Diversification and host switching in avian malaria parasites

Robert E. Ricklefs[*] and Sylvia M. Fallon

Department of Biology, University of Missouri-St Louis, St Louis, MO 63121-4499, USA

The switching of parasitic organisms to novel hosts, in which they may cause the emergence of new diseases, is of great concern to human health and the management of wild and domesticated populations of animals. We used a phylogenetic approach to develop a better statistical assessment of host switching in a large sample of vector-borne malaria parasites of birds (*Plasmodium* and *Haemoproteus*) over their history of parasite–host relations. Even with sparse sampling, the number of parasite lineages was almost equal to the number of avian hosts. We found that strongly supported sister lineages of parasites, averaging 1.2% sequence divergence, exhibited highly significant host and geographical fidelity. Event-based matching of host and parasite phylogenetic trees revealed significant cospeciation. However, the accumulated effects of host switching and long distance dispersal cause these signals to disappear before 4% sequence divergence is achieved. Mitochondrial DNA nucleotide substitution appears to occur about three times faster in hosts than in parasites, contrary to findings on other parasite–host systems. Using this mutual calibration, the phylogenies of the parasites and their hosts appear to be similar in age, suggesting that avian malaria parasites diversified along with their modern avian hosts. Although host switching has been a prominent feature over the evolutionary history of avian malaria parasites, it is infrequent and unpredictable on time scales germane to public health and wildlife management.

Keywords: cospeciation; emerging disease; *Haemoproteus*; phylogeny; *Plasmodium*

1. INTRODUCTION

Emerging diseases are a cause for concern as habitat alteration increasingly brings humans and domesticated animals into contact with wildlife and disease vectors, and introduced organisms carry potentially dangerous pathogens to the far corners of the earth (Lanciotti *et al.* 1999; Daszak *et al.* 2000). The appearance of new diseases is, however, highly unpredictable (Van Riper *et al.* 1986; Anderson *et al.* 1999). Of the four forms of human malaria, *Plasmodium malariae* and *P. vivax* differ little from lineages found in New World monkeys. Another *vivax*-like form is similar to *P. simiovale*, which is endemic to Old World macaques (Escalante *et al.* 1995). The clear implication is that humans have acquired these diseases recently from distantly related hosts. However, the probability that a disease might switch from one particular host to another is difficult to estimate. Here we use a phylogenetic approach to characterize aspects of host specialization and to evaluate host switching in a large sample of malaria parasites of birds (Page & Hafner 1996; Hoberg *et al.* 1997). Our broad geographical sampling also allows us to assess the movement of parasite lineages between continents. By examining host switching over the long history of parasite–host relationships within a particular group, we hope to develop a better understanding of the nature of emerging diseases.

The malaria parasites of birds comprise an unknown number of species in the genera *Plasmodium* and *Haemoproteus* (phylum Apicomplexa, class Haemosporida, family Plasmodiidae) (Atkinson & Van Riper 1991). Morphologically based taxonomic treatments of avian malaria suggest that individual parasite species are restricted to host taxonomic families (Bennett *et al.* 1993, 1994). However, a recent molecular phylogenetic study has shown that tits (*Parus*: Paridae) harbour lineages of *Haemoproteus* that are nested within parasite clades recovered from a variety of Old World warblers (*Acrocephalus* and *Phylloscopus*: Sylviidae) (Bensch *et al.* 2000). Clearly, host switching occurs among haemosporidian blood parasites. However, the frequency and the taxonomic and geographical limits of host switching are essentially unknown. Our results, combined with other available data, comprise a sample of 68 lineages of *Plasmodium/Haemoproteus* recovered from 79 species of birds in 20 avian families distributed over six continents. These data provide new insights into the evolutionary relationships and host switching of malaria parasites, and permit a preliminary estimate of the rate of genetic divergence of malaria parasites relative to that of their avian hosts.

