Phylogeography of *Psammodromus algirus* (Lacertidae) revisited: systematic implications

Joaquín Verdú-Ricoy[1,*], Salvador Carranza[2], Alfredo Salvador[1],

Stephen D. Busack[3], José A. Díaz[4]

Abstract. Relationships among *Psammodromus algirus* populations from the Iberian Peninsula and North Africa, including recently described *P. jeanneae* and *P. manuelae*, were estimated from mitochondrial DNA gene sequences. This enlarged data set confirmed the presence of two divergent eastern and western mitochondrial DNA lineages on the Iberian Peninsula, the distributions for which are separated by a narrow zone of contact across the centre of the Peninsula. Paratypes of *P. jeanneae* and topotypes of *P. manuelae* represent southern and northern clades of the western lineage, respectively, making *P. algirus* paraphyletic. This, together with the low level of allozymic and mitochondrial DNA substructuring within western populations, is not sufficient to retain *P. jeanneae* and *P. manuelae* as valid species, and we relegate them to the status of junior synonyms of *P. algirus*.

Keywords: Iberian Peninsula, lacertid lizards, mitochondrial DNA, *Psammodromus algirus*, *Psammodromus jeanneae*, *Psammodromus manuelae*.

Since its original designation as a species in 1758, *Psammodromus algirus* has been considered representative of a single, invariant, species throughout the Iberian Peninsula. Recently, however, Carranza et al. (2006) demonstrated that *P. algirus* in Iberia was actually comprised of eastern and western mtDNA clades. Working independently of Carranza et al., Busack and Lawson (2006) noted mtDNA differentiation and allozyme differentiation between northern and southern Iberian populations (2006: figs 2 and 3, respectively), and later described *P. manuelae* from Manzanares el Real (Madrid province) and *P. jeanneae* from 25.6 km NE of Facinas (Cádiz province) following a morphological analysis (Busack, Salvador and Lawson, 2006). In this analysis we utilize DNA sequences from seven mitochondrial genes and take advantage of a larger, more robust, specimen sample including the holotype and paratypes of *P. jeanneae* and topotypes of *P. manuelae* to revisit the phylogeography of Iberian *P. algirus.* This expanded data set allows us to infer genetic relationship among populations of *P. algirus*, *P. manuelae* and *P. jeanneae* throughout much of its current geographical range on Iberia and to revise the taxonomy of this group with special reference to the status of *P. manuelae* and *P. jeanneae*.

A total of 104 individuals from 68 populations, including most samples used by Carranza et al. (2006) and Busack and Lawson (2006), the holotype and one paratype of *Psammodromus jeanneae* (E232055 and 232056, respectively), topotypes of *P. manuelae* (E232060 and 232062) and 32 specimens from 16 previously unsampled populations (fig. 1), were sequenced. *Psammodromus h. hispanicus* from Encinasola (Huelva province) and *P. h. edwarsianus* from Sierra de Baza (Granada province) served as outgroups (following Carranza et al., 2006).

Total genomic DNA was extracted from ∼40 mg of tissue using the Qiagen BioSprint 15 DNA Kit® following the manufacturer's protocol. Resulting DNA was visually inspected after migration on agarose gels and quantified with a NanoDrop spectophotometer. Amplifications were performed in 50 μl of 1× reaction buffer, 2 mM MgCl$_2$, 0.4 μM each primer, 0.2 mM each dNTP, 1.25 U of Go-Taq Flexi DNA Polymerase (Promega), and 3 μl of previously extracted DNA (50-100 ng). A 709 bp fragment of

1 - Departamento de Ecología Evolutiva, Museo Nacional de Ciencias Naturales, CSIC, José Gutiérrez Abascal, 2, 28006 Madrid, Spain

2 - Instituto de Biología Evolutiva, CSIC-UPF, Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Spain

3 - North Carolina State Museum of Natural Sciences, 11 West Jones Street, Raleigh, NC 27601-1029, USA

