Biogeography and host-related factors trump parasite life history: limited congruence among the genetic structures of specific ectoparasitic lice and their rodent hosts

NINA DU TOIT,* BETTINE J. VAN VUUREN,† SONJA MATTHEE‡ and CONRAD A. MATTHEE*

*Evolutionary Genomics Group, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa, †Centre for Invasion Biology, Department of Zoology, University of Johannesburg, PO Box 524, Auckland Park, South Africa, ‡Department of Conservation Ecology and Entomology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

Abstract

Parasites and hosts interact across both micro- and macroevolutionary scales where congruence among their phylogeographic and phylogenetic structures may be observed. Within southern Africa, the four-striped mouse genus, *Rhabdomys*, is parasitized by the ectoparasitic sucking louse, *Polyplax arvicanthis*. Molecular data recently suggested the presence of two cryptic species within *P. arvicanthis* that are sympatrically distributed across the distributions of four putative *Rhabdomys* species. We tested the hypotheses of phylogeographic congruence and cophylogeny among the two parasite lineages and the four host taxa, utilizing mitochondrial and nuclear sequence data. Despite the documented host-specificity of *P. arvicanthis*, limited phylogeographic correspondence and nonsignificant cophylogeny was observed. Instead, the parasite–host evolutionary history is characterized by limited codivergence and several duplication, sorting and host-switching events. Despite the elevated mutational rates found for *P. arvicanthis*, the spatial genetic structure was not more pronounced in the parasite lineages compared with the hosts. These findings may be partly attributed to larger effective population sizes of the parasite lineages, the vagility and social behaviour of *Rhabdomys*, and the lack of host-specificity observed in areas of host sympatry. Further, the patterns of genetic divergence within parasite and host lineages may also be largely attributed to historical biogeographic changes (expansion-contraction cycles). It is thus evident that the association between *P. arvicanthis* and *Rhabdomys* has been shaped by the synergistic effects of parasite traits, host-related factors and biogeography over evolutionary time.

Keywords: comparative phylogeography, cophylogeny, mtDNA, *Polyplax*, *Rhabdomys*, southern Africa

Received 8 March 2013; accepted 12 July 2013

Introduction

Parasites and hosts interact along an evolutionary continuum across micro- and macroevolutionary scales (Brooks & McLennan 1993; Huyse *et al.* 2005). At the

Correspondence: Conrad A. Matthee, Fax: +27218082405;
E-mail: cam@sun.ac.za

macroevolutionary scale, repeated cospeciation (contemporaneous cladogenesis; Brooks 1979; Light & Hafner 2008) within a parasite–host system will lead to congruent phylogenetic relationships (cophylogeny), while erosion of congruence may occur via host-switching, parasite duplication and sorting events (Johnson *et al.* 2003; Page 2003; Clayton *et al.* 2004). Macroevolutionary patterns in turn are influenced by microevolutionary

processes such as selection, drift and dispersal, all of which may result in intraspecific codivergence and congruent phylogeographic structures (Nadler 1995; Clayton & Johnson 2003; Clayton *et al.* 2004; Criscione *et al.* 2005; Nieberding & Morand 2006; Štefka & Hypša 2008; Nieberding *et al.* 2010). While examples of near-perfect congruence between parasite and host genetic structures exist (Hafner *et al.* 1994; Light & Hafner 2007; Demastes *et al.* 2012), these patterns seem to be the exception rather than the rule (Paterson *et al.* 2000; Hoberg & Klassen 2002; Clayton *et al.* 2004; Weckstein 2004; Huyse *et al.* 2005) and parasite–host associations more often represent a complex mosaic of episodic evolutionary events (Hoberg & Brooks 2008).

Sucking lice (Phthiraptera: Anoplura) are obligate and permanent parasites of eutherian mammals (Kim 2006), and these close associations provide interesting models for the study of parasite–host coevolution (Kim 1985; Reed *et al.* 2004; Light & Hafner 2008). Within southern Africa, the widely distributed four-striped mouse genus, *Rhabdomys* (Rodentia: Muridae) is parasitized by the sucking louse *Polyplax arvicanthis* (Ledger 1980; Skinner & Chimimba 2005). Recent phylogenetic evidence based on multiple genes indicated that *P. arvicanthis* consists of two divergent genetic lineages (*P. arvicanthis* 1 and 2) with sympatric distributions

across most of the sampled range (du Toit *et al.* 2013). The apparent lack of gross morphological differences between the two lineages suggests the presence of two cryptic species within *P. arvicanthis* (du Toit *et al.* 2013; Table 1; Fig. 1). A concurrent phylogenetic study on the southern African distribution of *Rhabdomys* (du Toit *et al.* 2012) indicated the existence of at least four host species (*R. pumilio*, *R. intermedius*, *R. bechuanae* and *R. dilectus*; Fig. 1), which has distinct biome distributions with narrow areas of sympatry at biome edges. These findings provide a unique opportunity to investigate parasite–host associations at the interface between micro- and macroevolutionary scales.

The extent of parasite–host congruence is predicted to be largely determined by the intimacy of the interaction as dictated by several parasite life history, ecological and demographic traits (Johnson *et al.* 2003; Huyse *et al.* 2005; Charleston & Perkins 2006; Nieberding & Morand 2006; Nieberding *et al.* 2010) as well as host-related factors (Huyse *et al.* 2005; Barrett *et al.* 2008; Demastes *et al.* 2012). *Polyplax arvicanthis* possesses several traits expected to lead to congruence with the genetic structure of their hosts. They are reported to be specific to the host genus, have a direct life cycle with no free-living phase or intermediate hosts and may even complete several generations on a single host individual (Ledger

Table 1 Localities from which parasite and host specimens were collected (taxonomic designations follow du Toit *et al.* 2012, 2013), number of hosts captured, parasite prevalence and the number of specimens for each parasite lineage per locality

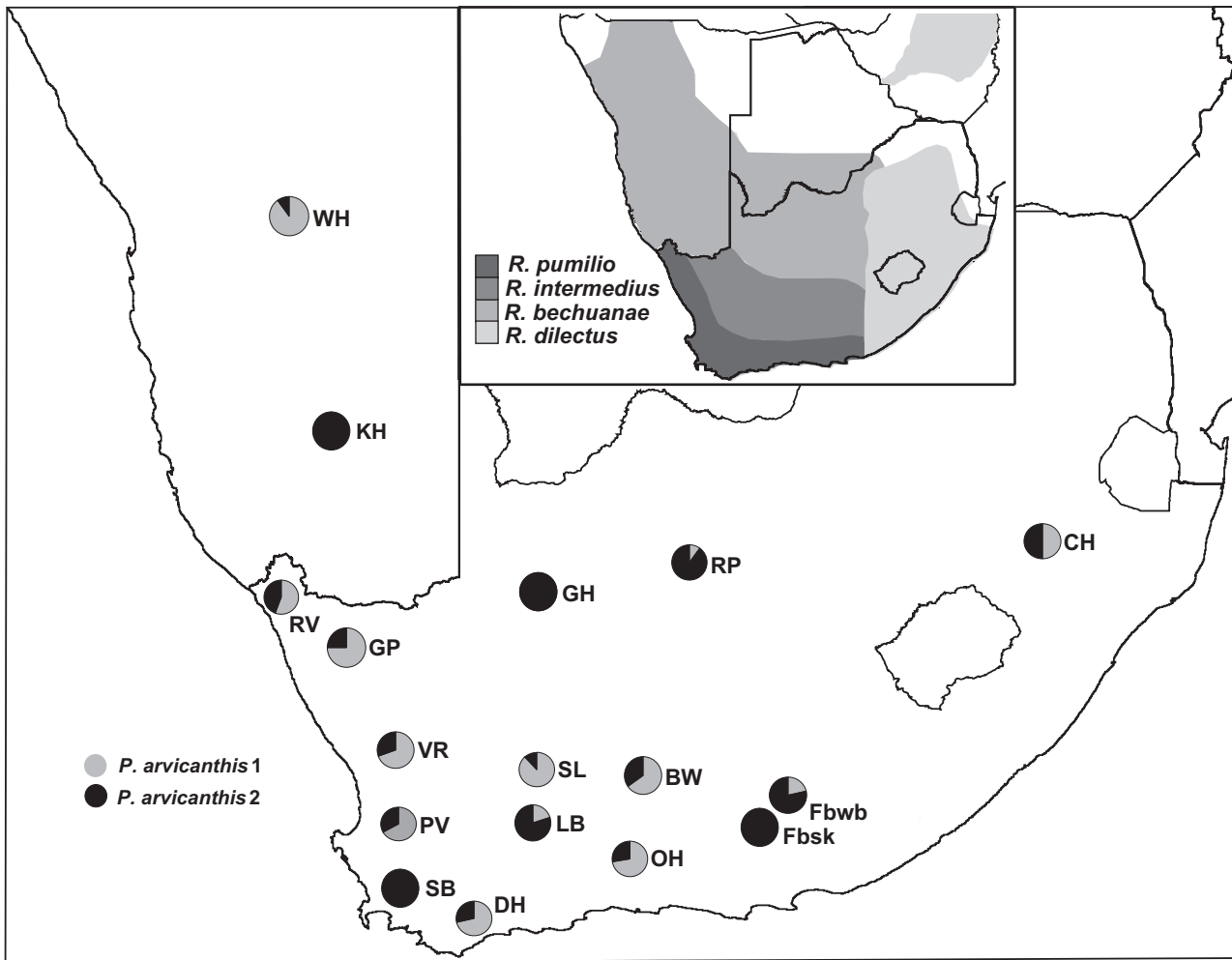| Country/ Province	Locality	Code	Geographic coordinates	Host clade	Hosts(*n*)	Parasite prevalence (%)	*Polyplax arvicanthis* 1 (*n*)	*Polyplax arvicanthis* 2 (*n*)
Namibia	Windhoek	WH	22°31′S, 17°25′E	*Rhabdomys bechuanae*	20	85	18z	2
	Keetmanshoop	KH	26°21′S, 18°29′E	*R. bechuanae*	21	57	Absent	20
South Africa								
Northern Cape	Groblershoop	GH	28°37′S, 21°42′E	*R. bechuanae*	14	57	Absent	18
	Rooipoort	RP	28°39′S, 24°08′E	*R. bechuanae*	15	73	1	27
	Richtersveld	RV	28°12′S, 17°06′E	*R. pumilio*	31	87	15	12
	Springbok	GP	29°42′S, 18°02′E	*R. pumilio*	30	93	21	7
	Sutherland	SL	32°24′S, 20°54′E	*R. intermedius*	13	46	15	2
Western Cape	Vanrhynsdorp	VR	31°44′S, 18°46′E	*R. pumilio*	30	76	16	7
	Porterville	PV	32°59′S, 19°01′E	*R. pumilio*	30	60	12	6
	Stellenbosch	SB	33°55′S, 18°49′E	*R. pumilio*	31	38	Absent	15
	De Hoop	DH	34°29′S, 20°24′E	*R. pumilio*	19	36	10	4
	Oudtshoorn	OH	33°36′S, 22°08′E	*R. pumilio*	31	93	21	8
	Laingsburg	LB	33°10′S, 20°55′E	*R. intermedius*	23	30	2	8
	Beaufort West	BW	32°13′S, 22°48′E	*R. intermedius*	33	45	13	7
Eastern Cape	Fort Beaufort_Smutskraal	FB_sk	32°50′S, 26°27′E	Contact zone	14	57	Absent	18
	Fort Beaufort_Winterberg	FB_wb	32°47′S, 26°37′E	Contact zone	15	73	3	11
Kwazulu-Natal	Chelmsford	CH	28°00′S, 29°54′E	*R. dilectus chakae*	7	85	4	4
Total					377		151	176

Fig. 1 Localities from which parasite and host specimens were collected (codes as in Table 1), indicating the frequency of the two *Polyplax arvicanthis* lineages at each locality and the distributions of the four putative *Rhabdomys* species (insert).

1980; Durden & Musser 1994; Kim 2006). The dispersal of *P. arvicanthis* is thus completely dependent on that of *Rhabdomys*, and vertical transmission through successive generations likely occurs (Blouin *et al.* 1995; Jerome & Ford 2002; Johnson *et al.* 2003; Johnson & Clayton 2004; Whiteman & Parker 2005; Wirth *et al.* 2005). This lifestyle is expected to impose impediments to parasite gene flow and limit opportunities for host switching to other potential host species (Blouin *et al.* 1995; Nadler 1995; Johnson *et al.* 2002; McCoy *et al.* 2003). The probability that a parasite will track the migration and differentiation patterns of its host is further strengthened by high prevalence and intensity (Clayton *et al.* 2003; Nieberding & Morand 2006; Nieberding *et al.* 2010), both of which have been found for *P. arvicanthis* on *Rhabdomys* (prevalence = 60%, relative mean infestation intensity = 12.43%; Matthee *et al.* 2007).

