MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 3052–3063 doi: 10.1111/j.1365-294X.2010.04681.x

Complete separation along matrilines in a social spider metapopulation inferred from hypervariable mitochondrial DNA region

I. AGNARSSON,*†‡§ W. P. MADDISON‡§ and L. AVILÉS‡

*Institute of Biology, Scientific Research Centre of the Slovenian Academy of Sciences and Arts, Novi trg 2, PO Box 306, SI-1001 Ljubljana, Slovenia, †Department of Biology, University of Puerto Rico, PO Box 23360, San Juan, PR 00931-3360, USA, ‡Department of Zoology, The University of British Columbia, 2370-6270 University Boulevard, Vancouver, BC, V6T 1Z4, Canada, §Department of Botany, The University of British Columbia, 3529-6270 University Boulevard, Vancouver, BC, V6T 1Z4, Canada

Abstract

The distribution and quantity of genetic diversity may be profoundly influenced by the emergence and dynamics of social groups. Permanent social living in spiders has resulted in the subdivision of their populations in more or less isolated colony lineages that grow, proliferate and become extinct without mixing with one another. A newly discovered hypervariable mitochondrial DNA region allowed us to examine the fine scale metapopulation structure in the social *Anelosimus eximius*. We sampled 39 colonies in Ecuador and French Guiana and identified 25 haplotypes. The majority of colonies contained one haplotype. Additional haplotypes occurred in approximately 15% of the colonies, and were always closely related to the common colony haplotype. Our findings confirm that colonies consist of single matrilines, with within-colony variation explained by mutations within the matriline. We thus found no evidence of mixing of matrilines. Likewise, colonies in a cluster often shared a haplotype, implying common colony ancestry. In few cases, however, haplotypes were shared between more distant colonies, providing evidence for occasional longer distance dispersal and/or widespread colony lineages. The geographical localities of colonies were incongruent with phylogenetic trees and haplotype networks, showing that some areas contained two or more matrilines. Hence, females do not migrate into foreign colonies, but faithfully remain within their own colony lineage, even when they disperse into new areas. These results indicate that the fine scale metapopulation structure of pure matrilines is maintained over the long term and that colony turnover is not extensive or radical enough to homogenize entire geographical areas. Genetic diversity is thus preserved to some extent at the metapopulation level.

Keywords: Anelosimus eximius, dispersal, inbred-social, matrilines, population structure

Received 15 June 2008; revision received 8 April 2010; accepted 15 April 2010

Introduction

The structuring of genetic diversity within species is influenced not only by environmental factors that control the sizes, ranges and fragmentation of populations (Bay *et al.* 2008; Kamiya *et al.* 2008; Ortega *et al.* 2008), but potentially also by behaviours that influence dispersal and breeding relationships (Hughes 1998; Storz 1999; Barrett *et al.* 2008). Social behaviour, in particular, may play a critical role in shaping the structure of genetic diversity in animal populations by causing members of one or both sexes to become associated in more or less permanent social groups (Storz 1999; Ross 2001; Hoelzel *et al.* 2008). Depending on the extent to which the social systems of males and females overlap,

Correspondence: I. Agnarsson, Fax: 787 7642610;
E-mail: iagnarsson@gmail.com

and on how social and breeding system relate to each other, such structuring may significantly influence the course of microevolutionary change (Storz 1999; Ross 2001) and the long-term evolutionary potential of many animal species (e.g. Agnarsson *et al.* 2006). In addition to introducing a new level of selection (Hamilton 1964; Wilson 1975; Maynard Smith & Szathmáry 1995; Ross 2001), the structuring of genetic diversity within and among groups may also change the balance between drift and selection by increasing the chances of fixation of random alleles when long-lived and expanding groups are established by few individuals (Storz 1999). Whether the net effect is increased or decreased genetic diversity in the overall metapopulation will in turn depend on the dynamics of group turnover and associated drift and selection at the level of the groups. In the long term, sociality may either increase evolutionary potential by allowing the colonization of new adaptive zones (Wilson 1987) or, in the cases when it is accompanied by inbreeding, it may constitute an evolutionary dead-end (Aviles 1997; Agnarsson *et al.* 2006). The study of sociality-modulated structuring is thus important to understand the set of evolutionary forces that shape many social animal species.

Although sociality could provide impetus for inbreeding, in general social organisms are outbred (Pusey & Wolf 1996; Cockburn 1998; Storz 1999; Ross 2001). And, although in many social species dispersal is often limited to short distances which may increase population genetic structuring, it seems sufficient to maintain genetic mixing at the species level (e.g. Lepais *et al.* 2010). Notably unusual in this regard are the non-territorial permanent social (*sensu* Aviles 1997) or, simply, social spiders (Aviles 1997; Lubin & Bilde 2007). Phylogenetic studies suggest that permanent sociality in spiders evolved by elimination of the dispersal phase that preceded mating in subsocial ancestral species (Johannesen & Lubin 1999, 2001; Agnarsson 2004, 2006; Agnarsson *et al.* 2006; Johannesen *et al.* 2007). Prior behavioural and genetic studies (Lubin & Robinson 1982; Riechert & Roeloffs 1993; Smith & Engel 1994; Smith & Hagen 1996; Johannesen *et al.* 2002, 2007, 2009) suggest that social spider colony lineages typically propagate with little or no mixing with one another. Colonies that reach a certain size give rise to daughter colonies by either fissioning, budding, or the emigration of single, or a small group of inseminated females (Aviles 1997, 2000; Crouch & Lubin 2001; Lubin & Robinson 1982; Roeloffs & Riechert 1988). The origin of spider sociality from outbred subsocial ancestors thus simultaneously altered population structure and breeding system (Riechert & Roeloffs 1993; Aviles 1997; Lubin & Bilde 2007). In other social systems, close and permanent inbreeding has only been documented in the eusocial mole rats (Jarvis *et al.* 1994), thrips (Chapman *et al.* 2000), and perhaps some ants that are social parasites or have atypical reproductive modes (reviewed in Ross 2001). Termites alternate phases of inbreeding and outbreeding (Shellman-Reeve 1997) and are thus quite distinct from the social spiders. Somewhat similar to the highly inbred social spiders are some clonal social insects, such as aphids, as clonality results in strong population genetic structuring. However, aphids differ from the social spiders in high levels of movement and frequent mixing of clones on host plants (Abbot 2009). Aphids also alternate sexual and asexual phases (Hales *et al.* 1997).