4 - Departamento de Zoología y Antropología Física, Facultad de Biología, Universidad Complutense, E-28040 Madrid, Spain

*Corresponding author; e-mail: jvricoy@mncn.csic.es
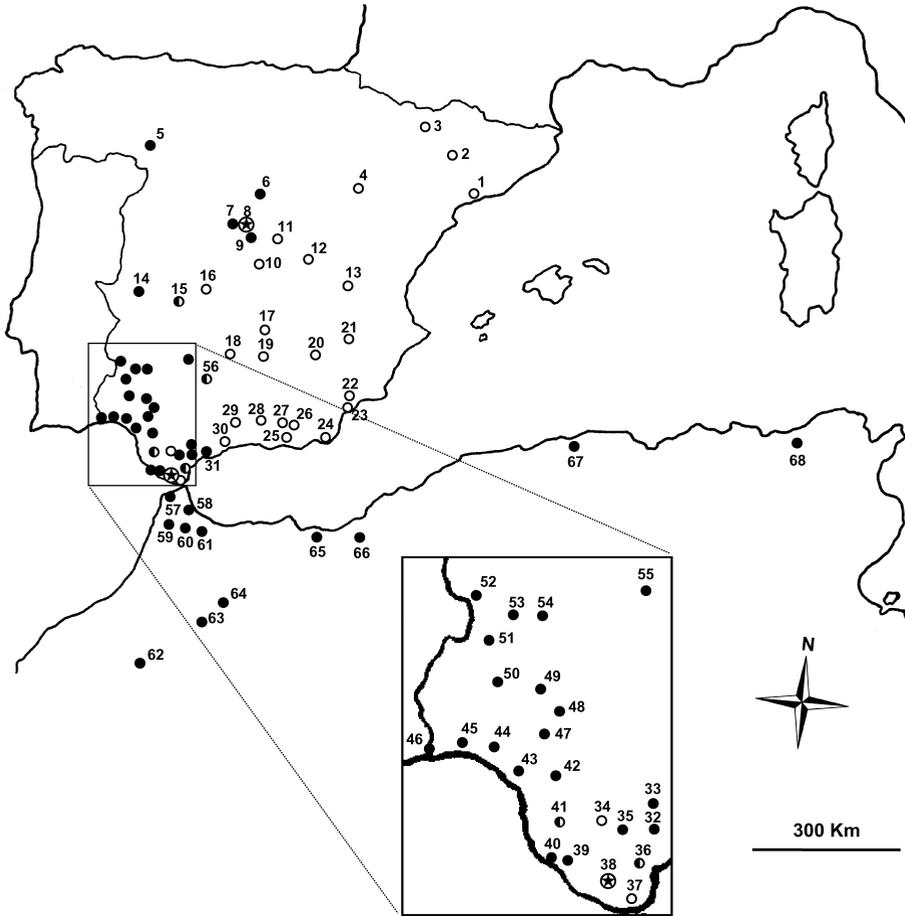
Figure 1. Localities sampled; solid symbols represent the western lineage, open symbols represent the eastern lineage and half-solid symbols indicate localities representing both lineages (type localities for *P. manuelae* [8] and *P. jeanneae* [38] highlighted with asterisk). Locality Key: <u>SPAIN</u>: TARRAGONA: (1) Tarragona. LERIDA: (2) Tartareu. HUESCA: (3) Ainsa. ZARAGOZA: (4) El Frasno. ZAMORA: (5) Cabañas de Tera. SEGOVIA: (6) Sotos de Sepúlveda. MADRID: (7) Navacerrada, (8) Manzanares el Real, (9) El Pardo, (10) Aranjuez. GUADALAJARA: (11) Pioz. CUENCA: (12) Torrejoncillo del Rey, (13) Hoces del Cabriel. CACERES: (14) Santiago del Campo, (15) Villuercas. TOLEDO: (16) Espinoso del Rey. CIUDAD REAL: (17) Valdepeñas, (18) Solana del Pino. JAEN: (19) Despeñaperros. ALBACETE: (20) Embalse Fuensanta, (21) Embalse de Bayco. MURCIA: (22) Morata, (23) Águilas. ALMERIA: (24) Cabo de Gata, (25) Sierra Gador, (26) Abrucena. GRANADA: (27) La Calahorra, (28) Cortijo del Ciprés, (29) Dehesa de los Montes. MALAGA: (30) Málaga, (31) Marbella, (32) Sierra Bermeja, (33) Genaguacil, (34) La Sauceda, (35) Río Hozgarganta. CADIZ: (36) Castellar de la Frontera, (37) Getares, (38) Facinas, (39) Barbate, (40) Caños de Meca, (41) Medina Sidonia. SEVILLA: (42) Lebrija, (47) Cañada de los Pájaros, (48) Gelves, (49) Gerena. HUELVA: (43) Matalascañas, (44) Bodegones, (45) El Portil, (46) Ayamonte, (50) Berrocal, (51) Linares de la Sierra. BADAJOZ: (52) Oliva de la Frontera, (53) Tentudia, (54) Pallarés. CORDOBA: (55) Doña Rama, (56) Virgen de la Cabeza. <u>MOROCCO</u>: (57) Tangier, (58) Ued Lau, (59) Chefchaouen, (60) Jebala, (61) Bab-Berret, (62) Beni-Mellal, (63) Boulemane, (64) Middle Atlas, (65) Berkana. <u>ALGERIA</u>: (66) Tlemcen, (67) Sidi Feredj. <u>TUNISIA</u>: (68) Ain Draham.