Parasites, in general, are also expected to show stronger signatures of genetic structure when compared with

their hosts as a result of the combined effects of limited dispersal ability, smaller effective population sizes than free-living organisms of equal census size (greater effect of drift, particularly in isolated populations) and elevated mutation rates (Nieberding & Morand 2006; Nieberding *et al.* 2010). In the present study, it is expected that *P. arvicanthis* will have small effective population sizes as the species is recorded to have a female-biased sex ratio (Matthee *et al.* 2007), which is also generally the case for parasites without intermediate hosts and free-living phases (Rannala & Michalakis 2003; Criscione & Blouin 2005). A more pronounced phylogeographic signature is also predicted for *P. arvicanthis* based on the elevated mtDNA mutation rates generally reported for chewing and sucking lice (Hafner *et al.* 1994; Page & Hafner 1996; Page *et al.* 1998, 2004; Paterson & Banks 2001; Light & Hafner 2007, 2008).

Biogeographic changes can also have significant impact on host–parasite interactions over evolutionary

time (Hoberg & Brooks 2008). A shared biogeographic history among closely interacting parasites and their hosts may lead to congruent genetic relationships through similar responses to vicariant events (Hoberg & Klassen 2002; Page 2003; Hoberg & Brooks 2008). Biogeography may also determine the genetic structure of parasites irrespective of host associations, particularly in systems involving multihost parasites, parasites with free-living phases and parasites with intermediate hosts (Clayton 1990; Weckstein 2004; Hoberg & Brooks 2008; Nieberding *et al.* 2008). In the current study, the host genus *Rhabdomys* has a complex biogeographic history. Given the specificity and apparent ecological adaptation of the various host taxa to their respective vegetational biomes (du Toit *et al.* 2012), cladogenesis within the genus probably occurred in response to climatic and landscape changes that lead to the establishment of these biomes around the Mio-Pliocene boundary (Scott *et al.* 1997; Tyson & Partridge 2000; Mucina & Rutherford 2006). Subsequently, several cycles of range expansions and contractions of the various host taxa probably occurred in response to shifting biome distributions during climatic oscillations in the region (Ellery *et al.* 1991; Partridge 1997; Zachos 2001; Scott & Nyakale 2002; Dupont 2006; Montgelard & Matthee 2012; du Toit *et al.* 2012).

In the present study, we compare the phylogeographic relationships within the two *P. arvicanthis* lineages and the four *Rhabdomys* taxa to test for spatial congruence of genealogical relationships among parasites and hosts, utilizing both mitochondrial and nuclear sequence data. Given the life history of *Polyplax,* we hypothesized that the parasites would have more pronounced spatial genetic structuring than the hosts and that phylogeographic and phylogenetic congruence would be observed among the parasite and host structures. Potential cophylogeny among the parasite and host genetic groups was evaluated using distance- and topology-based approaches. Contact zones, where host lineages occur sympatrically, were used to evaluate host-specificity by determining whether host switching occurs among parasites of ecological and genetically divergent host lineages.

Materials and methods

Taxon and gene sampling

Polyplax (*n* = 327) and *Rhabdomys* (*n* = 377) specimens were collected from 17 localities across the southern African distribution (Table 1; Fig. 1). Live traps (Sherman type) baited with a mixture of peanut butter and oats were used to capture host individuals. Mice were euthanized with 0.2–0.4 mL sodium pentobarbitone

(200 mg/kg) and placed in individual plastic bags to prevent the loss of ectoparasites postmortem. Host specimens were frozen in the field at −20 °C and subsequently thawed in the laboratory, where all lice were removed with forceps under a stereoscopic microscope. *Polyplax arvicanthis* specimens included in the genetic analyses were selected from as many different host individuals as possible per sampled locality (determined by parasite prevalence and abundance; Table 1) and preserved in 100% ethanol (correct identification of *P. arvicanthis* was morphologically confirmed by L.A. Durden; Department of Biology, Georgia Southern University, USA).

Mitochondrial Cytochrome Oxidase subunit I (COI) sequence data were generated for all host and parasite specimens (Tables S1 and S2, Supporting information). These data were augmented by sequencing nuclear introns. The eukaryotic translation elongation factor 1 alpha 1 (Eef1a1) and carbamoyl-phosphate synthetase 2, aspartate transcarbamylase and dihydroorotase (CAD) regions were sequenced for *Rhabdomys* (Table S1, Supporting information) and *Polyplax* (Table S2, Supporting information), respectively. Attempts were made to sequence as many mtDNA haplotypes as possible, and if haplotypes occurred at multiple sampled localities, efforts were also made to sequence at least one individual from each locality.

Molecular techniques and data analysis

Total genomic DNA was extracted from host tissue and whole individual louse specimens, and amplification and sequencing of the various gene fragments were performed following standard protocols and published primers (Table S3, Supporting information; du Toit *et al.* 2012, 2013). Sequences were edited in BIOEDIT Sequence Alignment editor 7.0.5 (Hall 2005), and the alignments were trimmed to avoid missing data (Table 2). When compared with other *Polyplax* species, a 6-bp or 3-bp insert was present within the *P. arvicanthis* 1 and 2 COI sequences, respectively. Translation into protein sequences using EMBOSStranseq (www.ebi.ac.uk/Tools/st/emboss_transeq) indicated that both variants represent functional copies. The insertions were coded as presence/absence data in the phylogenetic analyses and included in the population-level analyses, as the two parasite lineages were analysed separately for the latter. For the nuclear regions, heterozygous positions were resolved in PHASE, version 2.1.1 (Stephens *et al.* 2001; Stephens & Scheet 2005). The algorithm ran for 100 000 generations with a thinning interval of 1 and 10 000 burn-in generations. Phases with a 0.9 probability or higher were considered resolved (Stephens *et al.* 2001). Subsequent nuclear analyses were

Table 2 The total number of sequences and alleles after resolving heterozygous positions (nDNA), haplotypes retrieved, total sites, polymorphic sites, nucleotide diversity (π), haplotype diversity (h) and estimated alpha shape parameter of the gamma rate variation distribution for the mitochondrial and nuclear data sets. Theta estimates of mitochondrial effective population size are also indicated. Standard deviations from the sliding window analyses of *Rhabdomys* mitochondrial π and theta are indicated in brackets

	Sequences/ Alleles	Total haplotypes	Singleton haplotypes	bp	P	π	h	α	θ
mtDNA									
Rhabdomys (COI)	377	97	56	900	211	0.06 ± 0.03 (0.003)	0.96 ± 0.003	0.14	0.036 ± 0.002 (0.002)
Polyplax arvicanthis 1 (COI)	151	42	29	270*	107	0.09 ± 0.04	0.85 ± 0.02	0.34	0.069 ± 0.006
Polyplax arvicanthis 2 (COI)	176	67	45	270*	114	0.11 ± 0.05	0.92 ± 0.01	0.34	0.072 ± 0.006
nDNA									
Rhabdomys (Eef1a1)	83/166	92	63	230	30	0.01 ± 0.002	0.98 ± 0.004	0.90	—
Polyplax arvicanthis 1 (CAD)	46/92	36	22	348	40	0.01 ± 0.005	0.93 ± 0.01	0.17	—
Polyplax arvicanthis 2 (CAD)	66/132	60	44	348	36	0.01 ± 0.006	0.97 ± 0.006	0.17	—

*Excluding the two 3-bp insertions.

COI, cytochrome oxidase subunit I; Eef1a1, eukaryotic translation elongation factor 1 alpha 1; CAD, carbamoyl-phosphate synthetase 2, aspartate transcarbamylase and dihydroorotase.

performed on all alleles (Table 2). Haplotypes were identified using COLLAPSE 1.21 (Posada 2004). Parasite and host mitochondrial effective population sizes (estimated with theta per site from the number of polymorphic sites) and standard molecular diversity indices for the mitochondrial and nuclear data sets were calculated in DNASP, version 5 (Librado & Rozas 2009). To investigate the potential impact of different mitochondrial fragment lengths (hosts = 900 bp, parasites = 276 bp) on these estimates, sliding window analyses (window length = 270, step size = 100) were performed for *Rhabdomys*.

Genetic relationships

Statistical parsimony networks were constructed in TCS 1.21 (Clement *et al.* 2000) to depict relationships among both mitochondrial and nuclear haplotypes of *P. arvicanthis* 1, 2 and *Rhabdomys*, separately. Distinct haplogroups, which could not be connected within the 95% confidence interval, were retrieved from all three mtDNA data sets, and Bayesian and maximum-likelihood trees were therefore constructed from the parasite and host haplotype data sets to explore deeper evolutionary relationships. Bayesian COI topologies

were constructed in MRBAYES, version 3.2 (Ronquist *et al.* 2012), and best-fit models of sequence evolution, as determined in JMODELTEST, version 0.1.1 (Guindon & Gascuel 2003; Posada 2008), using the Akaike Information Criterion corrected (Akaike 1973; Burnham & Anderson 2004), were specified in each analysis (Table S4, Supporting information). All analyses were partitioned by codon, with individual models assigned to each partition, and parameters were unlinked across partitions. Each analysis consisted of two parallel Markov chain Monte Carlo (MCMC) simulations of five chains each that were run for 5 million generations with a sampling frequency of 100 generations. Parameter convergence and ESS values were monitored in TRACER, version 1.5 (Rambaut & Drummond 2007), and 25% of the total number of generations were discarded as burn-in. Maximum-likelihood topologies were constructed in RAXMLGUI, version 0.95 (Silvestro & Michalak 2012), implementing the thorough bootstrap option (10 runs, 1000 replicates). The recommended GTR+G model (Stamatakis 2006) was specified, and the data sets were partitioned by codon.

Bayesian analysis of population structure (BAPS) was conducted on the complete parasite and host mtDNA data sets (all individuals) in the software BAPS, ver-

sion 5.3 (Corander *et al.* 2008), using a nonspatial mixture model for linked loci (Corander & Tang 2007). The codon linkage model and a vector of maximum K values (ranging from 1 to the maximum number of populations, each replicated by 5) were specified. For both mitochondrial and nuclear complete data sets, three-level hierarchical analyses of molecular variance (AMOVA; Excoffier *et al.* 1992) were conducted and Φ statistics were calculated in ARLEQUIN, version 3.5.1.2 (Excoffier & Lischer 2010), to test for significant differentiation (i) among clusters identified by the mtDNA data; (ii) among sampled localities within these clusters; and (iii) within sampled localities. Gamma corrections estimated for each data set in JMODELTEST were applied (Table 2) and significance evaluated with 1000 permutations. Pairwise Φ_{st} values were calculated among localities, and sequential Bonferroni corrections were applied within each data set to account for multiple comparisons (Rice 1989).

Spatial mtDNA structure

Mantel tests (Mantel 1967) for matrix correspondence of genetic relatedness and geographic distance among individuals (Smouse *et al.* 1986; Smouse & Long 1992) were performed on the complete mtDNA data sets in GENALEX, version 6.5 (Peakall & Smouse 2006, 2012), to test for isolation by distance (IBD). Statistical significance was assessed with 10 000 random permutations. Partial mantel tests (Legendre & Legendre 1998) were conducted in IBDWS, version 3.23 (Bohonak 2002; Jensen *et al.* 2005), using 10 000 randomizations to test whether significant differentiation was present among genetic clusters when correcting for IBD (Drummond & Hamilton 2007; Meirmans 2012).

In the presence of hierarchical structuring, IBD may potentially act only at certain distances (Drummond & Hamilton 2007). To further investigate the potential extent of IBD and spatial structure, spatial autocorrelation analyses were performed in GENALEX, version 6.5 (Peakall & Smouse 2006, 2012), at the individual level for *P. arvicanthis* 1, 2 and *Rhabdomys*. The genetic autocorrelation coefficient (*r*) was estimated for increasing distance size classes. Statistical significance was evaluated through the 95% confidence interval, as defined from the distribution of 1000 randomly permuted *r* values generated under the null hypothesis of no autocorrelation (Smouse & Peakall 1999). Significant autocorrelation is inferred if the observed *r* value falls outside the confidence interval. Confidence intervals around the observed *r* values were calculated with 1000 bootstrap replicates (Peakall *et al.* 2003).