Driven by a high rate of colony turnover, evidenced by short lived colonies and colony replacement (Aviles 1993; Smith & Hagen 1996), the strongly subdivided population structure of the permanent-social spiders may result in sufficiently strong intercolony selection to produce and maintain the highly female-biased sex ratios characteristic of the majority of species (Aviles 1993, 1997, 1999). Other putative explanations of sex ratio bias, such as local mate competition, are insufficient as theoretical modelling predicts that within the colony lineages sex ratios should return towards a 1:1 equilibrium in the absence of intercolony selection (Aviles 1993). With some lineages proliferating and others going extinct, rapid colony turnover could then lead to some lineages sweeping through local areas. Thus, while population subdivision and inbreeding should lead to low genetic variability within the colony lineages (Johannesen *et al.* 2002, 2007, 2009), the latter process could result in a loss of genetic variability at the metapopulation or even species level (Aviles 1997), possibly leading to depressed diversification over long-term evolution (Agnarsson *et al.* 2006). Accordingly, allozyme studies have found allelic variation in social spiders to be generally low (see Table 12.5 in Riechert & Roeloffs 1993), and mostly limited to among-colony variation (Lubin & Crozier 1985; Roeloffs & Riechert 1988; Smith & Engel 1994; Smith & Hagen 1996).

The multiple origins of inbred sociality in spiders (Aviles 1997; Bilde *et al.* 2005; Agnarsson 2006; Agnarsson *et al.* 2006; Aviles *et al.* 2006; Johannesen *et al.* 2007; Lubin & Bilde 2007) provide an opportunity to study the short- and long-term evolutionary consequences of both sociality and inbreeding. Social origins are scattered across the order (Aviles 1997; Agnarsson *et al.* 2006; Lubin & Bilde 2007), but occur most frequently in two distantly related lineages, a group of cobweb spiders (Theridiidae) and in the genus *Stegodyphus* (Eresidae). Despite their vast phylogenetic distance, these lineages both have produced social species that are strikingly similar in their social structure. This

represents a unique opportunity to study the consequences of sociality and inbreeding both across related species where sociality has arisen in parallel, as well as across distantly related lineages that, convergently, are prone to generate these parallel social species. However, while theridiids have been relatively well studied phylogenetically (Agnarsson 2003, 2004; Arnedo *et al.* 2004, 2007), their population genetic structure is less well known, aside from allozymes (Smith & Engel 1994; Smith & Hagen 1996). In *Stegodyphus*, in contrast, population genetics studies are more advanced (Johannesen *et al.* 2002, 2007, 2009).

Our goal here is to understand the population genetics consequences of sociality in *Anelosimus eximius*, one of the best known social theridiid spiders. Previous studies on allozymes in *A. eximius* (Smith & Engel 1994; Smith & Hagen 1996) point to population subdivision, but they lack resolution because of the relatively low genetic variability of allozyme markers. Here we examine the patterns of genetic diversity in *A. eximius*, in part using a hypervariable mitochondrial DNA region, nested within the 16S ribosomal RNA loci in our study species *A. eximius*, whose discovery we report here. We test two primary hypotheses regarding female dispersal and nest fidelity: (i) that colonies primarily contain single matrilineages, thus showing that population subdivision has led to the structuring of genetic variability mostly among, rather than within colonies and, (ii) that relationships among haplotypes reflect the geographical distribution of colonies, as would be the case if high rates of colony turnover led to single haplotypes sweeping through local areas by either group-level drift or selection.

Methods

We sampled 204 individuals from 38 *A. eximius* colonies (Table 1) in Ecuador (2003–2005) and one colony from near Cayenne, French Guiana (2005). Colonies were visually searched for along transects in each area. Sampled colonies ranged in size from approximately 50 to 2500 adult females. To examine the correspondence of haplotype distribution within and among colonies with their geographical locality we sampled colonies at various geographical scales. Colonies were sampled along an approximately 30 km segment of the Cuyabeno River (Sucumbios region, 0.021–0.097S, 76.133–76.337W, 2003 and 2005) and an approximately 20 km transect along the road Jondachi–Loreto (Ecuador, Napo region, 0.7103–0.727S, 77.585–77.757W, 2005), and in Jatun Sacha (Napo region, approx. 1.067S, 77.617W, 2004), all on the E slope of the Andes in Northern Ecuador (Fig. 1). Colonies were also sampled from southern Ecuador (Morona Santiago, Limón, approximately 2.9833S, 78.3667W, 2003) and from the W slope of the Andes (Pedro Vincente, 0.08–0.10N 79.03–79.07). To serve as outgroups for phylogenetic analyses 10 specimens of the related social species *A. domingo* were sampled in Ecuador, from Jatun Sacha (colonies JS04JSD-1, -2, and -3), and Pedro Vincente (colonies PV04X-1 and -2). Distances between colonies ranged from a few meters (colonies within the same colony cluster) to 3000 km (southwestern Ecuador to French Guiana). Figure 1 shows the distribution of sampling areas in Ecuador. Voucher specimens are deposited at the University of British Columbia. GenBank accession numbers for the specimens are: FJ743932–FJ744066.

Specimens were placed in 95% ethanol in the field and upon return to the laboratory, kept frozen at −80 °C until DNA extraction. For details of DNA extraction and PCR protocol see Agnarsson *et al.* (2007). PCR products were purified and sequenced by the Macrogen Inc. (ABI 3730) sequencer and proofread using the Chromaseq module (Maddison & Maddison, unpublished) in Mesquite (Maddison & Maddison 2009) (further details in Agnarsson *et al.* 2007). We sequenced from 3 to 12 individuals per colony, although in some cases only a single individual was successfully sequenced. We sequenced a single 804-bp long (aligned) mitochondrial fragment obtaining partial sequences of the 16S and ND1 genes. This region is documented to show intraspecific variability in spiders (Hedin 1997; Agnarsson *et al.* 2007; Johannesen *et al.* 2007). Sequences were aligned using ClustalX (Thompson *et al.* 1997) with gap opening/gap extension costs set to 24/6 following prior work (Maddison & Hedin 2003). Two other alignments were made using MacClade (Maddison & Maddison 2005) by (i) adjusting the clustal alignment in obviously misaligned regions mostly near the sequence ends, and manually aligning the 'microsatellite' region (Fig. 3); and (ii) further modifying the latter alignment by minimizing the number of gaps in the microsatellite region, under the assumption that it mostly contains terminal repeats.

Secondary structure of the 16S RNA sequence was implied using the RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi), to structurally locate the microsatellite region.