the fourth subunit of the NADH dehydrogenase mitochondrial gene (ND4) and adjacent tRNA^{His} (68 bp), tRNA^{Ser} (67 bp) and tRNA^{Leu} (48 bp) were amplified using primers ND4 and Leu (Arévalo, Davis and Sites, 1994). PCR consisted of 3-min pre-denaturing step at 94°C, followed by 35 cycles of denaturing for 30 sec at 94°C, primer annealing for 30 sec at 54°C and elongation for 40 sec at 72°C with a final

4-min elongation step at 72°C. Fragments of the mitochondrial cytochrome *b* gene (CytB, 300 bp), 12S rRNA (363 bp) and 16S rRNA (410 bp) were amplified using primers cytb1 and cytb2 (Palumbi, 1996), 12Sa and 12Sb (Kocher et al., 1989), and 16Sa and 16Sb (Palumbi, 1996), respectively. PCR consisted of a 5-min pre-denaturing step at 94°C, followed by 35 cycles of denaturing for 30 sec at 94°C, anneal-

ing primers for 45 sec at 48°C and elongation for one min at 72°C with a final 5-min elongation step at 72°C. PCR effectiveness was visually quantified after migration of PCR products on agarose gels. PCR products were purified by an ammonium acetate/ethanol cleaning process and sequenced using the ABI Prism Big Dye Terminator Cycle sequencing protocol in an ABI Prism 310 automated sequencer (Applied Biosystems).