2. MATERIAL AND METHODS

(a) *Sequencing*

Parasite DNA was isolated from infected avian blood samples by cell lysis followed by salt precipitation (Puregene: Gentra systems, Minneapolis, MN). We amplified between 368 and 1128 base pairs (complete sequence) of the cytochrome *b* gene using the polymerase chain reaction (PCR). Primers were designed using published sequences of the cytochrome *b* gene for *Plasmodium* and *Haemoproteus* (Escalante *et al.* 1998). We used seven primers: (L14621 5′-ATG-CCT-AGA-CGT-ATT-CCT-GAT-TA-3′, L14902 5′-TTA-TTA-GCC-ACT-TGT-TAT-ACT-CC-3′, L15175 5′-GTG-CAA-CYG-TTA-TTA-CTA-A-3′, L15399 5′-AAA-AAT-ACC-CTT-CTA-TCC-AAA-TCT-3′, H15374 5′-CGA-GAC-TTA-ATA-GAT-TTG-GAT-AGA-3′, H15725 5′-CAT-CCA-ATC-CAT-AAT-AAA-GCA-T-3′,

[*] Author for correspondence (ricklefs@umsl.edu).

01PB0664.1

H15909 5′-TTG-TTC-YGC-TCA-ATA-CTY-AGA-3′). The primer number refers to the approximate nucleotide position of the 3′ end of the oligonucleotide according to the numbering system for the human sequence (Anderson *et al.* 1981). PCR reactions were run in 25 μl volumes that contained the following components in their final concentration: 1.5–2.5 mM of MgCl, 0.2 mM each of dNTP, 0.4 μM each of primer and 0.5 units of Taq polymerase. One millilitre of the DNA extraction solution was used for amplification. Cycling conditions were as follows: 1 min at 94 °C, followed by 35 cycles with 0.5 min denaturation at 94 °C, annealing at 44–50 °C, and elongation at 72 °C for 1.5 min. After 35 cycles, a final elongation step followed at 72 °C for 3 min. We purified the amplified product by gel extraction. Sequencing was carried out on an automated sequencer (ABI Prism 377: Perkin–Elmer) according to the manufacturer's protocol. The cytochrome *b* sequences were edited and aligned using DNA STAR software and are available through GenBank (accession numbers AF465547–AF465594). All sequences were translated and none contained stop codons.

We obtained cytochrome *b* sequences for 34 species of passerine hosts and sequences from another mitochondrial protein-coding region (ATPase 6,8; 842 bp) for five additional species for comparing host and parasite sequence divergence. These sequences were obtained from GenBank or from unpublished data of E. Bermingham and I. J. Lovette (available upon request). We generated additional sequences for cytochrome *b* using the primer pair L14987 5′-CCA-TCC-AAC-ATC-TCW-GCW-TGA-TG-3′, H15706 5′-TAT-GCG-AAT-AGG-AAR-TAY-CAY-TC-3′ (GenBank numbers AF465595–AF465598). Details concerning host taxon selection are available from the authors.

(b) *Phylogeny reconstruction*

The reconstruction was performed by a heuristic search for the shortest phylogenetic tree based on LogDet distances (Lockhart *et al.* 1994; Swofford *et al.* 1996). Phylogenetic analyses were performed with PAUP* (Swofford 1998). Phylogenies produced by parsimony and maximum-likelihood heuristic searches on a subset of 37 avian parasite lineages with greater than 800 bp of sequence plus the 17 mammal parasite lineages from GenBank did not contradict any of the supported LogDet clades and had similar overall topology. Mammalian *Plasmodium* sequences were used as a monophyletic outgroup. However, designating *Hepatocystis* as the only outgroup taxon did not change any of the relationships. A BLAST search of GenBank recovered no other cytochrome *b* sequences remotely related to those of malaria parasites. A likelihood-ratio test comparing maximum-likelihood trees for 54 taxa having more than 800 bp of sequence, with and without a molecular clock enforced, was significant overall ($\chi^2 = 185$). However, no individual clock-enforced branch length differed significantly from the non-clock-based branch length as determined from its standard error and a single degree of freedom.