Our primary purpose here is to address female movement and matrilineage separation. However, in order to get at male movement we explored two nuclear loci, 28S and ITS2. To scan for potentially informative variation, sequences of 26 individuals were obtained for 28S, and of 10 individuals for ITS2, in both cases sampling specimens that span the maximum observed mitochondrial variability. However, the nuclear data had almost no variability. 28S showed zero genetic variability among the sampled specimens. For ITS2 only two base

Table 1 List of haplotypes and geographical location (from east to west) of colonies of *A. Eximius*. Coloured boxes identify phylogenetic lineages, matching colors in Figure 2. N = the number of spiders sampled per colony (F = females, M = males, juv(s) = juvenile(s))

Area	Colony	N (F/M)	H1	H13	H14	H5	H6	H9	H10	H11	H12	H15	H17	H18	H19	H22	H23	H21	H3	H25	H4	H8	H2	H24	H16	H20	H7
Pedro Vincente	PV_04_X_1	**5 (3/2 juv)**																								5	
Pedro Vincente	PV_04_X_2	**4 (3/1 juv)**																								4	
Pedro Vincente	PV_04_X_3	**4 (3/1 juv)**																								4	
Limon	LI_04_X_PS1	**10 (6/3/1 juv)**										10															
Limon	LI_04_X_PS2	**1 (1/0)**										1															
Limon	LI_04_X_PS3	**1 (1/0)**										1															
Limon	LI_04_X_PS4	**1 (1/0)**										1															
Limon	LI_04_X_PS5	**5 (2/1/2 juv)**																						2	3		
Loreto	LO_05_28.5_0	**12 (10/2)**														5	7										
Loreto	LO_05_27.0_2	**9 (7/2)**											9														
Loreto	LO_O5_3.3_0	**5 (4/1)**								5																	
Loreto	LO_O5_32.7_1	**1 (1/0)**												1													
Loreto	LO_O5_32.7_2	**4 (3/1)**												4													
Loreto	LO_O5_45.1_1	**5 (5/0)**													5												
Loreto	LO_O5_45.1_3	**4 (3/1)**													4												
Jatun Sacha	JS_03_X_X	**1 (1/0)**			1																						
Jatun Sacha	JS_03_JSE8_0	**7 (7/0)**			7																						
Jatun Sacha	JS_03_JSE3_0	**3 (3/0)**								3																	
Jatun Sacha	JS_03_JSE2_2	**4 (4/0)**								4																	
Jatun Sacha	JS_03_JSE4_0	**5 (5/0)**								5																	
Jatun Sacha	JS_03_JSE6_0	**4 (juvs)**																4									
Cuyabeno	CU_O5_R14_1	**5 (3/2)**																				5					
Cuyabeno	CU_O5_R14_2	**5 (1/4)**																				5					
Cuyabeno	CU_03_b4_1 _extrano	**1 (1/0)**																									1
Cuyabeno	CU_03_23_1	**8 (7/1)**																					8				
Cuyabeno	CU_03_23_2	**4 (3/1)**																					4				
Cuyabeno	CU_03_5_1	**3 (3/0)**					3																				
Cuyabeno	CU_03_5_2	**3 (1/2)**					3																				
Cuyabeno	CU_03_14_1	**8 (6/2)**	8																								
Cuyabeno	CU_03_14_2	**4 (2/2)**	4																								
Cuyabeno	CU_03_28_1	**9 (7/2)**																	6	3							
Cuyabeno	CU_03_28_2	**6 (6/0)**									2										4						
Cuyabeno	CU_03_29_1	**4 (2/2)**				4																					
Cuyabeno	CU_03_35_1	**9 (7/2)**				9																					
Cuyabeno	CU_03_35_2	**5 (5/0)**				5																					
Cuyabeno	CU_05_R6_0	**9 (8/1)**																			9						
Cuyabeno	CU_O5_R10_1	**9 (7/2)**		1				2	6																		
Cuyabeno	CU_O5_R10_2	**10 (10/0)**						2	8																		
French Guiana	FG_05_X_1	**4 (2/2)**																									4

Fig. 1 Geographic distribution of sampling areas in Ecuador.

positions showed parsimony-informative variation, and those two were in conflict, indicating homoplasy. These loci thus are not informative at the population level in *A. eximius*, and further sampling of these nuclear loci was thus abandoned.

Phylogenetic reconstruction of sequences and haplotypes was performed in PAUP 4.08 (Swofford 1999) using parsimony with equal weights, either excluding indels, or treating gaps as a fifth state. Analyses were run with 1000 constrained random addition sequence searches using TBR branch rearrangements; the trees resulting were then used as starting points for an unconstrained TBR search limited only by MAXTREES of 100 000.

Haplotype networks were constructed using the CombineTrees software (Cassens *et al.* 2005) utilizing all the most parsimonious trees from the phylogenetic analysis, and in a statistical parsimony framework using TCS (Clement *et al.* 2000).

Population genetic structure was calculated with an AMOVA in Arlequin 3.1.1 (Excoffier *et al.* 2005). We estimated the distribution of variation within and among populations, and calculated fixation indices.

Geographic and genetic distance matrices were generated in Mesquite, and correlation between geographic and genetic distances between colonies was tested with the MANTEL statistic, using Mantel Nonparametric Test Calculator Version 2.0 (Liedloff 1999). Analyses were run both with the full dataset, as well as a pruned dataset. The data were pruned (and restricted to Ecuador) to test if correlation between geographic and genetic distances detected in the full dataset could be explained exclusively by the geographic clustering of identical colonies (colonies sharing the same haplotypes) that are expected to be the progeny of the same mother colony. By representing identical colonies within regions only once, we could test if different haplotypes within regions tend to be more similar than when comparing haplotypes among regions.

The program Estimates 8.2.0 (Colwell 2006) was used to estimate expected total number of haplotypes in the entire sampling, as well as within individual colonies, using the Chao1 statistic.

Results

Haplotypes

Among the 204 individuals sequenced from 39 colonies, we identified 25 haplotypes (Table 1), whose phylogenetic relationships are shown in Fig. 2. Changes within a single hyper-variable region (51 bp aligned) within the 16S gene separated many of the haplotypes (the aligned data are available as a Nexus file from the first author, and as Appendix S1, Supporting Information). This region spans bases 395–446 in the aligned matrix (Fig. 3, matrix available online and from the authors) and is characterized mostly by microsatellite-like repeats of the motif 'TA' (really UA, as these are RNA