Cophylogeny

Cophylogeny between *Rhabdomys* and *Polyplax* was investigated by applying a distance-based (global fit) approach to the mtDNA haplotype data sets (trees generated from the Bayesian analyses; see above) and a pruned data set consisting of sixteen mtDNA haplotypes representative of all the genetic clusters (*P. arvicanthis* 1 = 5 clusters; *P. arvicanthis* 2 = 7 clusters; *Rhabdomys* = 4clusters; contact zones excluded). Bayesian phylogenetic trees for the pruned haplotype data sets were constructed using the methods outlined above (also see Table S4, Supporting information), and the topological relationships were congruent with those obtained from the haplotype data sets. The distance-based cophylogenetic analyses were performed in AXPARAFIT (Stamatakis *et al.* 2007), which is an optimized version of PARAFIT (Legendre *et al.* 2002) that is implemented in COPYCAT, version 2.02 (Meier-Kolthoff *et al.* 2007). The hypothesis of cophylogeny between parasite and host is tested through the comparison of parasite and host patristic distances (calculated from the branch lengths of phylogenetic trees), incorporating the parasite–host associations. Significance was assessed with 10 000 permutations, generated under the null hypothesis of independent parasite and host evolution. To exclude the possibility that a lack of cophylogeny signal may arise from study artefacts (e.g. insufficiently resolved parasite topologies), the Shimodaira–Hasegawa test (SH test; Shimodaira & Hasegawa 1999) was performed in PAUP*, version4.0b10 (Swofford 2000), for both haplotype and pruned data sets. This test determines whether the host and parasite topologies are significantly different by comparing the likelihood scores of the parasite topology and a constrained host topology, given the parasite data (see Light & Hafner 2008). A significant difference between the parasite and host topologies would imply that the underlying cause of these topological differences are biologically meaningful and not a product of study artefacts.

Based on the outcome of the former analyses, event-based cophylogenetic analyses were conducted on the pruned cluster data sets. Event-based methods attempt to identify the most probable coevolutionary history of the taxa concerned (Merkle *et al.* 2010; Keller-Schmidt *et al.* 2011). A cost is assigned to each potential evolutionary event (i.e. cospeciation, duplication, host switch, sorting event) and the parasite phylogeny mapped onto that of the host using these different events, while attempting to find the solution with the minimum total cost (most parsimonious). Several software packages implementing this approach are available (see De Vienne *et al.* 2013), most of which rely on the a priori assignment of a cost scheme (Merkle *et al.* 2010). It has

been shown that the outcome of event-based analyses are heavily dependent on the cost scheme employed (Merkle *et al.* 2010), and choosing a biologically meaningful cost scheme a priori is often difficult (De Vienne *et al.* 2013). To evaluate the impact of varying cost schemes on cophylogeny reconstruction, reconciliation of *Polyplax* and *Rhabdomys* topologies was performed in JANE, version 4 (Conow *et al.* 2010), using several different cost schemes (including the defaults from other software programs; see results). The cost of failure to diverge events, which is exclusively implemented in Jane, was arbitrarily set to 1. JANE, version 4 (Conow *et al.* 2010), was chosen as it is currently the only software that allows multihost parasites (Althoff *et al.* 2012; Mendlová *et al.* 2012). The Vertex-based cost model method was implemented, the genetic algorithm was set to 500 generations with a population size of 300 and the statistical significance of reconstructions was evaluated with 1000 random tip mating permutations.

CORE-PA (Merkle *et al.* 2010) and TREEMAP, version 3b (Charleston 2011), follow alternative event-based reconciliation approaches that do not require a priori cost assignment (Keller-Schmidt *et al.* 2011). The presence of multihost parasites in our data, however, precluded the use of TREEMAP. Thus, event-based reconstruction for *Polyplax* and *Rhabdomys* was also performed in CORE-PA, version 0.3a (Merkle *et al.* 2010). CORE-PA implements a parameter-adaptive reconstruction approach where several random cost schemes are applied to the data, and the most parsimonious reconstructions are retained. These reconstructions are then rated according to their quality value, which is a measure of how well the reconstruction fits the particular cost model used (see Dilcher *et al.* 2011; Rosenblueth *et al.* 2013). We evaluated 100 000 random cost sets and statistical significance was assessed with 1000 randomized parasite–host associations.

Evolutionary rates

*BEAST as implemented in BEAST, version 1.7.5 (Drummond *et al.* 2012), uses the multispecies coalescent to jointly estimate multiple gene trees and a shared species tree from multilocus data using multiple individuals per species (Heled & Drummond 2010). This approach has been used previously to estimate the relative rates of evolution among hosts and parasites (Štefka *et al.* 2011) and was implemented here to estimate the COI evolutionary rates of the two *P. arvicanthis* lineages relative to that of *Rhabdomys*. The two *P. arvicanthis* lineages were analysed separately and the data sets consisted of all parasite individuals and the complementary host individuals. In each analysis, the rate for *Rhabdomys* was set to 1.0, allowing relative estimation of the para-

site rates using an uncorrelated lognormal relaxed clock (Drummond *et al.* 2006). Best-fit models of sequence evolution for whole (unpartitioned) host and parasite data sets were estimated in JMODELTEST, version 0.1.1 (Guindon & Gascuel 2003; Posada 2008). Subsequently, the GTR+G model was specified for *Rhabdomys* and the GTR+I model for *P. arvicanthis* 1 and 2. All data sets were unpartitioned to allow estimation of overall evolutionary rates, a birth–death process was implemented as species tree prior (Stadler 2010), and the piecewise linear and constant root population size model was used. The analyses ran for 500 million generations, sampling every 50 000 generations. The output was evaluated in TRACER, version 1.5 (Rambaut & Drummond 2007), to ensure that all ESS values were >200, ignoring the first 10% of generations as burn-in.

Results

Parasite prevalence and distribution

Relatively high prevalence levels were recorded for *P. arvicanthis*, with most localities exceeding 50% (Table 1). The two parasite lineages, however, were found to occur at different frequencies at the various localities, and <15% of the *Rhabdomys* individuals from which multiple louse specimens were sampled harboured both louse lineages (Table 1; Fig. 1).

Genetic relationships among sampling sites

Mitochondrial data were generated for 377 *Rhabdomys*, 151 *P. arvicanthis* 1 and 176 *P. arvicanthis* 2 specimens from 17 localities throughout southern Africa [see Tables S1 and S2 (Supporting information) for GenBank accessions]. A total of 83 *Rhabdomys*, 46 *P. arvicanthis* 1 and 66 *P. arvicanthis* 2 nuclear sequences were generated in our attempts to sequence all the mtDNA haplotypes identified [Table 2; see Tables S1 and S2 (Supporting information) for GenBank accessions]. *Polyplax arvicanthis* 1 had relatively fewer unique haplotypes and also a lower haplotype diversity when compared with *P. arvicanthis* 2 for both mtDNA and nDNA data sets (Table 2). As expected for species characterized by higher mutation rates (Nieberding & Morand 2006; Nieberding *et al.* 2010), both parasite lineages had higher mtDNA nucleotide diversities and proportions of polymorphic sites than the host. Estimated theta values generated for the mtDNA data indicated that both parasite taxa had larger effective population sizes than the host (Table 2). Sliding window analyses indicated that theta and nucleotide diversity estimates did not vary significantly along the length of the *Rhabdomys* COI fragment (Table 2). Thus, the different parasite and

host COI fragment length sizes did not lead to bias in these estimates.

A remarkably high level of genetic divergence was detected within the parasite and host mtDNA data sets as several geographically structured haplogroups were retrieved from the statistical parsimony analyses (Fig. 2). These groups could not be connected within the 95% confidence interval, which corresponds to more than 7 and 13 mutational steps for the parasites and hosts, respectively. Lowering the confidence interval to 90% did not result in additional connections among haplogroups (with the exception of *P. arvicanthis* 2 cluster 1 and 2 that collapsed to form a single clade; Fig. 2), lending further support for extensive genetic divergence among these groups. The distinct genetic clusters retrieved from BAPS (Table S5, Supporting information; Fig. 2) were largely congruent with the TCS

haplogroups found for both parasite and host taxa. These haplogroups were also retrieved from the phylogenetic analyses (Bayesian and maximum likelihood; Fig. 2); with most haplogroups forming statistically supported monophyletic clades. The topologies also indicate the relationships among the haplogroups, which were relatively well resolved overall (Fig. 2). In most instances, haplogroups in close geographic proximity cluster as sister assemblages in the topologies (Fig. 2).

At the mtDNA level, *Rhabdomys* consists of five genetic groups (Table S5, Supporting information; Fig. 2), corresponding to the four putative species (du Toit *et al.* 2012; Fig. 1), with further subdivision of *R. pumilio* into subclades 'A' and 'B'. Subclade 'B' consists exclusively of individuals sampled in the contact zones (FB_sk and FB_wb; Fig. 1; du Toit *et al.* 2012) where mtDNA haplotypes belonging to *R. pumilio*,
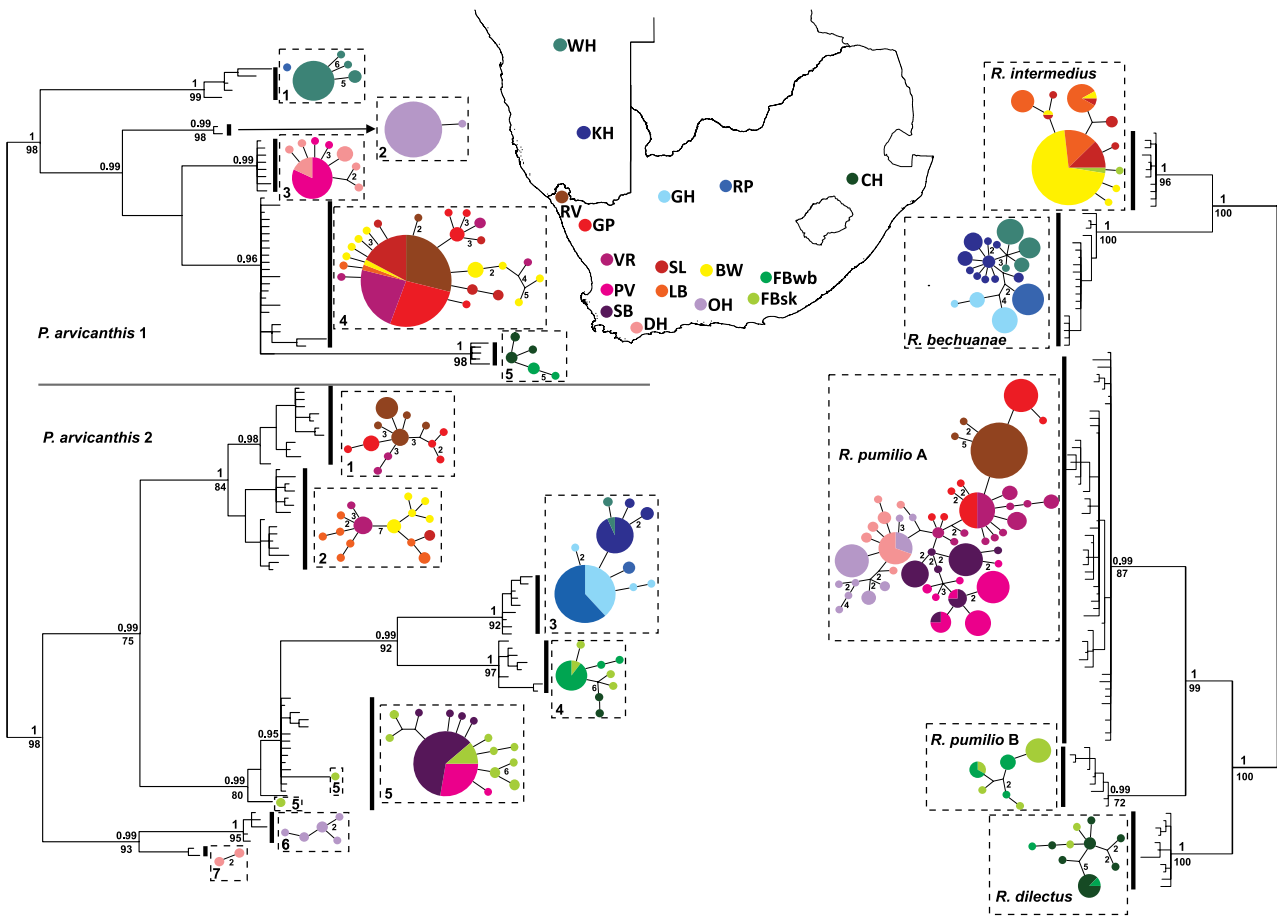

Fig. 2 Composite of the Cytochrome Oxidase subunit I (COI) statistical parsimony haplotype networks for both parasite taxa (left) and the host (right) with accompanying Bayesian topologies indicating the relationships among haplogroups. Bayesian posterior probabilities (>0.95) and maximum-likelihood bootstraps (>70) are indicated above and below nodes, respectively. Each circle constitutes a particular haplotype with size indicating the relative number of individuals per haplotype. Colours indicate the frequency of each haplotype within the various sampled localities (insert) and each connection constitutes a single mutational step with numbers under lines indicating number of steps if more than one. Dashed-line boxes indicate the genetic clusters retrieved from Bayesian analysis of population structure (BAPS), which for *Rhabdomys* coincides with the previously described species (du Toit *et al.* 2012).