All sequence chromatograms were edited with Sequencer (v. 4.2.2, Gene Codes). Once corrected, sequences were aligned independently for each gene with CLUSTALX (Thompson et al., 1997) under program default parameters (opening gap = 10; gap extension = 0.2) and visually inspected with Bioedit v.7.0.5 (Hall, 2005). Topological incongruence among partitions was tested using the incongruence length difference (ILD) test (Michkevich and Farris, 1981; Farris et al., 1994). In this test, 10,000 heuristic searches were carried out after removing all invariable characters from the data set (Cunningham, 1997). To test for incongruence among data sets, we also used a reciprocal 70% bootstrap proportion or a 95% posterior probability threshold (Mason-Gamer and Kellogg, 1996). Topological conflicts were considered significant if two different relationships for the same set of taxa were each supported. Results of all tests indicated that independent data sets were not incongruent (data not shown) and therefore a combined analysis involving 7 mitochondrial genes (ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB-12S-16S) was carried out. For many samples, however, it was not possible to amplify 12S and 16S genes and a smaller data set including only 5 genes (ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB) was also elaborated.
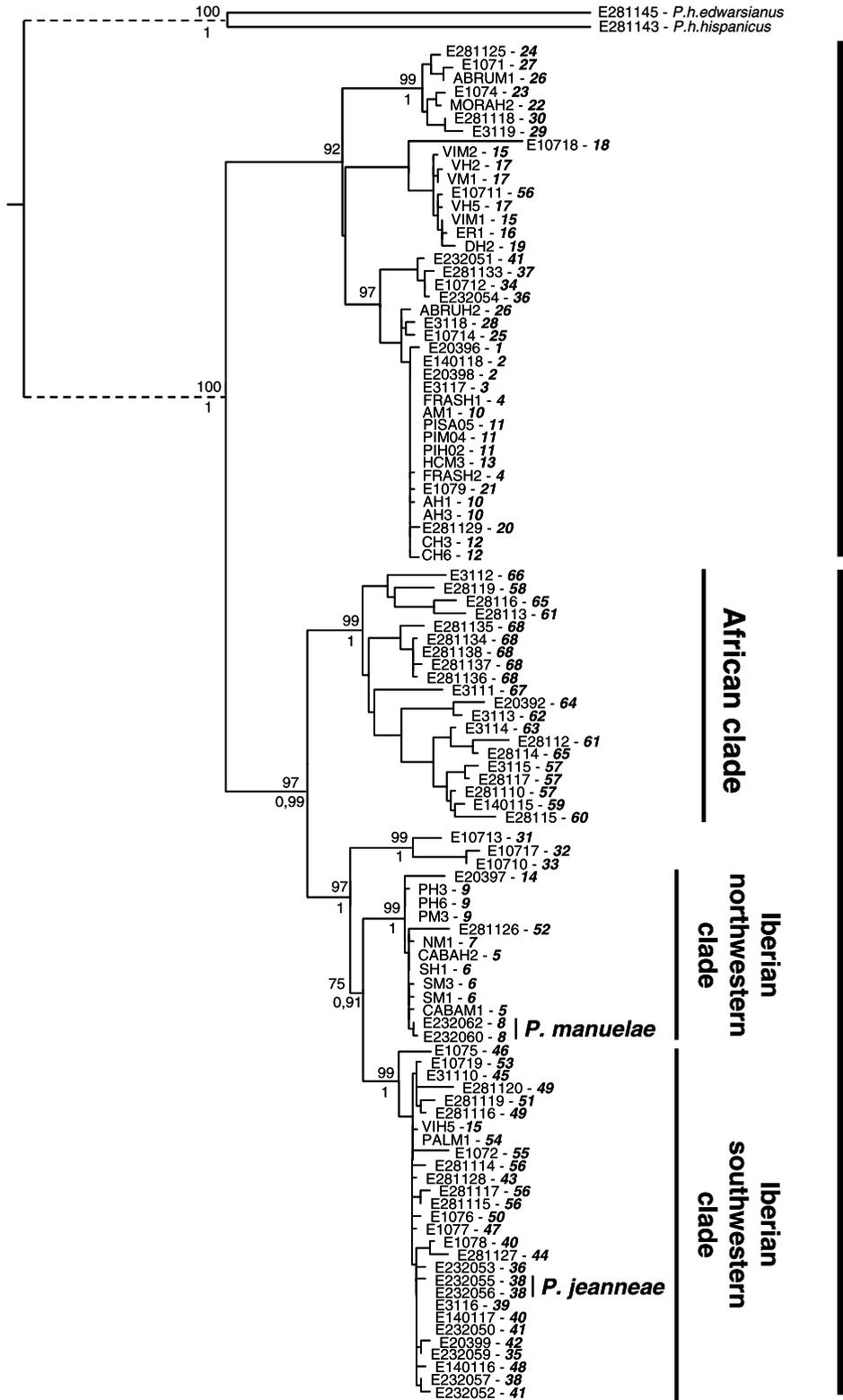