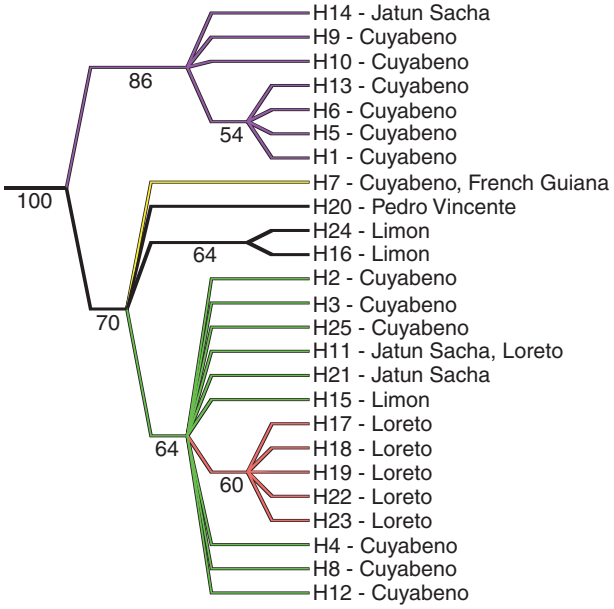

Fig. 2 Phylogenetic relationships among the 25 haplotypes (H), and their geographic distribution. Colours mark phylogenetic lineages, matching the colour scheme used to represent colonies sampled from Cuyabeno (see Fig. 3), additionally the Loreto lineage is marked in red so that all lineages correspond to colouring of Table 1. Numbers on nodes are bootstrap support values. It is notable that there is virtually no character conflict in the data (consistency index = 0.99, retention index = 1), suggesting that relatively low bootstrap values are due to few data, not conflicting hypotheses.

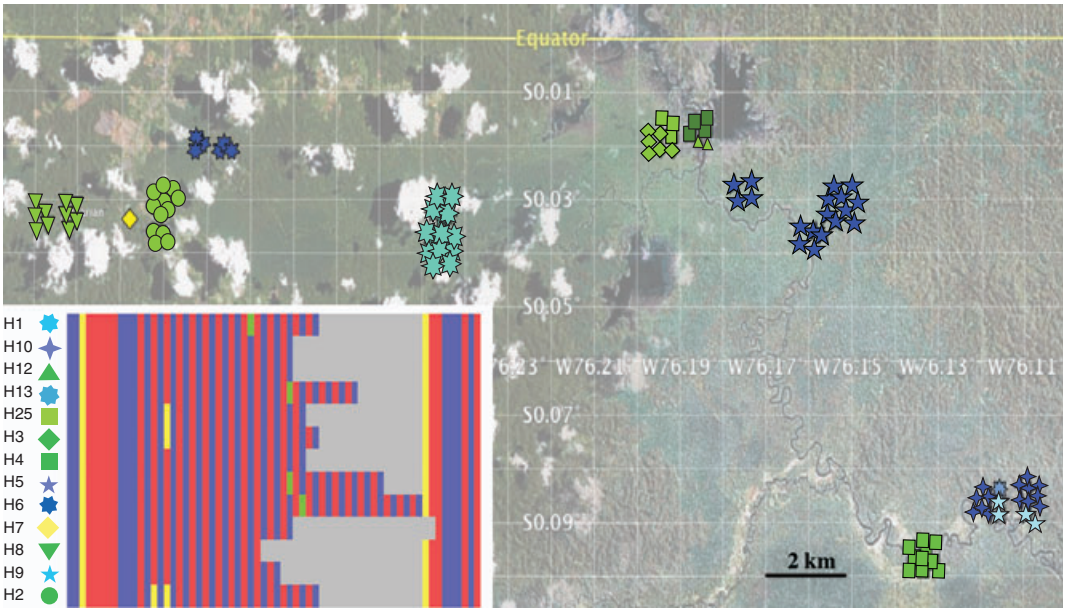

Fig. 3 Distribution of haplotypes among colonies along a 30 km transect in Cuyabeno, Ecuador. Haplotype colouring reflects phylogenetic lineages as in Fig. 2, with different colour shades of blue or green used to help separate colonies within the same phylogenetic lineage. The phylogenetic lineages in the area are clearly shuffled, without any evidence of mixing. The legend shows part of the 16S sequence of each haplotype including the microsatellite-like region of 16S for the 13 haplotypes found in Cuyabeno (note that H10 and H12 are identical in this region of the sequence, but differ by point mutations in other part of the entire sequence). Colours are standard DNA base colours A = red, T = blue, G = yellow, C = green. Grey indicates indels that accommodate for difference in sequence lengths. Each cluster of overlapping symbols represents a single colony. Similar symbols demark colonies from the same colony cluster (same locality data) that have been separated on the map for visual clarity.

sequences) that differs dramatically among individuals (Fig. 3, Supporting Information). Secondary structure reconstruction (Appendix S3, Supporting Information) suggests that the TA repeats form a portion of a stem, where T and A match up. Hence, individuals differ in the length of this stem in the 16S ribosome. This would explain why the sequences always differ by a TA repeat, rather than a single base pair as might be expected if this were a loop region. This region of TA repeats is much shorter and varies little in length in *A. domingo* and other *Anelosimus* species (Agnarsson *et al.* 2007, in preparation), hence it appears to be hypervariable only in *A. eximius*.

Of the 25 haplotypes one was sampled only in a single individual, the others were sampled twice or more. Chao1 estimates 25.13 haplotypes in the system with the 95% upper bound confidence interval at 25.18. Hence, we do not expect numerous unsampled haplotypes to occur in the sampled colonies. Likewise, in the vast majority of cases, within colonies and colony complexes sharing haplotypes, the expected number of haplotypes is the observed (see Appendix S2, Supporting Information), using the Chao 1 95% upper bound CI. Nevertheless, we expect that within some of the most poorly sampled colonies we have missed haplotypes. Our prediction would be that missing haplotypes represent single-mutational variation of the more common colony haplotype, as observed in several better sampled colonies.

The majority of colonies contained only a single haplotype in the sampled specimens (Table 1). In such cases each colony typically either had a unique haplotype (e.g. JS-03-JSE6-0, LO-05-28.5-0) or shared a haplotype with a nearby colony in the same, or a nearby colony cluster (e.g. CU-05-R14-1 and R14-2, CU-03-23-1 and 23-2). In some cases haplotypes were shared among more distant colonies, either in geographic proximity (e.g. JS-03-JSE2-2 from Jatun Sacha and LO-05-3.3-0 from Loreto approximately 42 km apart) or distant (CU-03-b4-1 from Cuyabeno Ecuador and the colony from French Guiana, approximately 3000 km apart). Approximately 18% of the colonies contained two haplotypes; one colony contained three haplotypes. In all but one case the second, rarer, haplotype differed from the other haplotype found within the colony by a single two base pair long 'repeat' in the microsatellite-like region. In the colony that had three haplotypes, the third haplotype differed from the other two by an eight adjacent bases-putatively a single four 'repeats' long insertion.