R. intermedius and *R. dilectus* were found sympatrically. *Polyplax arvicanthis* 1 and 2 contains 5 and 7 distinct genetic groups, respectively (Table S5, Supporting information; Fig. 2). No shared haplotypes were detected among the haplogroups of *P. arvicanthis* 1, while some mixture was present at Vanrhysdorp (VR) between haplogroups 1 and 2 of *P. arvicanthis* 2 (Table S5, Supporting information). Some spatial congruence is evident among the parasite and host genetic groups, particularly within the arid northwestern region (*P. arvicanthis* 1 cluster 1 and *P. arvicanthis* 2 cluster 3 with *R. bechuanae*; Fig. 2) and the mesic eastern region (*P. arvicanthis* 1 cluster 5 and *P. arvicanthis* 2 cluster 4 with *R. dilectus*; Fig. 2) of southern Africa. These parasite and host clusters appear to be genetically differentiated from all other clusters as indicated by the absence of shared haplotypes, although they do not necessarily form basal monophyletic groups to the exclusion of other clusters (Fig. 2). Contrary to our initial predictions, the parasites collected from *R. pumilio*, *R. dilectus* and *R. intermedius* within the contact zones (FB_sk and FB_wb) are not genetically differentiated (Fig. 3).

All haplotypes in the nuclear DNA statistical parsimony networks could be connected within the 95% confidence interval (Fig. 4 and Fig. S1, Supporting information). Within *Rhabdomys*, the highest frequency of shared haplotypes occurs among *R. pumilio* and *R. intermedius*, which is also reflected by the parasites collected from these two hosts (Fig. 4). The close relationship between *P. arvicanthis* 2 clusters 3, 4 and 5 at the mtDNA level is reflected by the high incidence of shared haplotypes among these clusters in the nuclear CAD network (Fig 2 and Fig. S1, Supporting information).

Using the mtDNA haplogroups as prior, the hierarchical analyses of molecular variance indicated that all parasite and host genetic groups are significantly differentiated at the mtDNA level (Table 3). All the respective mitochondrial genetic groups are also significantly differentiated at the nuclear level (stronger in the parasites), although the values suggest lower levels of differentiation for these data. Pairwise Φst values among geographic sampling sites of both *P. arvicanthis* lineages and *Rhabdomys* indicated significant differentiation among nearly all sites at the mtDNA level (Tables S6 and S7, Supporting information) and for several pairwise comparisons at the nuclear level (Tables S6 and S8, Supporting information).

Spatial mtDNA structure

The Mantel tests indicated significant IBD within *Rhabdomys* and *P. arvicanthis* 2 (Table 4). It has been shown that the presence of IBD may lead to the overestimation of distinct genetic clusters by amongst others, BAPS (Frantz *et al.* 2009; Safner *et al.* 2011; Meirmans 2012). In our case, this seems unlikely as the clusters identified by BAPS are congruent with the statistical parsimony haplogroups, which is also supported by the Bayesian topologies. Furthermore, the results of the partial Mantel tests indicated that the relationship between cluster membership and genetic distance remains significant when controlling for geographic distance (Table 4; also see Meirmans 2012).

To further evaluate the effects of geographic distance on the genetic signatures obtained in the present study, we evaluated and found significant positive spatial autocorrelation among sampling sites for both *P. arvicanthis* lineages and *Rhabdomys*, with *r* becoming nonsignificant at 1200 km (Fig. 5). This correlation was strongest at the shortest distances and declined with increasing distance, indicating that IBD is strongest among neighbouring populations and that the positive relationship between genetic and geographic distance deteriorates with increasing distance among sites. When compared among the groups, positive spatial autocorrelation was much stronger in *P. arvicanthis* 2 relative to *P. arvicanthis* 1 and *Rhabdomys*, as indicated by higher *r* values at the shortest distances with the magnitude of difference declining with increasing distance class size. *Polyplax arvicanthis* 1 had the lowest *r* values, which corresponds to the nonsignificant Mantel test for overall IBD (Table 4).

Cophylogeny

Although there is some phylogeographic congruence among parasite and host genetic structures when the haplogroups are compared, the Bayesian topologies indicate a fair amount of incongruence at the larger scale (Fig. 2). Furthermore, within the host contact zones (FB_sk and FB_wb; Fig. 2), parasites collected from *R. pumilio*, *R. dilectus* and *R. intermedius* are not genetically differentiated (Fig. 3). Thus, as expected, the AXPARAFIT global cospeciation tests indicated nonsignificant association between parasite and host distances for the mtDNA haplotype data set ($P = 0.06$) and pruned data set ($P = 0.88$). The near significant result of the haplotype analysis (considering the significance level of $P < 0.02$ suggested by AXPARAFIT) is the result of only three out of 103 significant individual links. SH tests on both the haplotypes and pruned data sets revealed significant differences ($P < 0.05$) among parasite and host topologies, thus indicating that the absence of significant cophylogeny is biologically meaningful and not due to study artefacts.

The JANE reconstructed scenarios varied according to the cost scheme employed with a maximum of two
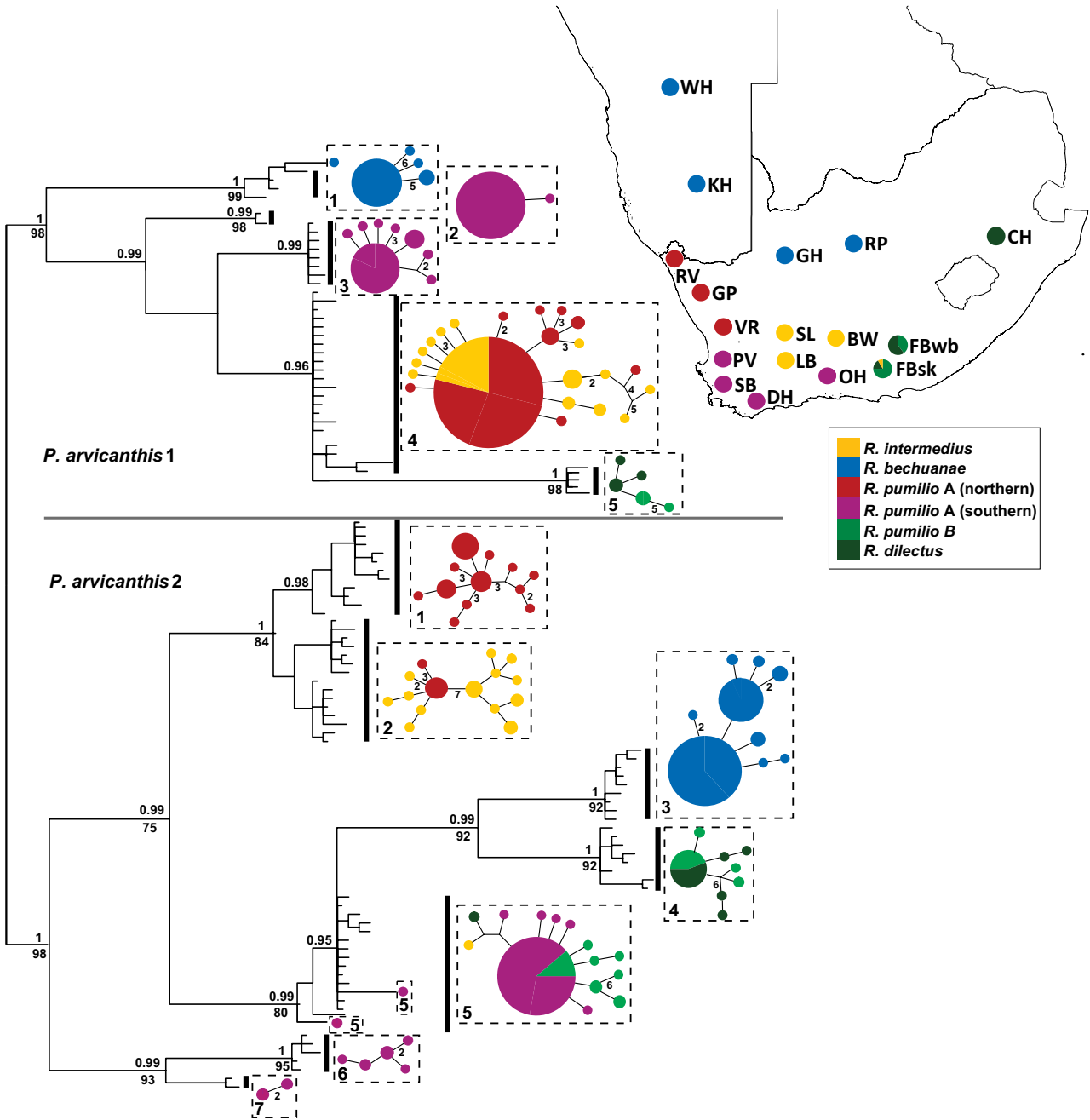
Fig. 3 *Polyplax arvicanthis* composite mtDNA haplotype network and Bayesian topology depicting spatial congruence with the host genetic groups. Bayesian posterior probabilities (>0.95) and maximum-likelihood bootstraps (>70) are indicated above and below nodes, respectively.

cospeciations retrieved (Table 5). All scenarios indicated nonsignificant cophylogeny, as the average total costs resulting from random solutions were lower than that of the observed reconstructions (Table 5). The number of inferred codivergence and failure to diverge events were stable across different cost schemes, while most cost schemes indicated either predominantly host switching or a combination of parasite duplications and