Phylogenetic trees were inferred using Maximum-Likelihood (ML; Felsenstein, 1981) and Bayesian methods. The most appropriate model of sequence evolution was determined with jModelTest v.0.1.1 (Posada, 2008) using the Akaike information criterion. In the Bayesian analyses each partition had its evolutionary model and these were: the HKY for tRNA^{His} and tRNA^{Leu}, HKY+G for tRNA^{Ser} and 16S, GTR+G for ND4 and 12S, and GTR+I+G for CytB. ML analyses were performed using PhyML version 2.4.3 (Guindon and Gascuel, 2003), with model parameters fitted to the data by likelihood maximization. In this case a single model of sequence evolution was selected for each concatenated data set (5 genes and 7 genes). In both cases the best model was the GTR + I + G. Reliability of the ML trees was assessed by bootstrap analysis with 1000 replications (Felsenstein, 1985). Bayesian analyses were performed with MrBayes version 3.0b4 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) for 2.5×10^6 generations, with a sampling frequency of 100 generations. After verifying that stationarity had been reached in terms of likelihood scores and parameter estimation, the first 5000 trees for the two data sets were discarded from both runs and independent majority-rule consensus trees were generated from the remaining (post burn-in) trees. The frequency of any particular clade of the consensus tree represents the posterior probability of that node (Huelsenbeck and Ronquist, 2001); onl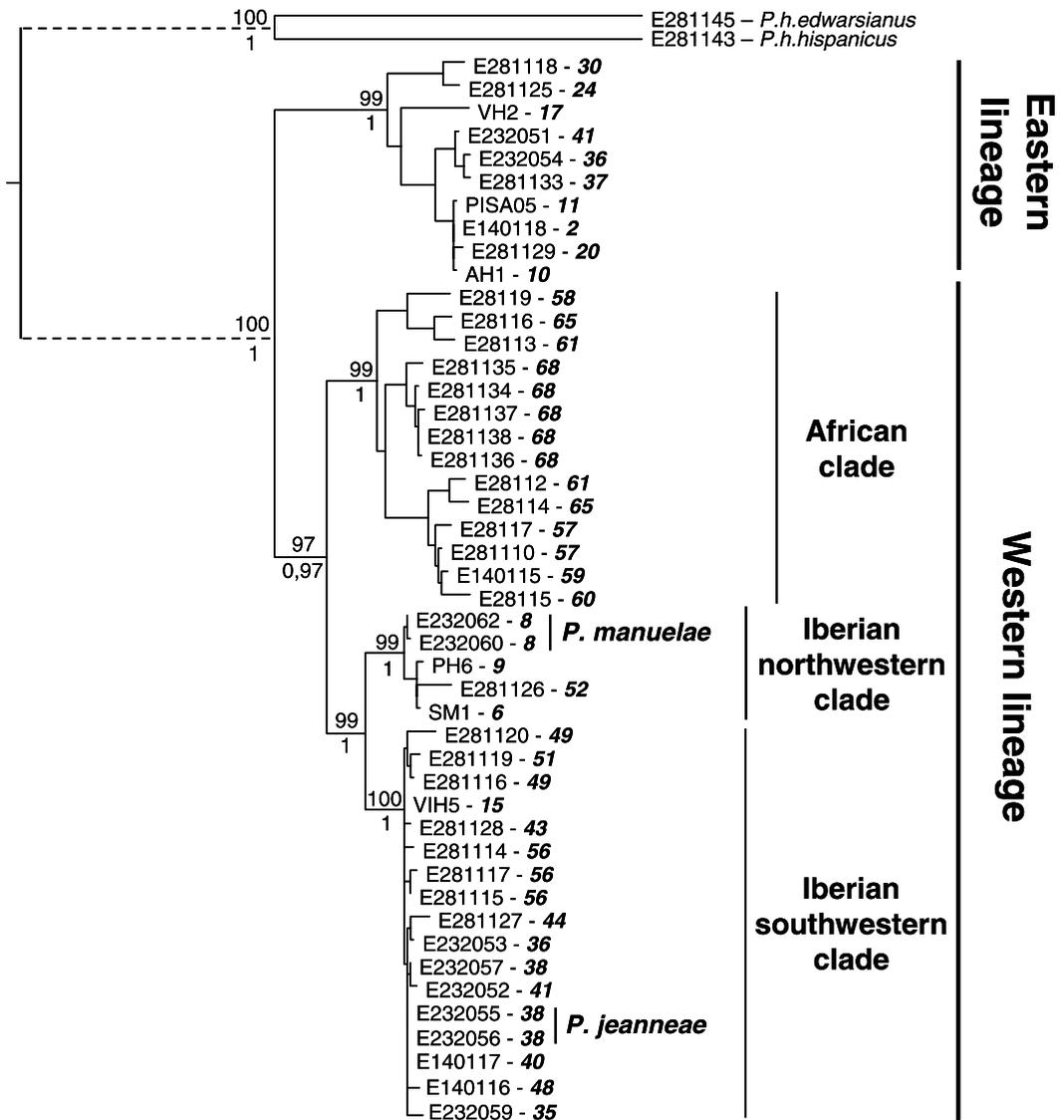y values equal to or greater than 95% were considered to indicate that nodes were significantly supported (Wilcox et al., 2002).