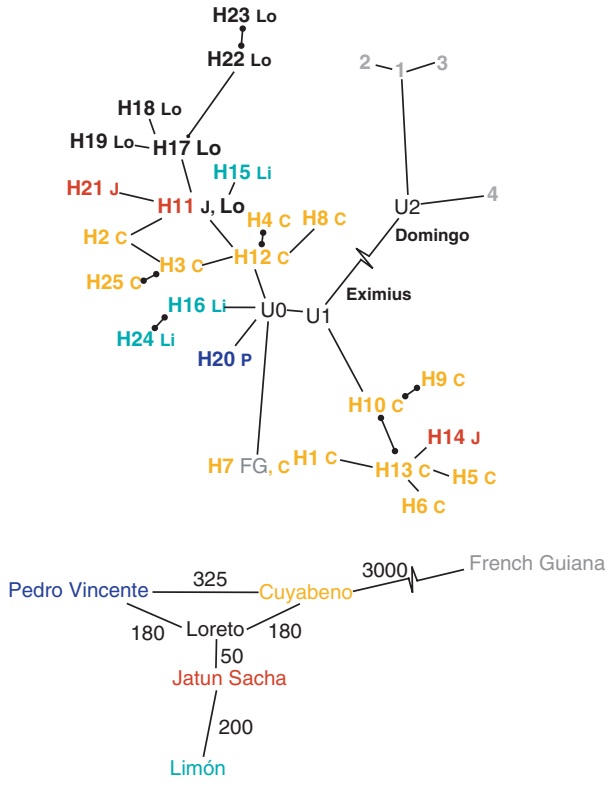

Fig. 4 A haplotype network constructed using CombineTrees. Note that colours here refer to geographic locations, not phylogenetic lineages. Haplotypes are numbered and locality is indicated with colour and letter following the haplotype number (C, Cuyabeno; PV, Pedro Vincente; Lo, Loreto; J, Jatun Sacha; Li, Limón). Branch lengths correspond to distances between haplotypes (distance between *A. domingo* and *A. eximius* is not to scale). Different haplotypes found in the same colony are connected by a line with circles on each end. Diagram below shows approximate relative geographic localities, and distances between, sample sites (distance to F. Guiana, not to scale) Numbers indicate kilometres.

In two cases colonies with more than one haplotype shared one, but not the other(s) with a neighbouring colony (Fig. 3). In each case the haplotypes involved were only one 'repeat' motif different from each other.

Alignment and haplotype phylogeny and network

The majority of the indel events, or gaps, were implicated in the microsatellite-like region that is seemingly characterized by repeated insertions and/or deletions of mostly 'TA' (motif repeats Fig. 3, Supporting Information). Therefore, it is near impossible to detect which repeats arose independently and which may be homologous across taxa, this region will be particularly prone to homoplasy under any given alignment scheme. Further, treating gaps as characters will increase the weight of this homoplastic region. We therefore preferred analyses that exclude indels, favouring information from unambiguously aligned regions. In support of this preference, phylogenetic analyses that used gaps as a fifth state (treating each two base pair 'gap' as a single event) resulted in gaps being in conflict with unambiguously aligned sequence variation. Furthermore, removing this region altogether yielded nearly identical results as analyses ignoring gaps.

Parsimony analyses of the sequence data, ignoring gaps, resulted in six equally most parsimonious trees (70 steps; consistency index 0.986; retention index 0.995), the strict consensus of which is shown in Fig. 2 (collapsing groups with less than 50% Bootstrap support). The alternative alignments gave near identical results (not shown, matrices available from first author). The most parsimonious trees were translated into a haplotype network (Fig. 4). The phylogeny and haplotype network poorly reflect geographic localities of colonies (Figs 3 and 4). While some areas more or less cluster (e.g. Loreto), consistent with the presence of a single matrilineage in the area and haplotype diversification occurring among related colonies, other areas, such as Jatun Sacha and Cuyabeno, contain two or more phylogenetically distinct haplotypes. Haplotypes of different phylogenetic lineages never co-occurred in any single colony. The haplotype network obtained with the TCS method broadly agreed with the CombineTrees network, although more difficult to interpret because of numerous reticulations. Nevertheless, the conclusion from the TCS analysis is the same, the haplotype network poorly reflects geographic localities of colonies.

The AMOVA confirms strong population structure, with less than 5% of genetic variation found within colonies (% var. = 4.25, sum of squares = 37.7, d.f. = 166) and very high fixation index (F_{ST} = 0.96, P < 0.0001).

Geographic and genetic distances were correlated in the full dataset (G = 4.87, Z = 39608.8, r = 0.239, P = 0.01). However that correlation was due exclusively to the clustering of identical colonies within regions. When considering each haplotype within a region only once, there is no correlation between geographic and genetic distances of the different haplotypes found within and among regions (G = 0.936, Z = 332.8, r = −0.01, P = 0.27).

Nuclear data

The small nuclear datasets collected (available from author upon request) showed nearly no intraspecific genetic variation, even across specimens with high mitochondrial sequence divergence. These nuclear loci have insufficient variation to resolve population level

structure such that further sampling was not deemed fruitful.

Discussion

Behaviour, such as sociality and breeding systems can profoundly affect the structure of populations and metapopulations, and, in turn, alter the distribution and quantity of genetic diversity (Ross 2001; Ansell *et al.* 2008; Barrett *et al.* 2008; Bay *et al.* 2008; Hoelzel *et al.* 2008; Kamiya *et al.* 2008; Ortega *et al.* 2008). As sociality often leads to reduced dispersal of at least one of the sexes, it reduces random mating within and among populations. Our results broadly confirm a highly subdivided matrilineage metapopulation structure of the social spider *A. eximius* that leads to repackaging of genetic variability, thereby corroborating the hypothesis that colonies contain single matrilines (Johannesen *et al.* 2009). Thus sociality has led to the structuring, at least of female inherited mitochondrial genetic variability mostly among, rather than within colonies (hypothesis 1). Much of the observed haplotype variation, and hence our ability to infer population history, rests within the newly discovered microsatellite-like region in *A. eximius*, where other *Anelosimus* species tend to be quite uniform (Agnarsson *et al.* 2007). This region seems to evolve very rapidly, and is characterized by insertions (and probably also deletions) of repeats of a 'TA' motif, although other point mutations clearly also occur in this area (Fig. 3).