losses. The CORE-PA analysis yielded 15 parsimonious reconstructions, each optimal to the particular event cost model used. The preferred reconstruction had a quality value of 0.05 and contained three codivergence, seven duplication, two host switching and 10 sorting events (Fig. 6; optimal costs = codivergence: 0.28, duplication: 0.14, host switch: 0.45 and sorting: 0.10). The reconstruction with the second best-cost model fit
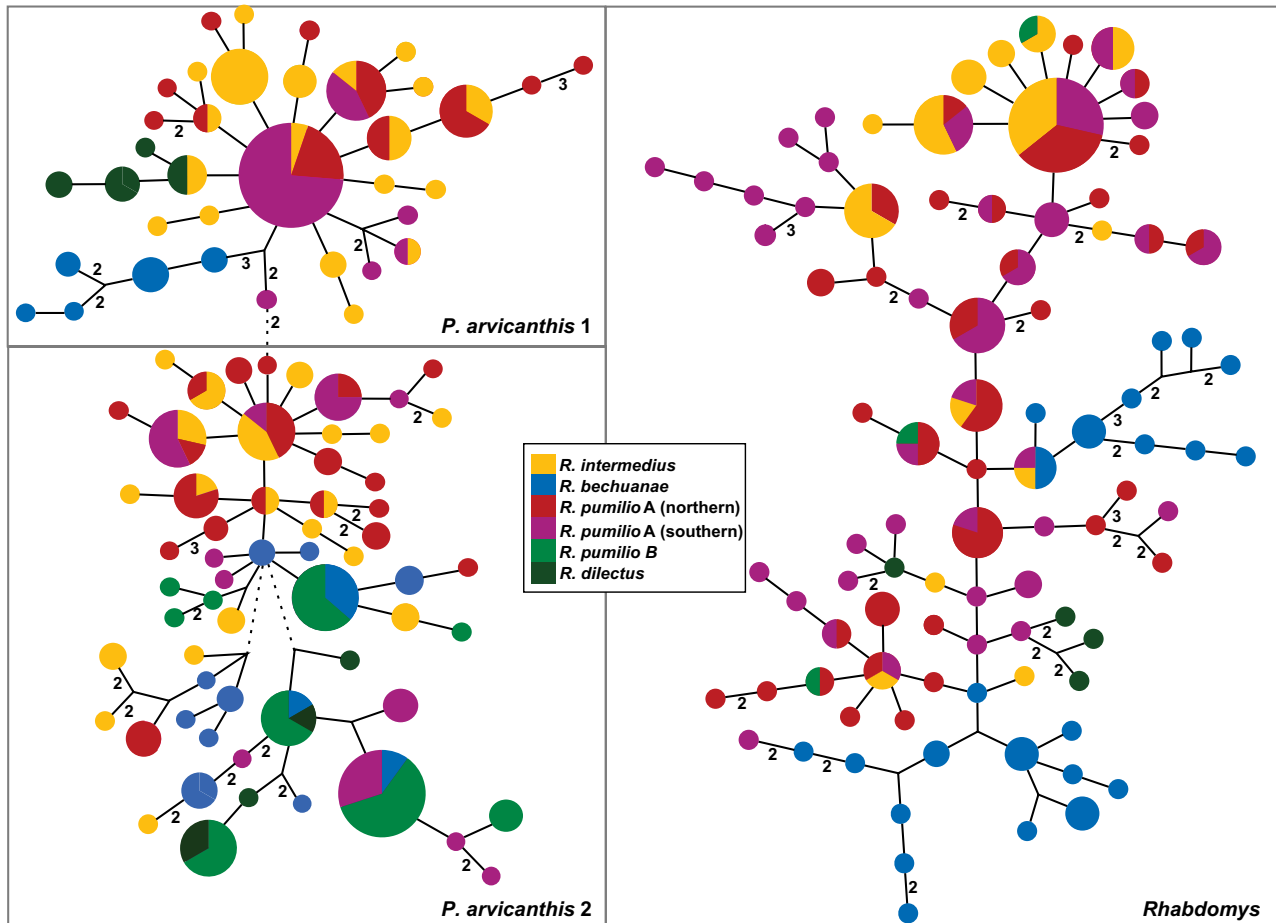
Fig. 4 Nuclear statistical parsimony networks for the parasite carbamoyl-phosphate synthetase 2, aspartate transcarbamylase and dihydroorotase (CAD) (left) and host eukaryotic translation elongation factor 1 alpha 1 (Eef1a1) (right) genes. Haplotypes are coloured according to host genetic groups (insert). Dotted lines indicate single mutational steps and are used for ease of representation. Clusters that have congruent population membership within the parasites and the host are indicated in the same colours.

(quality value = 0.08) contained three codivergence, seven duplication, three host switching and six sorting events (Fig. 6; optimal costs = codivergence: 0.26, duplication: 0.15, host switch: 0.43 and sorting: 0.13). All other reconstructions had considerably lower-quality values (0.19–0.73). Cophylogeny between parasite and host trees was nonsignificant with 77.3% of the random reconstructions having a total cost lower than the preferred reconstruction.

Evolutionary rates

The estimated COI evolutionary rates of *P. arvicanthis* 1 and 2 were 2.19 [95% highest posterior density interval (HPD): 1.20–3.26] and 2.58 (HPD: 1.56–3.71) times faster than those of *Rhabdomys*, respectively. Standard deviations of the uncorrelated lognormal relaxed clock were less than 1 for all parasite and host data sets (*P. arvicanthis* 1: 0.97, *P. arvicanthis* 2: 0.30, *Rhabdomys*: 0.35/0.68),

indicating moderate deviation from clock-like behaviour (Drummond *et al.* 2006).

Discussion

Polyplax arvicanthis has high specificity to *Rhabdomys*, a direct life cycle (Ledger 1980; Kim 2006) and high prevalence across the distribution (Table 1; Matthee *et al.* 2007, 2010). These factors are predicted to act in concert to enhance vertical transmission and limit host-switching opportunities (Blouin *et al.* 1995; Nadler 1995; Jerome & Ford 2002; Johnson *et al.* 2003; Johnson & Clayton 2004; Whiteman & Parker 2005; Wirth *et al.* 2005) that will tend to increase the probability of parasites tracking host movements and thus promote parasite–host congruence (Clayton *et al.* 2003; Nieberding & Morand 2006; Nieberding *et al.* 2010). In the present study, only limited congruence is present among the phylogeographic structures of the *P. arvicanthis* lineages and the

four *Rhabdomys* species (Figs 2 and 3). This finding is thus contrary to the expectation that highly host-dependant parasites should show high levels of congruence with the genetic structures of their hosts (Blouin *et al.* 1995; Whiteman & Parker 2005; Nieberding & Morand 2006; Nieberding *et al.* 2010). Instead, the outcome of our investigation emphasizes the complexity of parasite–host associations and clearly suggests that factors such as biogeography and host life history can be equally, or even more important, in shaping parasite–host evolutionary patterns (Huyse *et al.* 2005; Barrett *et al.* 2008; Hoberg & Brooks 2008).

The combined effects of elevated mutation rates, smaller effective population sizes and limited dispersal abilities generally observed in parasites are predicted to

reduce gene flow and increase genetic drift, thereby leading to more pronounced spatial genetic structure in parasites when compared with their hosts (Criscione & Blouin 2005; Huyse *et al.* 2005; Nieberding & Morand 2006; Matthee *et al.* 2007; Nieberding *et al.* 2010). In our study, the *P. arvicanthis* lineages do not fulfil the expectation of more pronounced spatial genetic structure, when compared with *Rhabdomys* (Nieberding & Morand 2006; Nieberding *et al.* 2010). In fact, geographic genetic structure is less pronounced in both parasite lineages when compared with *Rhabdomys*, as indicated by the high incidence of shared haplotypes among sampled localities (Fig. 2). Estimated mitochondrial COI evolutionary rates are elevated in *Polyplax* when compared with *Rhabdomys* (approximately 2.1 and 2.5 times higher for *P. arvicanthis* 1 and 2, respectively). Similar trends have been previously reported for chewing and sucking lice (see Hafner *et al.* 1994; Page & Hafner 1996; Page *et al.* 1998, 2004; Paterson & Banks 2001; Light & Hafner 2007, 2008), and it is postulated to result from the shorter generation times of lice leading to a faster accumulation of mutations (Hafner *et al.* 1994). Elevated mutation rates within the *P. arvicanthis* lineages are also corroborated by the higher levels of sequence diversity when compared with *Rhabdomys* (Table 2). Contrary to the expectation that the female-biased sex ratio (Matthee *et al.* 2007) and direct life cycle (Ledger 1980; Kim 2006) of *Polyplax* will promote smaller effective populations sizes compared with the host (Rannala & Michalakis 2003; Criscione & Blouin 2005; Huyse *et al.* 2005), our data indicates larger estimated mitochondrial effective population sizes for *P. arvicanthis* 1 and 2 when compared with *Rhabdomys* (Table 2). This suggests that in the current study system, sexual reproduction, the absence of species-specificity (see below) and large interconnected host populations probably act in concert to increase the effective population size of the parasites (Barrett *et al.* 2008). Furthermore, the theoreti-
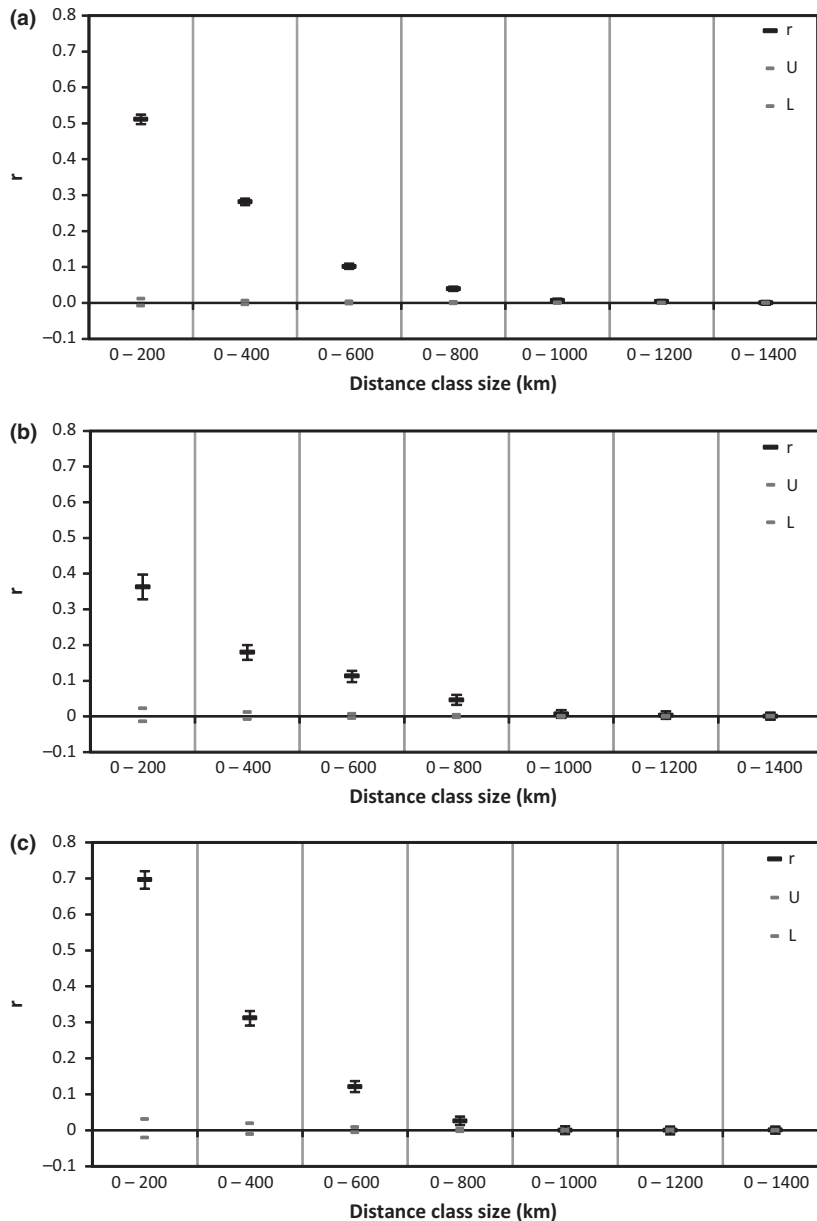
Table 3 Results from 3-level hierarchical analyses of molecular variance for the mitochondrial and nuclear data sets of the parasites and host

	Fixation index			% Variation		
	ΦST	ΦCT	ΦSC	Among clusters	Among sites within clusters	Within sites
mtDNA						
Rhabdomys	**0.97**	**0.94**	**0.60**	94.45	3.37	2.18
Polyplax arvicanthis 1	**0.96**	**0.96**	**0.19**	95.53	0.87	3.60
Polyplax arvicanthis 2	**0.90**	**0.87**	**0.31**	86.75	4.18	9.07
nDNA						
Rhabdomys	**0.29**	**0.21**	**0.10**	20.78	8.50	70.72
Polyplax arvicanthis 1	**0.43**	**0.37**	**0.11**	36.87	6.95	56.18
Polyplax arvicanthis 2	**0.40**	**0.28**	**0.17**	27.87	12.56	59.57

Significant values ($P < 0.05$) are indicated in bold.

Table 4 Mantel test and partial Mantel test results for the host and parasite mtDNA data sets. Correlation coefficients (*r*) and statistical significance (*p*) resulting from 10 000 permutations are indicated

	Mantel test*		Partial Mantel test[†]			
	H$_1$: geography		H$_1$: clusters		H$_2$: clusters (geography)	
	r	*p*	*r*	*p*	*r*	*p*
Rhabdomys	0.11	0.001	0.7	0.001	0.65	0.001
Polyplax arvicanthis 1	0.02	0.169	0.68	0.0002	0.67	0.0003
Polyplax arvicanthis 2	0.05	0.006	0.76	0.001	0.74	0.001

*H$_0$: Geographic and genetic distance is independent, H$_1$: Genetic distance increases with geographic distance.
[†]H$_0$: Genetic distance and cluster membership are independent, H$_1$: Genetic distance is greater among clusters, H$_2$: Genetic distance is greater among clusters, corrected for the effect of geography (distance).