(c) *Sampling*

Sources of parasite sequences and hosts associated with each parasite lineage are presented in electronic Appendices A and B available on The Royal Society's Publications Web site. We did not include a sample of *P. relictum* from Hawaii in this analysis because the original host and location are unknown; the sequence is identical to one recovered from a European sylviid warbler (no. 59, *Acrocephalus arundinaceus*). We also deleted the sequence for '*H. columbae*' from Escalante *et al.* (1998), which

may represent an incorrect parasite identification (A. Escalante, personal communication); it differs by only one nucleotide from a *Plasmodium* lineage obtained from *Acrocephalus arundinaceus* (no. 63). Parasites obtained from the commensal domestic pigeon (*Columbia livia*, no. 66), domestic fowl (*Gallus domesticus*, no. 58: *P. gallinaceum*), and house sparrow (*Passer domesticus*, no. 60: *P. elongatum*) were not scored for region. Named parasites are consistent with the assignment of lineages no. 1–46 to *Haemoproteus* and no. 47–68 to *Plasmodium*: no. 11, *H. sylvae*, no. 25, *H. majoris*, no. 63, *P. nucleophilum* (Bensch *et al.* 2000); no. 58, *P. gallinaceum*, no. 59, *P. relictum*, no. 60, *P. elongatum* (Escalante *et al.* 1998); no. 2, *H. paruli*, no. 28, *H. vireonis*, no. 35, *H. thraupi*, no. 55, *Haemoproteus* (?) spp., no. 56, *P. near elongatum* (C. Atkinson, personal communication); our observations of blood smears for lineages no. 1–8, 20, 21, 27–29, 31–36, 38, 41, 43, 45, 46, 48, 53–56, 61, 62, 64 and 66 are also consistent with these generic assignments. We use the host taxonomy of Sibley & Ahlquist (1990) in which taxonomic rank is defined by genetic distances obtained through DNA hybridization (family: 9–11 °C $\Delta T_H 50$).

(d) *Probability of co-occurrence*

We used a probability approach to determine whether related lineages of parasites tend to occur within the same host taxonomic family or region within our samples. Within each region (i), the probability of drawing a pair of species from a particular family (j) from our sample of hosts [$H(i,j)$] is

$$H(i,j) = \frac{n_{ij}(n_{ij} - 1)}{N_i(N_i - 1)},$$

where n_{ij} is the number of hosts sampled in family j within region i, and N_i is the total number of hosts sampled within region i. The probability of drawing a pair of species at random from the same family over the entire sample (H) is

$$H = \sum_{i=1}^{N_i} \frac{N_i}{N} \sum_{j=1} H(i,j),$$

where N is the total number of hosts sampled. The binomial probability of drawing at least k pairs of parasites in the same host family out of K tries is 1 minus the probability of obtaining up to $k - 1$ pairs, or

$$P = 1 - \sum_{x=0}^{k-1} \binom{K}{x} x^H (K - x)^{1-H}.$$

Additional analytical methods are described in § 3. All statistical analyses were carried out on the Statistical Analysis System (SAS) v. 6.12 (SAS Institute, Inc., Cary, NC).

3. RESULTS

(a) *Host switching*

Lineages of parasites sharing identical cytochrome *b* sequences and found in more than one host species (see electronic Appendix A) provide the clearest evidence of host switching, particularly when the hosts are well differentiated. Out of 10 such avian *Plasmodium* and *Haemoproteus* lineages, two were recovered from different species of host in the same genus (no. 1, *Sialia sialis* and *S. mexicanus*, Muscicapidae; no. 35, *Piranga rubra* and *P. olivacea*, Fringillidae), nine were from more distantly related species in the same host family, and in the tenth case the hosts belonged to the same superfamily. With respect to

geography, nine of the lineages shared hosts from the same region, and the tenth (no. 20) included two hosts from the West Indies (*Allenia fusca* and *Margarops fuscatus*) and one from North America (*Dumetella carolinensis*), all of which are in the family Mimidae. The ranges of *Margarops* and the migratory *Dumetella* overlap during the winter in the Bahamas.

We calculated the probability of such host and region conservatism among shared hosts by estimating the probability of drawing a single pair of host taxa at random from the same family or region, and then using the binomial distribution to estimate the probability of drawing the observed number of such pairs (see § 2). Because all the shared hosts are from the same region, the probability of drawing pairs from the same host family must be calculated within regions and then summed over regions. By this approach, the probability of obtaining 9 out of 10 pairs of host lineages from the same family within a region is $p = 0.000\ 17$, which demonstrates highly significant host–family conservatism.