The ILD test indicated that the 7 gene partitions (ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB-12S-16S) were not incongruent ($P = 0.23$); analyses of independent partitions confirmed there were no topological conflicts (Mason-Gamer and Kellogg, 1996) and two independent data sets of mitochondrial fragments were combined for further analysis. The first data set included 1192 bp for ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB from 104 individuals in which 318 bp were variable and 236 bp were parsimony-informative, and the second was comprised of 1965 bp for ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB-12S-16S from 46 individuals in which 459 bp were variable and 319 bp were parsimony-informative. All sequence data not currently available in GenBank will be added upon publication (see figs 2 and 3 for specimen identification information).

Results of Maximum-Likelihood (ML) and Bayesian analyses are illustrated in fig. 2 (ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB) and fig. 3 (ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB-12S-16S). Each method and dataset produced trees with very similar topologies. Log-likelihood values of the trees obtained by ML for ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB and ND4-tRNA^{His}-tRNA^{Ser}-tRNA^{Leu}-CytB-12S-16S combinations were −5824.03 and −6584.01, respectively.

These data confirm that *Psammodromus algirus* consists of two well-supported, reciprocally monophyletic, mitochondrial lineages (Carranza et al., 2006); an eastern lineage confined to Iberia, and a western lineage present in both Iberia and North Africa (figs 1-3). This analysis, which includes a larger and more geographically-representative sample of individuals and a more robust mitochondrial sampling than Carranza et al. (2006), strongly supports the African clade as being sister to the western lineage on Iberia. Iberian representatives of the western lineage are partitioned into three groups: a basal clade restricted to three

Eastern lineage

100
1
E281145 - *P.h.edwarsianus*
E281143 - *P.h.hispanicus*

99
1
E281125 - *24*
E1071 - *27*
ABRUM1 - *26*
E1074 - *23*
MORAH2 - *22*
E281118 - *30*
E3119 - *29*

92
E10718 - *18*

VIM2 - *15*
VH2 - *17*
VM1 - *17*
E10711 - *56*
VH5 - *17*
VIM1 - *15*
ER1 - *16*
DH2 - *19*

97
E232051 - *41*
E281133 - *37*
E10712 - *34*
E232054 - *36*

ABRUH2 - *26*
E3118 - *28*
E10714 - *25*
E20396 - *1*
E140118 - *2*
E20398 - *2*
E3117 - *3*
FRASH1 - *4*
AM1 - *10*
PISA05 - *11*
PIM04 - *11*
PIH02 - *11*
HCM3 - *13*
FRASH2 - *4*
E1079 - *21*
AH1 - *10*
AH3 - *10*
E281129 - *20*
CH3 - *12*
CH6 - *12*

