and Boulbon) did not match the HWE (table 3).

The Weir and Cockerham *F*-statistic showed a relatively low F_{IS} (0.285; P < 0.001; 95% confidence interval = 0.123–0.536) and a high F_{ST} (0.376; P < 0.001; 95% confidence interval = 0.259–0.461); this is reflected by significant F_{ST} in the majority of pairwise population comparisons (see table S3 in supplementary materials). In the PCA analysis (axis 1+axis 2=88.16% of the total variance), colonies from peninsular Italy, the Balkans and Greece cluster together, as well as Portoscuso sample with the other colonies from Sardinia and Corsica; the continental France samples were the most differentiated (fig. 5a).

Inter-SINE markers

The 20 loci detected were polymorphic throughout the entire dataset. The model choice analysis indicated the 'full model' as the best one to describe our dataset (DIC = 563.944).

Shannon index ranged between 0 (Estagel) and 1.160 (Toplou), while the genetic diversity range is 0.095 (Calvi)– 0.318 (Toplou; table 3). The inbreeding coefficient analysis showed a low value (f=0.168; 95% confidence interval= 0.006–0.541), while the inter-population divergences were higher (θ =0.418; 95% confidence interval=0.373–0.463). The PCA analysis (axis 1+axis 2=87.01% of the total variance) depicts a picture highly congruent with the one obtained from the microsatellite data. Samples from Greece, the Balkans and peninsular Italy form one group, and the samples from Portoscuso and Corsica-Sardinia form another one; the continental France colonies are the most divergent (fig. 5b).

Discussion

The very first mtDNA analysis of *K. flavicollis* (Luchetti *et al.*, 2004a) did not reveal any structuring of COI/tRNA^{Leu}/ COII haplotypes over a geographic area comprising peninsular Italy, Sicily, the Balkans, the Peloponnesus and the Cyclades. The results here presented, obtained from the same mitochondrial marker with the addition of the CR fragment, are in line with this previous analysis. Yet, the wider geographic area surveyed showed a more defined structuring due to the divergent and well-defined lineages present in Corsica/Sardinia, Portoscuso and continental France.

The range of within-group variability is, in all comparisons, separated by a large gap from the between-group divergence range; this recalls a 'barcoding' gap (Hebert *et al.*, 2004). Even considering the limits that the evolutionary history of each taxon imposes on these threshold estimates, the mean within- vs between-group divergences here observed are well above the conventional DNA barcode gap (Hebert *et al.*, 2004). With respect to the Hebert *et al.* (2004) dataset, our estimates are based on fewer samples and on different mtDNA genes; nevertheless, present values suggest that the here scored mtDNA variability could not represent a mere population-level divergence. With respect to our analysis, comparable results can also be found in other insect groups, such as Hemiptera (Žurovcová *et al.*, 2010) and Lepidoptera (Hajibabaei *et al.*, 2006).

A high level of differentiation is also supported by the microsatellite and I-SINE data, which suggest a clear-cut differentiation at the regional scale, with the French continental colonies being the most divergent. The existence of different entities in France and Italy was already hypothesized on physiological grounds (Luscher, 1956) and present molecular data support this possibility. Still other geographic taxa were suggested based on soldier morphometry, reproductive biology, and crossing experiments (Springhetti, 1967; Rossi & Springhetti, 1983); these data evidenced a significant divergence of Sardinian and Corsican colonies. Accordingly, the data presented here indicate that the Corsican and Sardinian colonies constitute a separate lineage that is differentiated from the continental France and peninsular Italy colonies. A more complex situation emerges when considering the Portoscuso sample. In our mitochondrial phylogeny, it seems to represent a separated lineage, but the level of divergence in the Portoscuso vs peninsular Italy, Balkans and Greece population comparison is low and falls in the within-species range. On the other hand, both nuclear markers define a strict relationship between Portoscuso and the Corsica-Sardinia lineage. On the whole, the data could be explained by a recent introduction of this population from peninsular Italy, followed by outbreeding events with local demes.