Fig. 5 Correlograms indicating the average genetic autocorrelation coefficient (r) as a function of increasing distance class size, for *Rhabdomys* (a), *Polyplax arvicanthis* 1 (b) and *Polyplax arvicanthis* 2 (c). Error bars indicate the 95% confidence interval around the observed r values, and grey dash marks indicate the 95% confidence interval surrounding the null hypothesis of no spatial autocorrelation (r = 0).

cal framework suggests that parasite features may act to reduce N_e below what is expected for free-living populations of equal census size (Criscione & Blouin 2005), and the census size of *Polyplax* is expected to be much larger than that of the host as a result of the high abundance on *Rhabdomys* (Matthee *et al.* 2010). The lower levels of divergence among parasite lineages (when compared with the hosts) may thus be partially explained by the larger effective population sizes of *Polyplax* counteracting the effects of the elevated mutation rates observed. Furthermore, despite the predicted low dispersal potential of *P. arvicanthis* due to its life history traits (Hopkins 1949; Kim 2006), the autocorrelation analyses indicated that the inferred levels of

historical gene flow of the parasites (particularly over shorter distances) are more extensive or only slightly more restricted compared with the host (Fig. 5).

In sucking lice, dispersal occurs predominantly through interhost contact from adult to offspring and during social interactions, but may also occur via shared nests or burrows (Ledger 1980; Marshall 1981). In the present study, the differentiation between parasite assemblages break down in the contact zones at Fort Beaufort (FB_sk and FB_wb; Fig. 3) where lice parasitizing *R. pumilio* and *R. dilectus* share haplotypes. It is thus evident that despite their presumable permanency and vertical mode of transmission, lice are able to disperse among host taxa within the contact zones. This implies

Table 5 Outcome of the cophylogenetic analyses in JANE, employing various different cost schemes. The number of the different event types retrieved and the statistical significance based on 1000 randomizations are indicated for each reconstruction

Cost scheme	Event costs*	Total cost	Codivergence	Duplication	Host switch	Parasite loss	Failure to diverge	*P*-value
JANE version 4 default	11211	21	1	5	5	3	2	0.89
TREEMAP version 2b default	0111(1)	13	2	2	7	2	2	0.74
Treefitter default	0021(1)	12	2	8	1	8	2	0.90
Equal weights	11111	15	2	2	7	2	2	0.70

*Event costs are ordered as codivergence, duplication, host switch, sorting, and failure to diverge. Values in brackets were arbitrarily set to 1 as failure to diverge events is implemented exclusively in Jane.
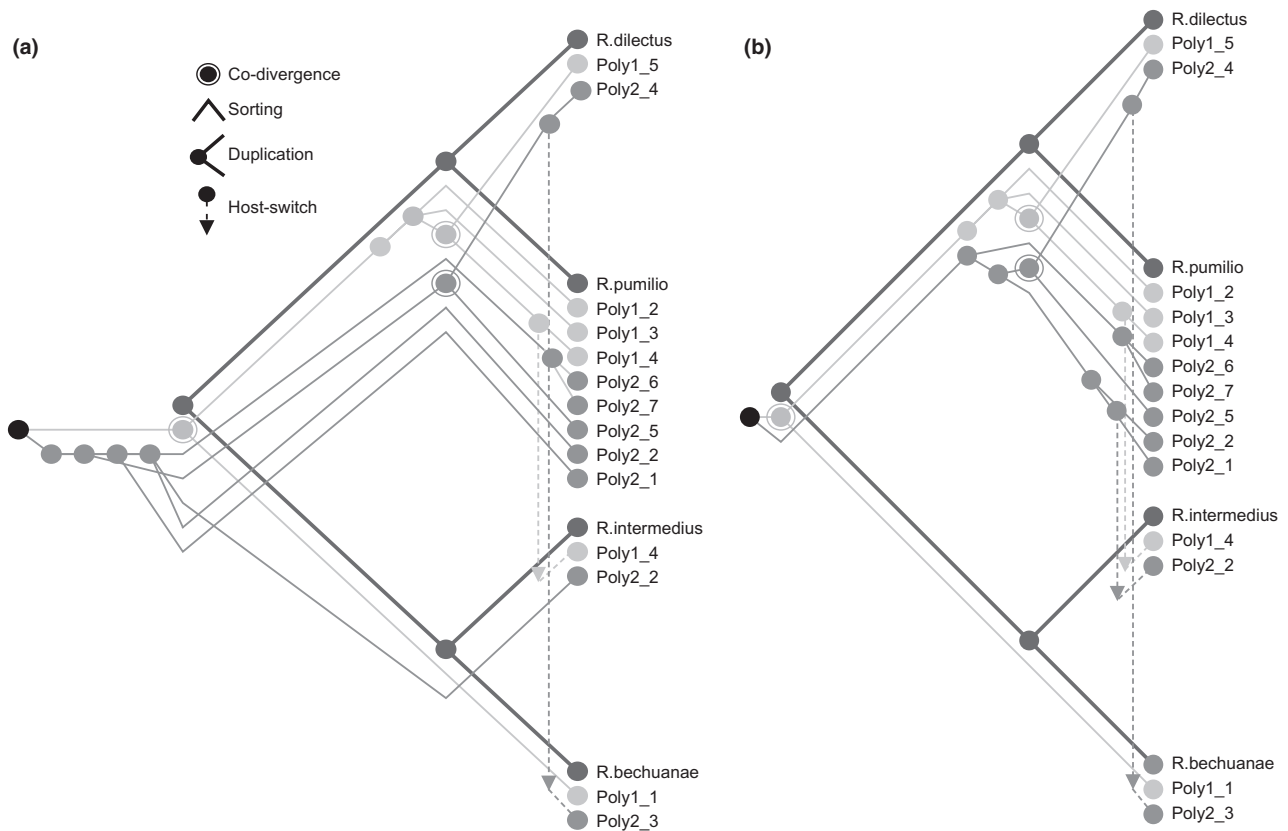
Fig. 6 The preferred reconstruction (a) and reconstruction with the second best-cost model fit (b) of the coevolutionary history of *Polyplax* and *Rhabdomys* retrieved from CORE-PA. Numbers following underscores indicate the respective *P. arvicanthis* 1 and 2 clusters.

that both host sympatry and syntopy (Rivas 1964; McCoy 2003) probably occur within these contact zones, as dispersal is expected to occur mostly via direct contact among hosts (Ledger 1980; Marshall 1981). There is also another host contact zone near Rooipoort (RP) in the Free State Province among *R. dilectus* and *R. bechuanae* (Ganem *et al.* 2012; du Toit *et al.* 2012) that could potentially facilitate gene flow especially within *P. arvicanthis* 2, for which the lice from these hosts are sister taxa and indicate a moderate amount of gene flow at the nuclear level (Figs 2 and 4). Further, we hypothesize that it is

indeed possible that more extensive geographic sampling within the distribution of *R. dilectus* and *R. bechuanae* may reveal breakdown in the reciprocal monophyly of the parasite mtDNA haplogroups associated with these species. Theory predicts that the evolutionary and biogeographic history of a parasite and host determines the range of hosts that can be exploited by the parasite (Poulin & Keeney 2007), and it is expected that the horizontal transmission of the parasite among divergent host lineages will prohibit the development of host-specificity (Nieberding & Olivieri 2007). In the present study, it

seems reasonable to suggest that the absence of genetic differentiation of parasites according to host lineages within contact zones indicates that *P. arvicanthis* is not adapted to specific hosts, and host switching frequently occurs where the opportunity arises. The previously documented host-specificity for this taxon is thus merely an index describing the current perceived association between *P. arvicanthis* and *Rhabdomys* (Ledger 1980; Clayton *et al*. 2004).

The dispersal of *Polyplax* is expected to be largely dependent on host vagility, given the importance of interhost contact (Criscione *et al*. 2005). Dispersal will thus be more restricted by a solitary lifestyle (Demastes *et al*. 2012) and enhanced by social contact among host individuals (Huyse *et al*. 2005). Some evidence exists to show that the arid-adapted *R. pumilio* is social and group-living while the mesic-adapted *R. dilectus* is more solitary (Schradin & Pillay 2005). Although it has not been explicitly tested, it follows that *R. intermedius,* which is also distributed in arid environments will probably also be social in nature. This premise is mainly based on a higher similarity in vegetation and climatic conditions between regions occupied by *R. pumilio* and *R. intermedius* (Mucina & Rutherford 2006). Host social interactions and shared nests are expected to facilitate *Polyplax* gene flow, and interestingly within both *P. arvicanthis* lineages, shared haplotypes occur mostly among the lice of the social *R. intermedius* and *R. pumilio* (Figs 2–4). While the females of *R. pumilio* tend to be philopatric, males may stay in their natal group as helpers or disperse, to either become the territorial breeding male of another group or solitary-living roamers (Solmsen *et al*. 2011). *Rhabdomys* males are reported to harbour more *P. arvicanthis* individuals than females, possibly as a result of testosterone-mediated immune depression (Matthee *et al*. 2010). Because solitary roaming *Rhabdomys* males have the highest reported levels of testosterone (Schradin *et al*. 2009), we propose that these individuals may be predominantly responsible for facilitating the high levels of *Polyplax* dispersal across the geographic landscape. If these arguments hold, our findings strongly support the premise that host life history characteristics (social behaviour and mobility) are equally important in shaping parasite population structure and determining the extent of genealogical congruence (Weckstein 2004; Huyse *et al*. 2005; Barrett *et al*. 2008).

Evolutionary relationships among the four *Rhabdomys* species, and the haplogroups detected for *P. arvicanthis* 1 and 2 were incongruent at both micro- and macroevolutionary levels according to the distance-based (global-fit) approach of PARAFIT. Event-based reconstructions of macroevolutionary relationships in JANE varied according to the cost scheme employed (Table 5) reiterating

that the outcome of cophylogeny reconciliation analyses is highly dependent on the cost scheme used (Merkle *et al*. 2010). All reconstructions, however, indicated non-significant cophylogeny with either predominantly duplications with sorting events or host switches. In instances, such as the current study system, where determining a biologically relevant cost scheme a priori is difficult, the parameter-adaptive approach of CORE-PA is a valuable alternative (De Vienne *et al*. 2013). CORE-PA also indicated nonsignificant cophylogeny among *Polyplax* and *Rhabdomys*. The two reconstructions with the best-cost model fit (Fig. 6) are dominated by duplication and sorting events with limited co-divergence and host-switching events. The major difference between the two scenarios was limited to the timing of duplication events within *P. arvicanthis* 2 with respect to host divergences. It should be noted, however, that CoRe-PA has a tendency to underestimate the number of host switches (Keller-Schmidt *et al*. 2011). Thus, while it is evident that codivergence did not significantly shape the coevolutionary history of *Polyplax* and *Rhabdomys*, the relative roles of host-switching, duplication and sorting events are not clear.

The limited genealogical congruence among *P. arvicanthis* and *Rhabdomys* at both micro- and macroevolutionary scales is probably also a direct result of the strong influence of biogeographic history, potentially also coupled to incomplete lineage sorting as a result of recent cladogenesis (Rannala & Michalakis 2003; Nieberding & Olivieri 2007; Demastes *et al*. 2012). Episodes of environmental change have been suggested as the main drivers for diversification in parasite–host systems by inducing cyclical episodes of expansion and contraction in geographic ranges ('Taxon pulse hypothesis'; Halas *et al*. 2005; Hoberg & Brooks 2008). Such expansion–contraction cycles may facilitate codivergence during periods of host allopatry and promote host switching during periods of sympatric contact (Clayton *et al*. 2004; Weckstein 2004; Brooks & Ferrao 2005). Temporary contraction and fragmentation of host ranges may also provide a mechanism for parasite duplication, if parasites were to diverge without accompanying host divergence (Johnson & Clayton 2004; Banks & Paterson 2005; Štefka & Hypša 2008). Parasite divergence in the absence of host divergence may be promoted by comparatively faster parasite mutation rates (as was found here), which would increase the likelihood of the fixation of differences within a given time period (Frankham *et al*. 2002).