Seventeen pairs of parasite sequences were selected from subterminal nodes representing pairs, triplets, or quartets of passerine hosts in figure 1 that had bootstrap support of greater than 60%. The average LogDet genetic distance (d_{ij}) between these lineages was 0.012 ± 0.011 s.d., based on cytochrome *b*. Of these pairs, 13 were from the same region ($p = 0.000\ 014$). Three additional pairs encompassing North America and the West Indies provide further evidence that Neotropical migrants link the parasite faunas of the two regions. These are lineages no. 3 versus 4,5: *Toxostoma rufum* (migrant, M) versus *Allenia fusca* (West Indian resident, WI; same host family: Mimidae); no. 53 versus 54,55: *Cinclocerthia ruficauda* (WI; Mimidae) versus *Helmitheros vermivora* (M; Parulidae), *Dendroica magnolia* (M; Parulidae), and *Baeolophus bicolor* (North American resident; Paridae); no. 27 versus 28,29: *Vireo altiloquus* (WI; Vireonidae), *V. olivaceus* (M), and *V. griseus* (M). When these are included as within-region pairs of lineages, the probability of 16 of 17 pairs becomes less than 10^{-5}.

The probability of drawing a pair of hosts from the same family was calculated within regions. Out of 17 pairs of parasite lineages, 7 were from hosts in the same genus and 13 in the same family ($p = 0.000\ 12$), again indicating highly significant host–family conservatism. This result is further supported by randomization of host families on the parasite phylogenetic tree and by using tree-based methods to estimate the frequency of cospeciation and switching events. We mapped the host family on the parasite phylogenetic tree, randomized the distribution of hosts on the tree by the shuffle utility of MACCLADE 4.0 (Maddison & Maddison 2000), and compared the number of steps, or host shifts, in the observed and randomized datasets. The minimum number of switches between host families mapped on the observed parasite phylogeny was 35. The number of steps for 100 randomized host distributions averaged 45.4 ± 1.8, with a minimum of 40 steps ($t = -5.7$, $p < 10^{-7}$). For geographical region, the 27 observed steps were significantly less than the average for 100 randomized distributions (35.7 ± 1.9, minimum = 31; $t = 4.7$, $p < 10^{-5}$).

We used the software TREEFITTER developed by Fredrik Ronquist (http://www.zoologi.uu.se/systzoo/research/treefitter)

to estimate, by means of a tree-based assessment, the number of codivergence (cospeciation), duplication (splitting of parasite lineage within a host), sorting (extinction) and switching events among the malaria parasites of birds (Page 1995; Ronquist 1998). The host phylogeny was based on Sibley & Ahlquist (1990) with cytochrome *b* sequence data used to resolve species relationships within genera or families. Branch lengths were equal. Results depended on 'costs' assigned to each of the event types. The default values in TREEFITTER of cospeciation = 0, duplication = 0, extinction = 1 and switching = 2 yielded significantly more cospeciation events (9–16; $p = 0.008$, 1000 permutations) and fewer switching events (49–52; $p = 0.012$) than found in randomized trees; duplication (2–6 events) and sorting (2–8 events) did not differ significantly from randomized trees. When costs were set to reflect a maximum cospeciation model (costs = -1, 0, 0, 0), neither cospeciation (34; $p = 0.12$) nor switching (6–32; $p = 0.27$) events were significantly more frequent than random, however many extinction events (97–366; $p = 0.001$) were required to balance the predominance of cospeciation. Cost assignments of 0, 2, 2 and 1 resulted in 11 cospeciation events ($p = 0.000$), 55 switching events ($p = 0.000$), and insignificant numbers of duplication (1) and extinction (0) events. Thus, tree-based estimates of event frequencies support a predominance of cospeciation and host switching (apparently often between closely related hosts), but provide support for extinction of parasite lineages only when parasite lineage splitting is biased in favour of cospeciation.

Host conservatism and cospeciation are a prominent feature of our data, however switching between hosts should cause this signal to weaken with depth in the phylogeny. To examine this possibility, we sampled 14 additional subterminal nodes with less than 60% bootstrap support, averaging 0.041 ± 0.020 s.d. genetic distance between terminal taxa, and tested them for geographical and taxonomic conservatism. Of the 14 pairs of taxa, three were from the same geographical region ($p = 0.73$). Because geography was not a significant effect, we calculated the probability of drawing two parasite lineages from the same host families within the global sample. Of 14 pairs, four were from the same family ($p = 0.085$). Thus, by a parasite genetic distance of 4% sequence divergence, both geographical and taxonomic signal had disappeared from the sample.