The Mediterranean basin's fauna shows a high level of diversity in several hotspots, as well as a number of endemic taxa, owing to various paleogeographic/paleoclimatic events. The range of sequence divergence observed among K. flavicollis clades points to roughly contemporary cladogenetic events. Quaternary climatic oscillations have been one of the major forces driving taxa evolution and distribution (Hewitt, 2004; Schmitt, 2007); during the glacial periods, Mediterranean peninsulas provided refugia for a number of organisms, while successive climate warming allowed northward recolonization. From present results, K. flavicollis colonies from the Balkans and peninsular Italy do not show the expected divergence, as populations lack mitochondrial lineage break across the entire area; only nuclear data suggest some degree of divergence. This agrees with the pattern observed, for example, in Maniola jurtina where the eastern European lineage spans from Italy to Balkans (Schmitt, 2007). Therefore, during the last glacial phase, an extensive gene flow within the Adriatic-Pontic area is suggested to have occurred, while the most western taxa remained isolated (Schmitt et al., 2005a). Data here presented may reflect a comparable situation.

The continental French clade may represent the leading edge of a northward recolonization from a glacial refugium. Since the scored genetic divergence (average $\sim 1.8\%$) excludes a close affinity with the Italian peninsular clade, the recolonization route may have initiated from an Iberian refugium. For example, this route has been followed by another European termite species, *Reticulitermes banyulensis* (Kutnik *et al.*, 2004) and by other strictly Mediterranean insects, such as the stick-insect *Leptynia hispanica* (Ghiselli *et al.*, 2007) or the



AXIS 1 (65.25%)

Fig. 5. Principal component analysis (PCA) built on (a) microsatellite dataset and (b) I-SINE dataset. In parenthesis the percentage of variance for each axis is reported.

butterfly *Polyommatus hispana* (Schmitt *et al.*, 2005b). However, the possibility of a refugial area in southern France, as hypothesized for other insects (Vasconcelos *et al.*, 2006; Wahlberg & Saccheri, 2007), cannot be ruled out. Further sampling in north-western Italy and, especially, in the Iberian peninsula will obviously highlight this point.

As for the origin of the Corsica-Sardinian colonies, except Portoscuso, the genetic divergence observed between this clade and the other eastern populations ($\sim 1.3\%$) well correlates with the Pliocene-Quaternary transgression, which, as argued for *Solatopupa* spp. (Ketmaier *et al.*, 2006), may have caused the speciation by vicariance. As proposed by Arbogast & Kenagy (2001), the comparative phylogeography approach is a good tool in detecting historical biogeography. The comparison of present results with those from the sympatric *Reticulitermes* spp. termites is of interest since, despite their largely overlapped distribution, these two groups show differences in present-day biodiversity. Taking into account the area considered in this paper, *Reticulitermes* spp. comprises eight species (Clément *et al.*, 2001; Luchetti *et al.*, 2004b, 2007; Uva *et al.*, 2004; Velonà *et al.*, 2010), while *K. flavicollis* shows only a half of lineages.

In the western part of the examined area, the biodiversity level of the two termite genera is similar. This does not necessarily mean that they share the same biogeographic history, but geological and paleoclimatic events might have similarly affected their diversity. *K. flavicollis* populations from Sardinia and Corsica are differentiated from the peninsular Italy ones, but with a level of divergence that could be associated with a relatively recent event. Differently, a speciation event may have originated in the same area, in a vicariance context, *Reticulitermes lucifugus corsicus* from *R. lucifugus lucifugus* around 9 million years ago (Velonà *et al.*, 2010).

On the other hand, east of Italy, K. flavicollis is genetically more homogeneous than Reticulitermes. The Aegean area is known as a hotspot of high biodiversity (Sfenthourakis & Legakis, 2001), mainly because of a complex paleogeographic history (Perissoratis & Conispoliatis, 2003); this also has been demonstrated in a number of animal systems (see for example Parmakelis et al., 2006 and references therein). Accordingly, *Reticulitermes* spp. show a consistent degree of diversification; in particular, known diversity hotspots, such as Peloponnesus, Cyclades Islands and Crete (Sfenthourakis & Legakis, 2001), show their own subterranean termite lineages (Luchetti et al., 2007; Velonà et al., 2010). At the same localities, K. flavicollis diversity is limited to a population level of differentiation. Therefore, in the eastern Mediterranean range, the two termite taxa do not share the same taxon radiation. The explanation is probably linked with the different colonization histories. In fact, K. flavicollis species could have colonized this area later with respect to Reticulitermes spp.; in this way, it was not processed by the geological/paleoclimatic events shaping eastern biodiversity (see Velonà et al., 2010 and references therein).