Cladogenesis and ecological divergence within *Rhabdomys* likely occurred as a result of biogeographic changes in response to climatic and landscape changes around the Mio-Pliocene boundary (Scott *et al*. 1997; Tyson & Partridge 2000; Mucina & Rutherford 2006; du

Toit *et al.* 2012), and subsequent climatic oscillations may have led to range expansion–contraction cycles (Ellery *et al.* 1991; Partridge 1997; Zachos 2001; Scott & Nyakale 2002; Dupont 2006; Montgelard & Matthee 2012; du Toit *et al.* 2012). It thus seems reasonable to suggest that biogeographic changes would have also significantly shaped the patterns of genetic divergence within *P. arvicanthis*. The occurrence of codivergence, duplication, sorting and host-switching events in association with particular host lineages is probably best explained by stochasticity coupled to the relative abundance of parasite lineages, as parasite abundance determines the likelihood of tracking host movements (also see Clayton *et al.* 2003, 2004; Nieberding & Morand 2006; Nieberding *et al.* 2010). Further, distinct mtDNA assemblages are present within both parasite lineages and the host in the arid northwestern region and mesic eastern region (see *P. arvicanthis* 1 cluster 1, *P. arvicanthis* 2 cluster 3 and *R. bechuanae* as well as *P. arvicanthis* 1 cluster 5, *P. arvicanthis* 2 cluster 4, and *R. dilectus*; Figs 2 and 3), which is also partly supported by the nuclear DNA data (Fig S1, Supporting information). Similar genetic disjunctions for nonrelated species in the region (Lamb & Bauer 2000; Matthee & Flemming 2002; Redman & Hamer 2003; Tolley *et al.* 2004, 2010; Bauer *et al.* 2006; Smit *et al.* 2007; Engelbrecht *et al.* 2011; Willows-Munro & Matthee 2011; Montgelard & Matthee 2012) support the premise that the geographic pattern observed in both *P. arvicanthis* lineages, and *Rhabdomys* is the result of vicariance and emphasizes a strong influence of biogeography.

From the current study, it is evident that biogeographic history coupled to host-related factors, such as vagility and sociality, has important consequences for the genetic outcome of parasite–host associations. We propose that a combination of these factors led to the limited congruence observed, despite the presence of parasite traits that are expected to favour congruence.

Acknowledgements

References

Akaike H (1973) Information theory and an extention of the maximum likelihood principle. In: *Second International Symposium on Information Theory* (eds Petrov PN & Csaki F), pp. 267–281. Budapest, Adad. Kiado.

Althoff DM, Segraves KA, Smith CI, Leebens-Mack J, Pellmyr O (2012) Geographic isolation trumps coevolution as a driver of yucca and yucca moth diversification. *Molecular Phylogenetics and Evolution*, **62**, 898–906.

Banks JC, Paterson AM (2005) Multi-host parasite species in cophylogenetic studies. *International Journal for Parasitology*, **35**, 741–746.

Barrett LG, Thrall PH, Burdon JJ, Linde CC (2008) Life history determines genetic structure and evolutionary potential of host-parasite interactions. *Trends in Ecology and Evolution*, **23**, 678–685.

Bauer AM, Lamb T, Branch WR (2006) A revision of the *Pachydactylus serval* and *P. weberi* groups (Reptilia: Gekkota: Gekkonidae) of Southern Africa, with the description of eight new species. *Proceedings of the California Academy of Sciences*, **57**, 595–709.

Blouin MS, Yowell CA, Courtney CH, Dame JB (1995) Host movement and the genetic structure of populations of parasitic nematodes. *Genetics*, **141**, 1007–1014.

Bohonak AJ (2002) IBD (isolation by distance): a program of analyses of isolation by distance. *Journal of Heredity*, **93**, 153–154.

Brooks DR (1979) Testing the context and extent of host-parasite coevolution. *Systematic Zoology*, **28**, 299–307.

Brooks DR, Ferrao AL (2005) The historical biogeography of co-evolution: emerging infectious diseases are evolutionary accidents waiting to happen. *Journal of Biogeography*, **32**, 1291–1299.

Brooks DR, McLennan DA (1993) *Parascript*. Parasites and the language of evolution Smithsonian Institution Press, Washington, District of Columbia

Burnham KP, Anderson DA (2004) Multi-model inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, **33**, 261–304.

Charleston MA (2011) TreeMap 3b. Available from [http://sites.google.com/site/cophylogeny].

Charleston MA, Perkins SL (2006) Traversing the tangle: algorithms and applications for cophylogenetic studies. *Journal of Biomedical Informatics*, **39**, 62–71.

Clayton DH (1990) Host specificity of *Strigiphilus* owl lice (Ischnocera: Philopteridae), with the description of new species and host associations. *Journal of Medical Entomology*, **27**, 257–265.

Clayton DH, Johnson KP (2003) Linking coevolutionary history to ecological process: doves and lice. *Evolution*, **57**, 2335–2341.

Clayton DH, Al-Tamimi S, Johnson KP (2003) The ecological basis of coevolutionary history. In: *Tangled Trees: Phylogeny, Cospeciation, and Coevolution*(ed. Page RDM), pp. 330–350. University of Chicago Press, Chicago.

Clayton DH, Bush SE, Johnson KP (2004) Ecology of congruence: Past meets present. *Systematic Biology*, **53**, 165–173.

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

Conow C, Fielder D, Ovadia Y, Libeskind-Hadas R (2010) Jane: a new tool for the cophylogeny reconstruction problem. *Algorithms for Molecular Biology*, **5**, 16.

Corander J, Tang J (2007) Bayesian analysis of population structure based on linked molecular information. *Mathematical Biosciences*, **205**, 19–31.

Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics*, **9**, 539.

Criscione CD, Blouin MS (2005) Effective sizes of macroparasite populations: a conceptual model. *Trends in Parasitology*, **21**, 212–217.

Criscione CD, Poulin R, Blouin MS (2005) Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Molecular Ecology*, **12**, 2247–2257.

De Vienne DM, Refrégier G, López-Villavicencio M, Tellier A, Hood ME, Giraud T (2013) Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. *New Phytologist*, **198**, 347–385.

Demastes JW, Spradling TA, Hafner MS et al. (2012) Cophylogeny on a fine scale: *Geomydoecus* chewing lice and their pocket gopher hosts, *Pappogeomys bulleri*. *Journal of Parasitology*, **98**, 262–270.

Dilcher M, Hasib L, Lechner M et al. (2011) Genetic characterization of Tribeč virus and Kemerovo virus, two tick-transmitted human-pathogenic Orbiviruses. *Virology*, **423**, 68–76.

Drummond CS, Hamilton MB (2007) Hierarchical components of genetic variation at a species boundary: population structure in two sympatric varieties of *Lupinus microcarpus* (Leguminosae). *Molecular Ecology*, **16**, 753–769.

Drummond AJ, Ho S, Phillips M, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLOS Biology*, **4**, e88.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUTi and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Dupont LM (2006) Late Pliocene vegetation and climate in Namibia (southern Africa) derived from palynology of ODP Site 1082. *Geochemistry, Geophysics, Geosystems*, **7**, 1–15.

Durden LA, Musser GG (1994) The sucking lice (Insecta, Anoplura) of the world: a taxonomic checklist with records of Mammalian hosts and geographical distributions. *Bulletin of the American Museum of Natural History*, **218**, 1–90.

Ellery WN, Scholes RJ, Mentis MT (1991) An initial approach to predicting the sensitivity of the grassland biome to climate change. *South African Journal of Science*, **87**, 499–503.

Engelbrecht A, Taylor PJ, Daniels SR, Rambau RV (2011) Cryptic speciation in the southern African vlei rat *Otomys irroratus* complex: evidence derived from mitochondrial cyt b and niche modelling. *Biological Journal of the Linnean Society*, **104**, 192–206.

Excoffier L, Lischer HEL (2010) ARLEQUIN suite ver 3.5: a new series of programs to perform population genetic analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.

Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance ? *Journal of Applied Ecology*, **46**, 493–505.

Ganem G, Meynard CN, Perigault M et al. (2012) Environmental correlates of co-occurrence of three mitochondrial lineages of striped mice (*Rhabdomys*) in the Free State Province (South Africa). *Acta Oecologica*, **42**, 30–40.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

Hafner MS, Sudman PD, Villablanca FX et al. (1994) Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science*, **265**, 1087–1090.

Halas D, Zamparo D, Brooks DR (2005) A historical biogeographical protocol for studying diversification by taxon pulses. *Journal of Biogeography*, **32**, 249–260.

Hall T (2005) ioEdit, Biological sequence alignment editor for Win95/98/NT/2K/XP. Available from http://www.mbio.ncsu.edu/BioEdit/bioedit.html.

Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.

Hoberg EP, Brooks DR (2008) A macroevolutionary mosaic: episodic host-switching, geographical colonization and diversification in complex host-parasite systems. *Journal of Biogeography*, **35**, 1533–1550.

Hoberg EP, Klassen GJ (2002) Revealing the faunal tapestry: co-evolution and historical biogeography of hosts and parasites in marine systems. *Parasitology*, **124**, 3–22.

Hopkins GHE (1949) The host association of the lice of mammals. *Proceedings of the Zoological Society, London*, **119**, 387–604.

Huyse T, Poulin R, Theron A (2005) Speciation in parasites: a population genetics approach. *Trends in Parasitology*, **21**, 469–475.

Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. *BMC Genetics*, **6**, 13.

Jerome CA, Ford BA (2002) Comparative population structure and genetic diversity of *Arceuthobium americanum* (Viscaceae) and its *Pinus* host species: insight into host-parasite evolution in parasitic angiosperms. *Molecular Ecology*, **11**, 407–420.

Johnson KP, Clayton DH (2004) Untangling coevolutionary history. *Systematic Biology*, **53**, 92–94.

Johnson KP, Williams BL, Drown DM, Adams RJ, Clayton DH (2002) The population genetics of host specificity: genetic differentiation in dove lice (Insecta: Phthiraptera). *Molecular Ecology*, **11**, 25–38.

Johnson KP, Adams RJ, Page RDM, Clayton DH (2003) When do parasites fail to speciate in response to host speciation? *Systematic Biology*, **52**, 37–47.

Keller-Schmidt S, Wieseke N, Klemm K, Middendorf M (2011) Evaluation of host parasite reconciliation methods using a new apporach for cophylogeny generation. Working paper

from Bioinformatics Leipzig. Available from http://www.bioinf.uni-leipzig.de/working/11-013.

Kim KC (1985) *Coevolution of Parasitic Arthropods and Mammals.* John Wiley, New York.

Kim KC (2006) Blood-sucking lice (Anoplura) of small mammals: True parasites. In: *Micromammals and Macroparasites: From Evolutionary Ecology to Management* (eds Morand S, Krasnov BR & Poulin R), pp. 141–160. Springer-Verlag, Tokyo.

Lamb T, Bauer AM (2000) Relationships of the *Pachydactylus rugosus* group of geckos (Reptilia: Squamata: Gekkonidae). *African Zoology*, **35**, 55–67.

Ledger JA (1980) *The Arthropod Parasites of Vertebrates in Africa south of the Sahara. Volume IV. Phthiraptera (Insecta). Publications of the South African Institute for Medical Research No. 56.* South African Institute for Medical Research, Johannesburg, South Africa.

Legendre P, Legendre L (1998) *Numerical Ecology*, 2nd edn. Elsevier, New York.

Legendre P, Desdevises Y, Bazin E (2002) A statistical test for host-parasite coevolution. *Systematic Biology*, **51**, 217–234.

Librado P, Rozas J (2009) DNASP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Light JE, Hafner MS (2007) Disparate rates of evolution in sympatric lineages of chewing lice on pocket gophers. *Molecular Phylogenetics and Evolution*, **45**, 997–1013.

Light JE, Hafner MS (2008) Codivergence in Heteromyid rodents (Rodentia: Heteromyidae) and their sucking lice of the genus *Fahrenholzia* (Phthiraptera: Anoplura). *Systematic Biology*, **57**, 449–465.

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

Marshall AG (1981) *The Ecology of Ectoparasitic Insects.* Aberdeen University Academic Press, London.

Matthee CA, Flemming AF (2002) Population fragmentation in the southern rock agama, Agama atra: more evidence for vicariance in Southern Africa. *Molecular Ecology*, **11**, 465–471.

Matthee S, Horak IG, Beaucournu J-C *et al.* (2007) Epifaunistic arthropod parasites of the four-striped mouse, *Rhabdomys pumilio*, in the Western Cape Province, South Africa. *Journal of Parasitology*, **93**, 47–59.

Matthee S, McGeoch MA, Krasnov B (2010) Parasite-specific variation and the extent of male-biased parasitism; an example with a South African rodent and ectoparasitic arthropods. *Parasitology*, **137**, 651–660.