100
1

African clade

99
1
E3112 - *66*
E28119 - *58*
E28116 - *65*
E28113 - *61*

E281135 - *68*
E281134 - *68*
E281138 - *68*
E281137 - *68*
E281136 - *68*

E3111 - *67*
E20392 - *64*
E3113 - *62*
E3114 - *63*
E28112 - *61*
E28114 - *65*
E3115 - *57*
E28117 - *57*
E281110 - *57*
E140115 - *59*
E28115 - *60*

97
0,99

Western lineage

Iberian northwestern clade

99
1
E10713 - *31*
E10717 - *32*
E10710 - *33*

97
1
E20397 - *14*

99
1
PH3 - *9*
PH6 - *9*
PM3 - *9*
E281126 - *52*
NM1 - *7*
CABAH2 - *5*
SH1 - *6*
SM3 - *6*
SM1 - *6*
CABAM1 - *5*
E232062 - *8* | *P. manuelae*
E232060 - *8*

75
0,91

Iberian southwestern clade

99
1
E1075 - *46*
E10719 - *53*
E31110 - *45*
E281120 - *49*
E281119 - *51*
E281116 - *49*
VIH5 - *15*
PALM1 - *54*
E1072 - *55*
E281114 - *56*
E281128 - *43*
E281117 - *56*
E281115 - *56*
E1076 - *50*
E1077 - *47*
E1078 - *40*
E281127 - *44*
E232053 - *36*
E232055 - *38* | *P. jeanneae*
E232056 - *38*
E3116 - *39*
E140117 - *40*
E232050 - *41*
E20399 - *42*
E232059 - *35*
E140116 - *48*
E232057 - *38*
E232052 - *41*

Figure 3. Estimate of relationships of *Psammodromus algirus*, *P. jeanneae* and *P. manuelae* derived from ML and Bayesian analyses using ND4, tRNA^{His}, tRNA^{Ser}, tRNA^{Leu}, CytB, 12S and 16S (see text for details). Numbers above and below nodes represent bootstrap support (>70%) for ML analysis and Posterior Probabilities (>0.95) for Bayesian analyses (not shown in polytomous nodes), respectively. Dashed lines indicate basal branch lengths not proportional to total number of changes. Numbers after specimens refer to the geographic locations shown in fig. 1. Plain text letter and number allocation identifies individuals; boldface number refers to locality on fig. 1.

Figure 2. Phylogenetic relationship of *Psammodromus algirus*, *P. jeanneae* and *P. manuelae* derived from ML and Bayesian analyses using ND4, tRNA^{His}, tRNA^{Ser}, tRNA^{Leu} and CytB (see text for details). Numbers above and below nodes represent bootstrap support (>70%) from ML analysis and Posterior Probabilities (>0.95) for Bayesian analysis (not shown in polytomous nodes), respectively. Dashed lines indicate basal branch lengths not proportional to total number of changes. Plain text letter and number allocation identifies individuals; boldface number refers to locality on fig. 1.

samples from southern localities, and two main northwestern and southwestern clades (figs 2-3). Colonization of the African continent occurred after differentiation between the western and eastern lineages on Iberia, but before the split between southwestern and northwestern clades within the western lineage.

Phylogenetic analyses clearly place the type specimens of *Psammodromus jeanneae* in the southwestern clade of the western lineage and specimens of *P. manuelae* from Manzanares el Real (Madrid province, the type locality) in the northwestern clade (figs 2 and 3). If *P. jeanneae* and *P. manuelae* were, in fact, well-differentiated species, *P. algirus* would be a paraphyletic unit. Within the western clade, however, whose African populations are associated with *P. algirus*, the level of genetic substructuring is much lower than that between western and eastern clades (Carranza et al., 2006; this study) and our molecular data (figs 2 and 3) do not support *P. jeanneae* and *P. manuelae* as well-differentiated species. As a result, we hereby relegate these names to the status of junior synonyms of *P. algirus*.