Considering our results and those obtained in a previous paper (Velonà *et al.*, 2010), it is possible to hypothesize that *K. flavicollis* species colonized recently the European area, and this event was followed by a diversification in at least three lineages.

Acknowledgements

We wish to thank Anne-Geneviève Bagnères and Simon Dupont for kindly providing samples from southern France, Megan Maurano for the linguistic revision and Federico Plazzi for statistical support. This work has been funded by Canziani donation and RFO (University of Bologna) to B.M.

Supplementary Material

The online tables can be viewed at http://journals. cambridge.org/ber

References

- Arbogast, B.S. & Kenagy, G.J. (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography* 28, 819–825.
- Austin, J.W., Szalanski, A.L., Uva, P., Bagnères, A.G. & Kence, A. (2002) A comparative genetic analysis of the subterranean termite genus *Reticulitermes* (Isoptera: Rhinotermitidae). *Annals of the Entomological Society of America* **95**, 753–760.
- Belkhir, K., Borsa, P., Chikki, L., Raufaste, N. & Bonhomme, F. (2003) GENETIX, logiciel sous Windows TM pour la génétique des populations. Available online at http:// www.genetix.univ-montp2.fr/genetix/genetix.htm.

- Benson, G. (1999) Tandem Repeat Finder: a program to analyze DNA sequences. Nucleic Acids Research 27, 573–580.
- Clement, M., Posada, D. & Crandall, K. (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9, 1657–1660.
- Clément, J.L., Bagnères, A.G., Uva, P., Wilfert, L., Quintana, A., Reinhard, J. & Dronnet, S. (2001) Biosystematics of *Reticulitermes* termites in Europe: morphological, chemical and molecular data. *Insectes Sociaux* 48, 202–215.
- Doyle, J.J. & Doyle, J.L. (1987) A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochemical Bulletin* 19, 11–15.
- Eggleton, P. & Tayasu, I. (2001) Feeding groups, lifetype and the global ecology of termites. *Ecological Research* 16, 941–960.
- Engel, M.S. & Krishna, K. (2004) Family-group names for termites (Isoptera). American Museum Novitates 2570, 1–31.
- Engel, M.S., Grimaldi, D.A. & Krishna, K. (2009) Termites (Isoptera): their phylogeny, classification, and rise to ecological dominance. *American Museum Novitates* 3650, 1–27.
- Ghiselli, F., Milani, L., Scali, V. & Passamonti, M. (2007) The Leptynia hispanica species complex (Insecta Phasmida): polyploidy, parthenogenesis, hybridization and more. Molecular Ecology 16, 4256–4268.
- **Goudet, J.** (2002) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2). Available online at http://www2.unil.ch/popgen/softwares/fstat. htm.
- Hajibabaei, M., Janzen, D.H., Burns, J.M., Hallwachs, W. & Hebert, P.D.N. (2006) DNA barcodes distinguish species of tropical Lepidoptera. PNAS 103, 968–971.
- Hammer, Ø., Harper, D. & Ryan, P. (2001) PAST– PAleontological STatisitics software package for education and data analysis. *Palaeontologia Electronica* 4, 9 pp.
- Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S. & Francis, C.M. (2004) Identification of birds through DNA barcodes. *PLoS Biology* 10, e312.
- Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London, Series B* 359, 183–195.
- Holsinger, K.E., Lewis, P.O. & Dey, D.K. (2002) A Bayesian approach to inferring population structure from dominant markers. *Molecular Ecology* 11, 1157–1164.
- Huelsenbeck, J.P. & Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17, 754–755.
- Inward, D., Vogler, A. & Eggleton, P. (2007) A comprehensive phylogenetic analysis of termites (Isoptera) illuminates key aspects of their evolutionary biology. *Molecular Phylogenetics* and Evolution 44, 953–967.
- Jenkins, T.M., Basten, C.J., Dean, R., Mitchell, S.E., Kresovich, S. & Forschler, B.T. (1999) Matriarchal genetic structure of *Reticulitermes* (Isoptera: Rhinotermitidae) populations. *Sociobiology* 33, 239–263.
- Ketmaier, V., Giusti, F. & Caccone, A. (2006) Molecular phylogeny and historical biogeography of the land snail genus *Solatopupa* (Pulmonata) in the peri-Tyrrhenian area. *Molecular Phylogenetics and Evolution* **39**, 439–451.
- Korb, J. (2007) Termites. Current Biology 17, R995-999.
- Korb, J. (2008) Termites, hemimetabolous diploid white ants? Frontiers in Zoology 2008, 5–15.
- Kostia, S., Ruohonen-Lehto, M., Väinölä, R. & Varvio, S.L. (2000) Phylogenetic information in inter-SINE and inter-SSR fingerprints of the artiodactyla and evolution of the bov-tA SINE. *Heredity* 84, 37–45.