We found strong population subdivision (F_{ST} = 0.96), with extremely low within-colony variability—typically a single haplotype—but considerable between-colony differentiation. The little within-colony variability found typically consisted of a second rarer haplotype differing from the more common one by a single 'TA' motif repeat in the microsatellite-like region. Such haplotypes most likely arose by single mutations within the colony lineage. As indicated by the Chao 1 statistic, undersampling of colonies and haplotypes is unlikely to be an important confounding factor. The Chao 1 statistic indicates that in the system as a whole and in the vast majority of cases within colonies, sampling was adequate, with the estimated number of haplotypes equalling the observed. We can still expect to have missed haplotypes, especially in those colonies sampled by four or less individuals. However, none of the polymorphic colonies that were relatively well sampled showed a pattern suggestive of intercolony mixing. We thus suspect that even if additional haplotypes were to be uncovered in the less extensively sampled colonies, it is unlikely that they will exhibit a pattern dramatically different from that seen in the better sampled ones. Our data, therefore, suggest colony propagation by single matrilineages and an absence of colony mixing, at least involving female migration, as also seen in the distantly related inbred-social *Stegodyphus* (Johannesen *et al.* 2002, 2007, 2009; Bilde *et al.* 2005). Colonies within the same nest cluster, or up to a few tens of meters away, typically shared the same haplotype (Fig. 3; Table 1), suggesting they were derived from the same ancestral colony via budding or short distance dispersal.

In two interesting cases colonies much farther apart shared the same unique haplotype. Haplotypes were shared between a colony from Jatun Sacha (JS-03-JSE2-2) and one from Loreto (LO-05-3.3-0), localities found approximately 42 km apart, and between a colony from Cuyabeno, Ecuador (CU-03-b4-1) and the colony from French Guiana, found 3000 km apart. These cases are suggestive of either a widespread lineage occupying, in the past or present, areas between the two localities or of dispersal over relatively large distances. We consider the latter possibility less likely, however, as nearly all observations or inferences of dispersal in this species are by colony budding or walking of one or a number of individuals over short distances. Note that although homoplasy cannot be ruled out in either case, it seems much less likely than identity by descent. In the case of the Jatun Sacha and Loreto colonies homoplasy would have involved converging on identical sequences of over 800 bp, an unlikely possibility unless the sequences were very recently identical by descent, in which case relatedness is indicated anyway. The argument against homoplasy is even stronger for the Cuyabeno and French Guiana colonies. The unique haplotype that these colonies shared differed from all others not only in the microsatellite region (where convergence is likely), but also in a number of point mutations along the entire 804-bp long sequence. Furthermore, individuals from both of these colonies amplified at atypically low success rates, suggesting that this haplotype differs from all others in the region where the primers bind. The single specimen from the Cuyabeno colony that we were able to amplify was processed prior to the collection of the colony in French Guiana, ruling out mislabelling or contamination. These findings, along with the observation that the Ecuadorian specimens of this haplotype looked rather different from those typical in the area (adult females were smaller in size and more brightly coloured than typical *A. eximius*, L. Aviles field observation) suggest that the French Guianan and Cuyabeno CU-03-b4-1 specimens belong to a lineage that has been separate from the others here studied long enough to have accumulated phenotypic differences. Nevertheless, there is zero genetic variability between these mitochondrial lineages in the nuclear 28S sequences. Sampling intervening areas will cast further light on the occurence of this haplotype in Ecuador.

Typically, however, colonies within a region were not homogenous in the haplotypes they contained. In addition to haplotypes belonging to the same phylogenetic lineage, and thus probably derived from a common matriline, some regions contained phylogenetically distinct haplotypes even for colonies found in relatively close geographical proximity (Fig. 3 and Table 1). Hence, apart from immediate colony neighbours, the haplotype phylogeny (Fig. 2) and the haplotype network (Fig. 4) poorly reflect geographical localities of, and distance between, colonies. Similarly, the genetic distance between these distinct haplotypes, and distance between the colonies that contain them were not correlated. Both observations lead to rejection of hypothesis 2. For example, Cuyabeno contains at least two distinct haplotype lineages that seem not closely related (Figs 2 and 3). These two lineages are shuffled at a fine geographic scale in this area while no single colony contains individuals representing both lineages (Fig. 3). This pattern offers clear evidence that even in areas with more than one matriline, they rarely, if ever, mix. Rather, areas seem to have been colonized independently by different matrilineages that both have become widespread and interspersed in the region.

Haplotypes from one of the areas (Loreto), however, did approximately cluster, consistent with the presence of a single matrilineage in this area and haplotype diversification occurring between related colonies. Thus, although the Loreto area was not strictly monophyletic as it shared a haplotype with a colony in Jatun Sacha, the shared haplotype could have belonged to the same colony lineage as other colonies in Loreto, or possibly even been ancestral to them, as implied by the haplotype network (Fig. 4). Overall, however, the strong population subdivision and structuring evidenced by our haplotype data indicate that area-wide 'sweeps'—single genotypes rapidly spreading through entire areas (by either drift or selection) and replacing other lineages—have not occurred recently in the areas we sampled most densely, as sweeps would have rendered entire areas mostly, or entirely, monomorphic.

Taken together the high correspondence between individual colony membership and haplotype, the presence of more than one haplotype lineage in some areas, lack of correlation between geographic and genetic distances of distinct haplotypes, and high fixation index (F_{ST} = 0.96), all indicate a highly subdivided metapopulation structure where female colony members remain 'faithful' to their matriline even when invading novel areas, but where no single lineage comes to replace all others, even within relatively small local areas. The observation that the most densely sampled areas contained haplotypes of diverse phylogenetic origin (Fig. 3) suggests that area-wide sweeps that would cause some lineages to replace all others in an area have not occurred recently in the areas sampled.

The extent to which female faithfulness translates to mating between close relatives within the colony lineages will depend on the extent to which males also remain within their natal nests to mate (e.g. Bilde *et al.* 2005). Being inherited through the maternal line, our mitochondrial data do not allow us to make inferences about the history of gene flow among colonies because of male migration, but can be used to infer recent male migration events. In our case, 25% of our mitochondrial sequences belonged to males. In all cases male haplotypes were congruent with female haplotypes in their respective colonies, suggesting absence of recent male migration. We are in the process of confirming this inference using nuclear microsatellite markers (L. Aviles, unpublished). Sequence data of two nuclear markers we obtained (28S, ITS2) were unfortunately uninformative for this level of analysis. The near complete lack of sequence variation observed at these nuclear markers, however, when contrasted with the relatively normal levels of sequences divergence found in the same markers in related outbred subsocial species (Agnarsson *et al.* unpublished), is consistent with a pattern of strong population subdivision and lineage turnover in *A. eximius*. Further evidence of lack of male movement between colonies comes from allozyme data (Smith & Hagen 1996), preliminary analyses of microsatellite data (Aviles *et al.* unpublished), and from the strongly female-biased sex ratios characteristic of this species (Avilés & Maddison 1991; Aviles 1993).