McCoy KD (2003) Sympatric speciation in parasites – what is sympatry? *Trends in Parasitology*, **19**, 400–404.

McCoy KD, Boulinier T, Tirard C, Michalakis Y (2003) Host-dependant genetic structure of parasite populations: differential dispersal of seabird tick host races. *Evolution*, **57**, 288–296.

Meier-Kolthoff JP, Auch AF, Huson DH, Göker M (2007) Copy-cat: Co-phylogenetic Analysis Tool. *Bioinformatics*, **23**, 898–900.

Meirmans PG (2012) The trouble with isolation by distance. *Molecular Ecology*, **21**, 2839–2846.

Mendlová M, Desdevises Y, Civáňová K, Pariselle A, Šimková A (2012) Monogeneans of West African cichlid fish: evolution and cophylogenetic interactions. *PLoS ONE*, **7**, e37268.

Merkle D, Middendorf M, Wieseke N (2010) A parameter-adaptive dynamic programming aproach for inferring cophylogenies. *BMC Bioinformatics*, **11**, S60.

Montgelard C, Matthee CA (2012) Tempo of genetic diversification in southern African rodents: The role of Plio-Pleistocene climatic oscillations as drivers for speciation. *Acta Oecologica*, **42**, 50–57.

Mucina L, Rutherford MC (2006) *The Vegetation of South Africa.* Lesotho and Swaziland, South African National Biodiversity Institute, Pretoria.

Nadler SA (1995) Microevolution and the genetic structure of parasite populations. *Journal of Parasitology*, **81**, 395–403.

Nieberding CM, Morand S (2006) Comparative phylogeography: The use of parasites for insights into host history. In: *Micromammals and Macroparasites: From Evolutionary Ecology to Management* (eds Morand S, Krasnov BR & Poulin R), pp. 277–293. Springer-Verlag, Tokyo.

Nieberding CM, Olivieri I (2007) Parasites: proxies for host genealogy and ecology? *Trends in Ecology & Evolution*, **22**, 156–165.

Nieberding CM, Durette-Desset MC, Vanderpoorten A *et al.* (2008) Geography and host biogeography matter for understanding the phylogeography of a parasite. *Molecular Phylogenetics and Evolution*, **47**, 538–554.

Nieberding CM, Jousselin E, Desdevises Y (2010) The use of co-phylogeographic patterns to predict the nature of host-parasite interactions, and *vice versa*. In: *The Biogeography of Host-Parasite Interactions* (eds Morand S & Krasnov BR), pp. 59–69. Oxford University Press, Oxford.

Page RDM (2003) *Tangled trees: Phylogeny, Cospeciation, and Coevolution.* University of Chicago press, Chicago.

Page DM, Hafner MS (1996) Molecular phylogenies and host-parasite co-speciation: gophers and lice as a model system. In: *New Uses for New Phylogenies* (eds Harvey PH, Brown AJL, Smith JM & Nee S), pp. 255–270. Oxford University Press, Oxford.

Page RDM, Lee PLM, Becher SA, Griffiths R, Clayton DH (1998) A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. *Molecular Phylogenetics and Evolution*, **9**, 276–293.

Page RDM, Cruickshank RH, Dickens M *et al.* (2004) Phylogeny of "*Philoceanus* complex" seabird lice (Phthiraptera: Ischnocera) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **30**, 633–652.

Partridge TC (1997) Evolution of landscapes. In: *Vegetation of Southern Africa* (eds Cowling RM, Richardson DM & Pierce SM), pp. 5–20. Cambridge University Press, Cambridge.

Paterson AM, Banks J (2001) Analytical approaches to measuring cospeciation of host and parasites: through a looking glass, darkly. *International Journal for Parasitology*, **31**, 1012–1022.

Paterson AM, Wallis GP, Wallis LJ, Gray RD (2000) Seabird and louse co-evolution: complex histories revealed by 12S rDNA sequences and reconciliation analyses. *Systematic Biology*, **49**, 383–399.

Peakall ROD, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peakall R, Smouse PE (2012) GENALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. *Bioinformatics*, **28**, 2537–2539.

Peakall R, Ruibal M, Lindenmayer DB (2003) Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution*, **57**, 1182–1195.

Posada D (2004) COLLAPSE 1.2: Describing haplotypes from sequence alignments. Available from http://darwin.uvigo.es/software/collapse.html.

Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Poulin R, Keeney DB (2007) Host specificity under molecular and experimental scrutiny. *Trends in Parasitology*, **24**, 24–28.

Rambaut A, Drummond AJ (2007) Tracer v1.5. Available from http://beast.bio.ed.ac.uk/Tracer.

Rannala B, Michalakis Y (2003) Population genetics and cospeciation: from process to pattern. In: *Tangled trees. Phylogeny, Cospeciation and Coevolution* (ed. Page RDM), pp. 120–143. Chicago University Press, Chicago.

Redman GT, Hamer ML (2003) The distribution of southern African Harpagophoridae Attems, 1909 (Diplopoda: Spirostreptida). *African Invertebrates*, **44**, 213–226.

Reed DL, Smith VS, Hammond SL, Rogers AR, Clayton DH (2004) Genetic analysis of lice supports direct contact between modern and archaic humans. *PLOS Biology*, **2**, 1972–1983.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Rivas LR (1964) A reinterpretation of the concepts "sympatric" and "allopatric" with proposal of the additional terms "syntopic" and "allotopic". *Systematic Zoology*, **13**, 42–43.

Ronquist F, Teslenko M, van der Mark P *et al.* (2012) MRBAYES 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Rosenblueth M, Sayavedra L, Sámano-Sánchez H, Roth A, Martínez-Romero E (2013) Evolutionary relationships of flavobacterial and enterobacterial endosymbionts with their scale insect hosts (Hemiptera: Coccoidea). *Journal of Evolutionary Biology*, **25**, 2357–2368.

Safner T, Miller MP, McRae BH, Fortin M-J, Manel S (2011) Comparison of Bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. *International Journal of Molecular Sciences*, **12**, 865–889.

Schradin C, Pillay N (2005) Intraspecific variation in the spatial and social organization of the African striped mouse. *Journal of Mammalogy*, **86**, 99–107.

Schradin C, Scantlebury M, Pillay N, König B (2009) Testosterone levels in dominant sociable males are lower than in solitary roamers: physiological differences between three male reproductive tactis in a socially flexible mammal. *The American Naturalist*, **173**, 376–388.

Scott L, Nyakale M (2002) Pollen indicators of Holocene palaeoenvironments at Florisbad in the central Free State, South Africa. *The Holocene*, **12**, 497–503.

Scott L, Anderson HM, Anderson JM (1997) Vegetation history. In: *Vegetation of Southern Africa* (eds Cowling RM, Richardson DM & Pierce SM), pp. 62–84. Cambridge University Press, Cambridge.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, **16**, 1114–1116.

Silvestro D, Michalak K (2012) raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution*, **12**, 335–337.

Skinner JD, Chimimba CT (2005) *The Mammals of the South African subregion*, 3rd edn. Cambridge University Press, Cambridge.

Smit HA, Robinson TJ, Jansen van Vuuren B (2007) Coalescence methods reveal the impact of vicariance on the spatial genetic structure of *Elephantulus edwardii* (Afrotheria, Macroscelidea). *Molecular Ecology*, **16**, 2680–2692.

Smouse PE, Long JC (1992) Matrix correlation-analysis in anthropology and genetics. *Yearbook Of Physical Anthropology*, **35**, 187–213.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity*, **82**, 561–573.

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, **35**, 627–632.

Solmsen N, Johannesen J, Schradin C (2011) Highly asymmetric fine-scale genetic structure between sexes of African striped mice and indication for condition dependent alternative male dispersal tactics. *Molecular Ecology*, **20**, 1624–1634.

Stadler T (2010) Sampling-through-time in birth-death trees. *Journal of Theoretical Biology*, **267**, 396–404.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analysis with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2268–2690.

Stamatakis A, Auch AF, Meier-Kolthoff JP, Göker M (2007) AxPcoords & parallel AxParafit: statistical co-phylogenetic analyses on thousands of taxa. *BMC Bioinformatics*, **8**, 405.

Štefka J, Hypša V (2008) Host specificity and genealogy of the louse *Polyplax serrata* on field mice, *Apodemus* species: a case of parasite duplication or colonization? *International Journal for Parasitology*, **38**, 731–741.

Štefka J, Hoeck PEA, Keller LF, Smith VS (2011) A hitchhikers guide to the Galápagos: co-phylogeography of Galápagos mockingbirds and their parasites. *BMC Evolutionary Biology*, **11**, 284.

Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *American Journal of Human Genetics*, **76**, 449–462.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.

Swofford DL (2000) *PAUP: Phylogenetic Analysis Using Parsimony (and other methods)*. Sinauer Associates, Sunderland, Massachusetts.

du Toit N, Jansen van Vuuren B, Matthee S, Matthee CA (2012) Biome specificity of distinct genetic lineages within the four-striped mouse *Rhabdomys pumilio* (Rodentia: Muridae) from southern Africa with implications for taxonomy. *Molecular Phylogenetics and Evolution*, **65**, 75–86.

du Toit N, Matthee S, Matthee CA (2013) The sympatric occurrence of two genetically divergent lineages of sucking louse, *Polyplax arvicanthis* (Phthiraptera: Anoplura), on the four-striped mouse genus, *Rhabdomys* (Rodentia: Muridae). *Parasitology*, **140**, 604–616.

Tolley KA, Tilbury CR, Branch WR, Matthee CA (2004) Phylogenetics of the southern African dwarf chameleons, Bradypodion (Squamata:Chamaeleonidae). *Molecular Phylogenetics and Evolution*, **30**, 354–365.

Tolley KA, Braae A, Cunningham M (2010) Phylogeography of the Clicking Stream Frog *Strongylopus grayii* (Anura, Pyxicephalidae) reveals cryptic divergence across climatic zones in an abundant and widespread taxon. *African Journal of Herpetology*, **59**, 17–32.

Tyson PD, Partridge TC (2000) The evolution of Cenozoic climates. In: *The Cenozoic of Southern Africa* (eds Partridge TC & Maud RR), pp. 371–387. Oxford University Press, New York.

Weckstein JD (2004) Biogeography explains cophylogenetic patterns in toucan chewing lice. *Systematic Biology*, **53**, 154–164.

Whiteman NK, Parker PG (2005) Using parasites to infer host population history: a new rationale for parasite conservation. *Animal Conservation*, **8**, 175–181.

Willows-Munro S, Matthee CA (2011) Linking lineage diversification to climate and habitat heterogeneity: phylogeography of the southern African shrew *Myosorex varius*. *Journal of Biogeography*, **38**, 1976–1991.

Wirth T, Meyer A, Achtman M (2005) Deciphering host migrations and origins by means of their microbes. *Molecular Ecology*, **14**, 3289–3306.

Zachos J (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*, **292**, 686–693.

Data accessibility

DNA sequence data deposited in GenBank, see supplementary online material for accession numbers. Sequence alignments available on Dryad (doi:10.5061/dryad.j0p62).

Supporting information

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

Table S1 Mitochondrial COI haplotypes obtained from the various host taxa sampled (locality codes as in Table 1) with GenBank accession numbers.

Table S2 Mitochondrial COI haplotypes identified within the two *Polyplax arvicanthis* lineages from the various sampled localities (codes as in Table 1) with GenBank accession numbers.

Table S3 Primers used for PCR amplification of parasite and host mitochondrial and nuclear genes.

Table S4 Models of evolution specified for the various codon positions in the mtDNA Bayesian analyses of all haplotypes and the reduced datasets (clusters) used as input for the co-phylogeny analyses.

Table S5 The distinct genetic clusters identified from BAPS within the host and two parasite taxa (locality codes as in Table 1).

Table S6 Pairwise Φst values among *Rhabdomys* sampled localities with COI above and Eef1a1 below the diagonal.

Table S7 Φst values of the pairwise comparisons among parasite localities for COI with *P. arvicanthis* 1 above and *P. arvicanthis* 2 below the diagonal.

Table S8 Nuclear CAD pairwise Φst values among sampled localities with *P. arvicanthis* 1 above and *P. arvicanthis* 2 below the diagonal.

Fig. S1 Nuclear statistical parsimony networks (CAD gene) for *P. arvicanthis* 1 and 2 with haplotypes coloured according to mtDNA clusters as in Fig. 2.