- Kutnik, M., Uva, P., Brinkworth, L. & Bagnères, A.G. (2004) Phylogeography of two European *Reticulitermes* (Isoptera) species: the Iberian refugium. *Molecular Ecology* 13, 3099– 3113.
- Luchetti, A. (2005) Identification of a short interspersed repeat in the *Reticulitermes lucifugus* (Isoptera Rinothermitidae) genome. DNA Sequence 16, 304–307.
- Luchetti, A. & Mantovani, B. (2009) Talua SINE Biology in the Genome of the *Reticulitermes* Subterranean Termites (Isoptera, Rhinotermitidae). Journal of Molecular Evolution 69, 589–600.
- Luchetti, A., Bergamaschi, S., Marini, M. & Mantovani, B. (2004a) Mitochondrial DNA analysis of native European Isoptera: a comparison between *Reticulitermes* (Rhinotermitidae) and *Kalotermes* (Kalotermitidae) colonies from Italy and Balkans. *Redia* 87, 149–153.
- Luchetti, A., Trenta, M., Mantovani, B. & Marini, M. (2004b) Taxonomy and phylogeny of north mediterranean *Reticulitermes* termites (Isoptera, Rhinotermitidae): a new insight. *Insectes Sociaux* 51, 117–122.
- Luchetti, A., Marini, M. & Mantovani, M. (2005) Mitochondrial evolutionary rate and speciation in termites: data on European *Reticulitermes* taxa (Isoptera, Rhinotermitidae). *Insectes Sociaux* 52, 218–221.
- Luchetti, A., Marini, M. & Mantovani, B. (2007) Filling the European gap: Biosystematics of the eusocial system *Reticulitermes* (Isoptera, Rhinotermitidae) in the Balkanic Peninsula and Aegean area. *Molecular Phylogenetics and Evolution* 45, 377–383.
- Luscher, M. (1956) Hemmende und fordernde Faktoren bei der Entstehung der Ersatzgeschlechtstiere bei der Termite Kalotermes flavicollis Fabr. Revue Suisse Zoologie 63, 261–267.
- Marini, M. & Mantovani, B. (2002) Molecular Relationship among European Samples of *Reticulitermes* (Isoptera Rhinotermitidae). *Molecular Phylogenetics and Evolution* 22, 454–459.
- Nishihara, H. & Okada, N. (2008) Retrotransposons: genetic footprints on the evolutionary paths of life. pp. 201–225 in Murphy, W.J. (Ed.) Methods in Molecular Biology: Phylogenomics. Totowa, NJ, USA, Humana Press.
- Ohshima, K. & Okada, N. (2005) SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. *Cytogenetic and Genome Research* 110, 475–490.
- Parmakelis, A., Stathi, I., Chatzaki, M., Simaiakis, S., Spanos, L., Louis, C. & Mylonas, M. (2006) Evolution of Mesobuthus gibbosus (Brullé, 1832) (Scorpiones: Buthidae) in the northeastern Mediterranean region. Molecular Ecology 15, 2883–2894.
- Perissoratis, C. & Conispoliatis, N. (2003) The impacts of sea-level changes during latest Pleistocene and Holocene times on the morphology of the Ionian and Aegean seas (SE Alpine Europe). *Marine Geology* 196, 145–156.
- Posada, D. & Crandall, K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Prota, R. (1962) L'infestazione termitica in Sardegna. Bollettino dell'Istituto di Patologia del Libro XXI, 1–35.
- Raymond, M. & Rousset, F. (1995) GENEPOP (version 1.2): a population genetics software for exact test and ecumenicism. *Journal of Heredity* 86, 248–249.
- Rossi, R. & Springhetti, A. (1983) Morphometric research on soldiers of *Kalotermes flavicollis* Fabr. from Italy. *Annali dell'Università di Ferrara - Biologia* 3, 41–50.
- Rozen, S. & Skaletsky, H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. pp. 365–386 in

Misener, S., Kravetz, S.A. (*Eds*) *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Totowa, NJ, USA, Humana Press.