Overall, the consequences of inbred sociality in spiders can best be described by considering various time scales. In the short term, the faithful matrilineal descent of social spider colonies generates a highly structured population with a severe reduction in diversity within colony lineages (Smith & Hagen 1996; Johannesen *et al.* 2009). Over the medium term, however, the species survives and retains considerable genetic diversity both within a region and across its range: broad selective sweeps are not implicated. Phylogenetic studies, however, suggest that medium term success does not translate to long term evolutionary potential, as social spider species lineages apparently either fail to speciate or fail to persist (Agnarsson *et al.* 2006; Johannesen *et al.* 2007).

Acknowledgements

We are indebted to the Museo Ecuatoriano de Ciencias Naturales for sponsoring our research in Ecuador and the Instituto Ecuatoriano de Areas Naturales y Vida Silvestre (Ecuador) for collecting permits. We are grateful to Patricio Salazar, Gabriel

Iturralde, and Jessica Purcell for help collecting specimens, and to Carol Ritland, Karen Needham, and Laura May-Collado for help with the molecular work. This research was funded by grants from the National Sciences and Engineering Research Council of Canada to L.A and W.P.M, and a Killam postdoctoral fellowship and a Slovenian Research Agency grant (Z1-9799-0618-07) to I.A. We are grateful to eight anonymous reviewers and Dr Brent Emerson for comments that improved the manuscript.

References

Abbot P (2009) On the evolution of dispersal and altruism in aphids. *Evolution*, **63**, 2687–2696.

Agnarsson I (2003) The phylogenetic placement and circumscription of the genus *Synotaxus* (Araneae: Synotaxidae), a new species from Guyana, and notes on theridioid phylogeny. *Invertebrate Systematics*, **17**, 719–734.

Agnarsson I (2004) Morphological phylogeny of cobweb spiders and their relatives (Araneae, Araneoidea, Theridiidae). *Zoological Journal of the Linnean Society*, **141**, 447–626.

Agnarsson I (2006) A revision of the New World *eximius* lineage of *Anelosimus* (Araneae, Theridiidae) and a phylogenetic analysis using worldwide exemplars. *Zoological Journal of the Linnean Society*, **146**, 453–593.

Agnarsson I, Aviles L, Coddington JA, Maddison WP (2006) Sociality in Theridiid spiders: repeated origins of an evolutionary dead end. *Evolution*, **60**, 2342–2351.

Agnarsson I, Maddison WP, Aviles L (2007) The phylogeny of the social *Anelosimus* spiders (Araneae : Theridiidae) inferred from six molecular loci and morphology. *Molecular Phylogenetics and Evolution*, **43**, 833–851.

Ansell SW, Grundmann M, Russell SJ, Schneider H, Vogel JC (2008) Genetic discontinuity, breeding-system change and population history of *Arabis alpina* in the Italian Peninsula and adjacent Alps. *Molecular Ecology*, **17**, 2245–2257.

Arnedo MA, Coddington J, Agnarsson I, Gillespie RG (2004) From a comb to a tree: phylogenetic relationships of the comb-footed spiders (Araneae, Theridiidae) inferred from nuclear and mitochondrial genes. *Molecular Phylogenetics and Evolution*, **31**, 225–245.

Arnedo MA, Agnarsson I, Gillespie RG (2007) Molecular insights into the phylogenetic structure of the spider genus *Theridion* (Araneae, Theridiidae) and the origin of the Hawaiian *Theridion*-like fauna. *Zoologica Scripta*, **36**, 337–352.

Aviles L (1993) Interdemic selection and the sex-ratio—a social spider perspective. *American Naturalist*, **142**, 320–345.

Aviles L (1997) Causes and consequences of cooperation and permanent-sociality in spiders. In:*The Evolution of Social Insects and Arachnid* (eds Choe JC, Crespi BJ). pp. 476–498, Cambridge University Press, Cambridge.

Aviles L (1999) Cooperation and non-linear dynamics: an ecological perspective on the evolution of sociality. *Evolutionary Ecology Research*, **1**, 459–477.

Aviles L (2000) Nomadic behaviour and colony fission in a cooperative spider: life history evolution at the level of the colony? *Biological Journal of the Linnean Society*, **70**, 325–339.

Avilés L, Maddison W (1991) When is the sex ratio biased in social spiders? Chromosome studies of embryos and male meiosis in Anelosimus species (Araneae, Theridiidae). *Journal of Arachnology*, **19**, 126–135.

Aviles L, Maddison WP, Agnarsson I (2006) A new independently derived social spider with explosive colony proliferation and a female size dimorphism. *Biotropica*, **38**, 743–753.

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.

Bay LK, Caley MJM, Crozier RH (2008) Meta-population structure in a coral reef fish demonstrated by genetic data on patterns of migration, extinction and re-colonisation. *BMC Evolutionary Biology*, **8**, 248.

Bilde T, Lubin Y, Smith D, Schneider JM, Maklakov A (2005) The transition to social inbred mating systems in spiders—role of inbreeding tolerance in a subsocial predecessor. *Evolution*, **59**, 160–174.

Cassens I, Mardulyn P, Milinkovitch MC (2005) Evaluating intraspecific ''network'' construction methods using simulated sequence data: do existing algorithms outperform the global maximum parsimony approach? *Systematic Biology*, **54**, 363–372.

Chapman TW, Crespi BJ, Kranz BD, Schwarz MP (2000) High relatedness and inbreeding at the origin of eusociality in gall-inducing thrips. *Proceedings of the National Academy of Sciences, USA*, **97**, 1648–1650.

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

Cockburn A (1998) Evolution of helping behavior in cooperatively breeding birds. *Annual Review of Ecology and Systematics*, **29**, 141–177.

Colwell RK (2006) EstimateS: statistical estimation of species richness and shared species from samples. Version 8. Available from: http://purl.oclc.org/estimates.

Crouch T, Lubin Y (2001) Population stability and extinction in a social spider *Stegodyphus mimosarum* (Araneae: Eresidae). *Biological Journal of the Linnean Society*, **72**, 409–417.

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

Hales DF, Tomiuk J, Wohrmann K *et al*. (1997) Evolutionary and genetic aspects of aphid biology: a review. *European Journal of Entomology*, **94**, 1–55.

Hamilton WD (1964) The genetical evolution of social behavior. I & II. *Journal of Theoretical Biology*, **7**, 1–52.

Hedin MC (1997) Molecular phylogenetics at the population/species interface in cave spiders of the southern Appalachians (Araneae: Nesticidae: Nesticus). *Molecular Biology and Evolution*, **14**, 309–324.

Hoelzel AR, Hey J, Dahlheim ME, Nicholson C, Burkanov V, Black N (2008) Evolution of population structure in a highly social top predator, the killer whale. *Molecular Biology and Evolution*, **24**, 1407–1415.