Our greater geographic coverage, relative to that initially reported by Carranza et al. (2006) and Busack and Lawson (2006), allows us to reject the hypothesis that variation in mtDNA haplotypes is gradual. Currently available data do, however, suggest that highly divergent colour patterns in *Psammodromus algirus* are found in central Spain and in three northeastern populations (Carretero, 2002). If such (or other) phenotipic differences were consistent on a broader geographical scale, the eastern lineage of *P. algirus* might represent a separate species. Combined, this information suggests that ecological or behavioural mechanisms may be currently acting to maintain differentiation, resulting in a relatively narrow contact zone between eastern and western lizards across the centre of the Iberian Peninsula (fig. 1). Additional work, which should include studies of nuclear markers (Godinho et al., 2008) and proper morphometric analyses of these clades, is needed to fully understand genetic and phenotypic variation between and within these lineages, especially at contact zones.

References

Arévalo, E., Davis, S.K., Sites, J.W. (1994): Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. Syst. Biol. **43**: 387-418.

Busack, S.D., Lawson, R. (2006): Historical biogeography, mitochondrial DNA, and allozymes of *Psammodromus algirus* (Lacertidae): a preliminary hypothesis. Amphibia-Reptilia **27**: 181-193.

Busack, S.D., Salvador, A., Lawson, R. (2006): Two new species in the genus *Psammodromus* (Reptilia: Lacertidae) from the Iberian Peninsula. Ann. Carnegie Mus. **75**: 1-10.

Carranza, S., Harris, D.J., Arnold, E.N., Batista, V., González de la Vega, J.P. (2006): Phylogeography of the lacertid lizard, *Psammodromus algirus*, in Iberia and across the Strait of Gibraltar. J. Biogeogr. **33**: 1279-1288.

Carretero, M.A. (2002): Sources of colour pattern variation in Mediterranean *Psammodromus algirus*. Netherlands J. Zool. **52**: 43-60.

Cunningham, C.W. (1997): Is congruence between data partitions a reliable predictor of phylogenetic accuracy? Empirically testing an iterative procedure for choosing among phylogenetic methods. Syst. Biol. **46**: 464-478.

Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C. (1994): Testing significance of incongruence. Cladistics **10**: 315-319.

Felsenstein, J. (1981): Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17**: 368-376.

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

Godinho, R., Crespo, E.G., Ferrand, N. (2008): The limits of mtDNA phylogeography: complex patterns of population history in a highly structured Iberian lizard are only revealed by the use of nuclear markers. Mol. Ecol. **17**: 4670-4683.

Guindon, S., Gascuel, O. (2003): A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**: 696-704.

Hall, T. (2005): Bioedit. Carlsbad, CA, Ibis Therapeutics.

Huelsenbeck, J.P., Ronquist, F.R. (2001): MrBayes: Bayesian inference of the phylogeny. Bioinformatics **17**: 754-755.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X., Wilson, A.C. (1989): Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA **86**: 6196-6200.

Mason-Gamer, R.J., Kellogg, E.A. (1996): Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramyneae). Syst. Biol. **45**: 524-545.

Michkevich, M.F., Farris, J.S. (1981): The implications of congruence in Menidia. Syst. Zool. **30**: 351-370.

Palumbi, S.R. (1996): Nucleic acids. II: the polymerase chain reaction. In: Molecular Systematics, p. 205-247. Hillis, D., Moritz, C., Mable, B.K., Eds, Sunderland, MA, Sinauer Associates.

Posada, D. (2008): jModelTest: Phylogenetic model averaging. Mol. Biol. Evol. **25**: 1253-1256.

Ronquist, F.R., Huelsenbeck, J.P. (2003): MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**: 1572-1574.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmourgin, F., Higgins, D.G. (1997): The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. **24**: 4876-4882.

Wilcox, T.P., Zwickl, D.J., Heath, T.A., Hillis, D.M. (2002): Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. Mol. Phyl. Evol. **25**: 361-371.