- Schlüter, P.M. & Harris, S.A. (2006) Analysis of multilocus fingerprinting data sets containing missing data. *Molecular Ecology Notes* 6, 569–572.
- Schmitt, T. (2007) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology* 2007, 4–11.
- Schmitt, T., Rober, S. & Seitz, A. (2005a) Is the last glaciation the only relevant event for the present genetic population structure of the meadow brown butterfly *Maniola jurtina* (Lepidoptera: Nymphalidae)? *Biological Journal of the Linnean Society* 85, 419–431.
- Schmitt, T., Varga, Z. & Seitz, A. (2005b) Are Polyommatus hispana and Polyommatus slovacus bivoltine Polyommatus coridon (Lepidoptera: Lycaenidae)? The discriminatory value of genetics in taxonomy. Organisms Diversity and Evolution 5, 297–307.
- Sfenthourakis, S. & Legakis, A. (2001) Hotspots of endemic terrestrial invertebrates in southern Greece. *Biodiversity and Conservation* 10, 1387–1417.
- Shafer, A.B.A. & Stewart, D.T. (2007) Phylogenetic relationships among Nearctic shrews of the genus *Sorex* (Insectivora, Soricidae) inferred from combined cytochrome *b* and inter-SINE fingerprint data using Bayesian analysis. *Molecular Phylogenetics and Evolution* 44, 192–203.
- Shimodaira, H. & Hasegawa, M. (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* 16, 1114–1116.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & van der Linde, A. (2002) Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series Series B 64, 483–689.
- Springhetti, A. (1967) Incroci tra reali di alcune popolazioni italiane di Kalotermes flavicollis Fabr. Annali dell'Università di Ferrara - Biologia III, 11–17.
- Swofford, D.L. (2001) PAUP*-Phylogenetic Analysis Using Parsimony (* and Other Methods), Ver 4b. Sunderland, MA, USA, Sinauer Associates.
- Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24, 1596–1599.
- Uva, P., Clément, J.L., Austin, J.W., Aubert, J., Zaffagnini, V., Quintana, A. & Bagnères, A.G. (2004) Origin of a new *Reticulitermes* termite (Isoptera, Rhinotermitidae) inferred from mitochondrial and nuclear DNA data. *Molecular Phylogenetics and Evolution* 30, 344–353.
- Van Oosterhout, C., Hutchinson, W., Wills, D. & Shipley, P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4, 535–538.
- Vasconcelos, T., Horn, A., Lieutier, F., Branco, M. & Kerdelhué, C. (2006) Distribution and population genetic structure of the Mediterranean pine shoot beetle *Tomicus destruens* in the Iberian Peninsula and Southern France. *Agricultural and Forest Entomology* 8, 103–111.
- Velonà, A., Luchetti, A., Scanabissi, F. & Mantovani, B. (2009) Genetic variability and reproductive modalities in European populations of *Triops cancriformis* (Crustacea, Branchiopoda, Notostraca). *Italian Journal of Zoology* **76**, 366–375.
- Velonà, A., Ghesini, S., Luchetti, A., Marini, M. & Mantovani, B. (2010) Starting from Crete, a phylogenetic re-analysis of the genus *Reticulitermes* in the Mediterranean area. *Molecular Phylogenetics and Evolution* 56, 1051–1058.

- Wahlberg, N. & Saccheri, I. (2007) The effects of Pleistocene glaciations on the phylogeography of *Melitaea cinxia* (Lepidoptera: Nymphalidae). *European Journal of Entomology* 104, 675–684.
- Zane, L., Bargelloni, L. & Patarnello, T. (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology* 11, 1–16.
- Žurovcová, M., Havelca, J., Stary, P., Věchtová, P., Chundelová, D., Jarošová, A. & Kučerová, L. (2010) 'DNA barcoding' is of limited value for identifying adelgids (Hemiptera: Adelgidae) but supports traditional morphological taxonomy. *European Journal of Entomology* 107, 147–156.