Hughes C (1998) Integrating molecular techniques with field methods in studies of social behavior: a revolution results. *Ecology*, **79**, 383–399.

Jarvis JUM, Oriain MJ, Bennet NC, Sherman PW (1994) Mammalian eusociality—a family affair. *Trends in Ecology and Evolution*, **9**, 47–51.

Johannesen J, Lubin Y (1999) Group founding and breeding structure in the subsocial spider *Stegodyphus lineatus* (Eresidae). *Heredity*, **82**, 677–686.

Johannesen J, Lubin Y (2001) Evidence for kin-structured group founding and limited juvenile dispersal in the subsocial spider *Stegodyphus lineatus* (Araneae, Eresidae). *Journal of Arachnology*, **29**, 413–422.

Johannesen J, Hennig A, Dommermuth B, Schneider JM (2002) Mitochondrial DNA distributions indicate colony propagation by single matri-lineages in the social spider *Stegodyphus dumicola* (Eresidae). *Biological Journal of the Linnean Society*, **76**, 591–600.

Johannesen J, Lubin Y, Smith DR, Bilde T, Schneider JM (2007) The age and evolution of sociality in *Stegodyphus* spiders: a molecular phylogenetic perspective. *Proceedings of the Royal Society B-Biological Sciences*, **274**, 231–237.

Johannesen J, Wickler W, Seibt U, Moritz RFA (2009) Population history in social spiders repeated: colony structure and lineage evolution in *Stegodyphus mimosarum* (Eresidae). *Molecular Ecology*, **18**, 2812–2818.

Kamiya K, Moritsuka E, Yoshida T, Yahara T, Tachida H (2008) High population differentiation and unusual haplotype structure in a shade-intolerant pioneer tree species, *Zanthoxylum ailanthoides* (Rutaceae) revealed by analysis of DNA polymorphism at four nuclear loci. *Molecular Ecology*, **17**, 2329–2338.

Lepais P, Darvill B, O'Connor S *et al.* (2010) Estimation of bumblebee queen dispersal distances using sibship reconstruction method. *Molecular Ecology*, **29**, 819–831.

Liedloff AC (1999) Mantel Nonparametric Test Calculator. Version 2.0. School of Natural Resource Sciences, Queensland University of Technology, Australia.

Lubin Y, Bilde T (2007) The evolution of sociality in spiders. In: *Advances in the Study of Behavior* (ed. Brockman HJ, Roper I, Naguib M, Wynne-Edwards K, Barnard C, Mitani J), Vol. 37, pp. 83–145. Elsevier Academic Press Inc., San Diego.

Lubin YD, Crozier RH (1985) Electrophoretic evidence for population differentation in a social spider *Achaearanea wau* (Theridiidae). *Insectes Sociaux*, **32**, 297–304.

Lubin YD, Robinson MH (1982) Dispersal by swarming in a social spider. *Science*, **216**, 319–321.

Maddison WP, Hedin MC (2003) Jumping spider phylogeny (Araneae: Salticidae). *Invertebrate Systematics*, **17**, 529–549.

Maddison DR, Maddison WP (2005) MacClade 4.07. Sinauer Associates, Sunderland, Massachusetts.

Maddison WP, Maddison DR (2009) Mesquite: a modular system for evolutionary analysis. Version 2.5 Available from: http://mesquiteproject.org.

Maynard Smith J, Szathmáry E (1995) *The Major Transitions in Evolution*. WH Freeman, Oxford.

Ortega J, Aparicio JM, Cordero PJ, Calabuig G (2008) Individual genetic diversity correlates with the size and spatial isolation of natal colonies in a bird metapopulation. *Proceedings of the Royal Society B-Biological Sciences*, **275**, 2039–2047.

Pusey A, Wolf M (1996) Inbreeding avoidance in animals. *Trends in Ecology and Evolution*, **11**, 201–206.

Riechert SE, Roeloffs RM (1993) *Evidence for and Consequences of Inbreeding in the Cooperative Spiders*. University of Chicago Press, Illinois.

Roeloffs R, Riechert SE (1988) Disperal and population-genetic structure of the cooperative spider, *Agelena consociata*, in West-African rainforest. *Evolution*, **42**, 173–183.

Ross KG (2001) Molecular ecology of social behavior: analysis of breeding systems and genetic structure. *Molecular Ecology*, **10**, 265–284.

Shellman-Reeve JS (1997) Advantages of biparental care in the wood-dwelling termite, *Zootermopsis nevadensis*. *Animal Behaviour*, **54**, 163–170.

Smith DR, Engel MS (1994) Population-structure in an Indian cooperative spider, *Stegodyphus sarasinorum* Karsch (Eresidae). *Journal of Arachnology*, **22**, 108–113.

Smith DR, Hagen RH (1996) Population structure and interdemic selection in the cooperative spider *Anelosimus eximius*. *Journal of Evolutionary Biology*, **9**, 589–608.

Storz JF (1999) Genetic consequences of mammalian social structure. *Journal of Mammalogy*, **80**, 553–569.

Swofford DL (1999) PAUP*. Phylogenetic Analysis Using Parsimony (and Other Methods). Sinauer Associates, Sunderland, Massachusetts.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876–4882.

Wilson DS (1975) A theory of group selection. *Proceedings of the National Academy of Sciences, USA*, **72**, 143–146.

Wilson EO (1987) Causes of ecological success: the case of the ants. *Journal of Animal Ecology*, **56**, 1–9.

I. A. studies a variety of topics including systematic, morphology, genetics, and behaviour of spiders, web evolution and biomechanical properties of spider silks, and phylogeny and evolution of mammals. W. P. M. studies spider phylogeny and evolution, theoretical approaches to understanding evolutionary history, and computational methods to implement them. L. A. is an evolutionary ecologist interested in understanding the forces that bring together individuals into social groups and the short and long term consequences of social evolution. Current areas of research include the ecology and biogeography of sociality, the interplay between kinship and demography in shaping the diversity of animal societies, the role of selection at multiple levels in evolution, and sociality and community ecology.

Supporting Information

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

Appendix S1 The aligned data matrix.

Appendix S2 List of haplotypes (as in Table 1) and estimates of the total expected number of haplotypes within individual colonies and within colony complexes sharing haplotypes, based on the Chao1 statistic.

Appendix S3 Secondary reconstruction of the 16S ribosome for two *A. eximius* individuals, chosen as exemplars with very short (CU O5 R14 1 3) and very long (CU 03 5 1 4) 'microsatellite' region. This region is indicated with a circle, showing that increased number of the TA repeats translate into a longer stem. The color bar indicates base-pair probabilities.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.