We used logistic regression to quantify this decline as the genetic divergence at which the probability of a pair of species belonging to the same region or family equals 0.5 (figure 2). Logistic regression (SAS CATMOD procedure) models the probability of co-occurrence (p) as a function of genetic distance (d) by the expression $p = \exp(a + bd)/[1 + \exp(a + bd)]$. For geography, with North America and the West Indies combined, $a = 2.56 \pm 0.86$ s.e. ($p = 0.03$) and $b = -64.9 \pm 7.0$ s.e. ($p = 0.008$). For taxonomy, $a = 1.76 \pm 0.57$ s.e. ($p = 0.002$) and $b = -67.7 \pm 23.4$ s.e.($p = 0.004$). For geographical region, with North America and the West Indies considered as a single region, a probability of 0.5 is reached at a depth of $d_{ij} = 0.039$. For host family, including parasite lineages with multiple hosts (i.e. $d_{ij} = 0$, of which 9 of 10 were confamilial), the balance point occurs at
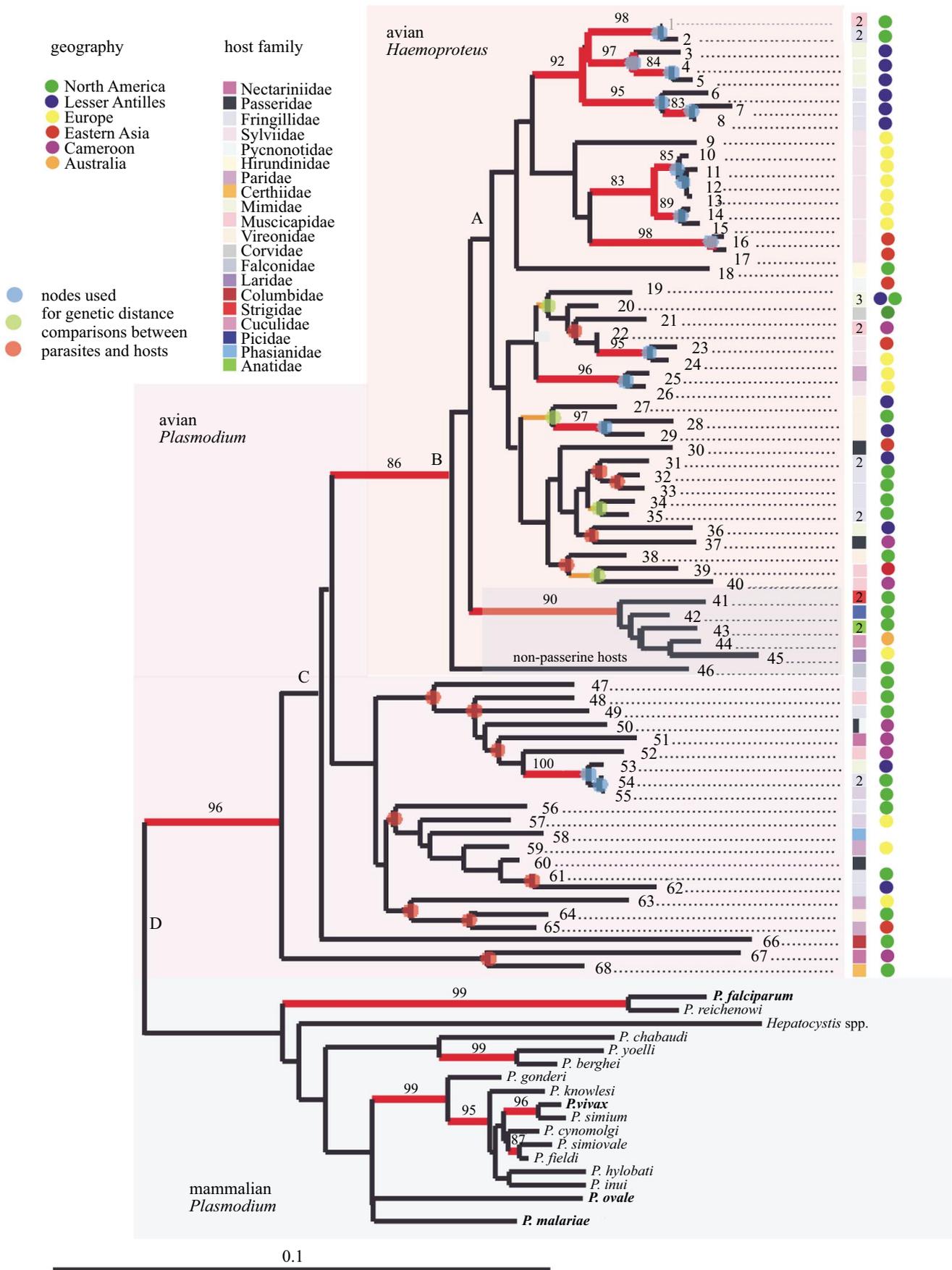
Figure 1. Phylogenetic relationships of 68 lineages of avian malaria parasites and 17 lineages of mammalian *Plasmodium*. Branches having bootstrap support (100 replicates) are shown as red (+ 80%) or orange (60–79%) lines. Lineages are numbered rather than named because most have not been described or are not diagnosable (see electronic Appendices A and B). Host family and geographical region are colour coded at right. Lineages recovered from more than one host or region are indicated by number of instances in the coloured key symbols. Subterminal nodes used for subsequent sister taxon comparisons are indicated by coloured dots. Mammalian lineages in bold type refer to human malaria parasites.
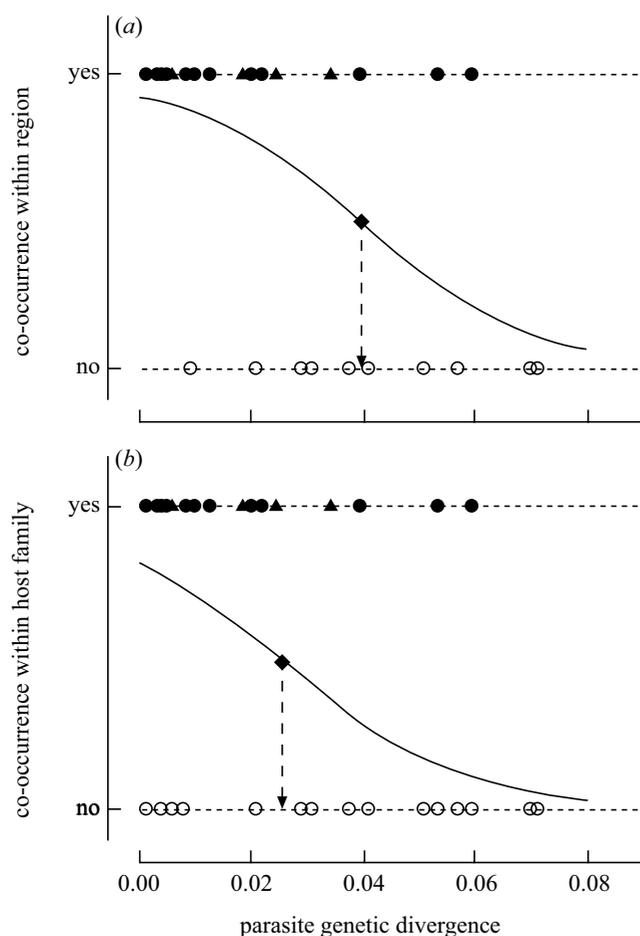
Figure 2. Logistic regressions of co-occurrence within regions (*a*) and within families (*b*) as a function of genetic divergence between parasite lineages (see § 2). In this analysis, North America and the West Indies are considered part of the same region. The lineage pairs are indicated in figure 1 by blue or green dots. Filled symbols, parasite lineages within the same family or region: open symbols, different families or regions. Filled triangles indicate host pairs between North America and the West Indies.

$d_{ij} = 0.026$. However, although parasite lineages appear to spread to new families of hosts within regions more readily than they spread to new regions, the logistic regression parameters did not differ significantly ($\chi^2 = 1.94$, d.f. $= 1$, $p = 0.164$).

Although our sampling was sparse, several host taxa harboured two or more lineages of parasites. Two lineages were recovered from each of eight host species (*Margarops fuscatus*, no. 20,36; *Coereba flaveola*, no. 6,31; *Vireo griseus*, no. 38,64; *Acrocephalus palustris*, no. 13,14; *Phylloscopus trochilus*, no. 24,26; *Nectarinia olivacea*, no. 51,67; *Piranga rubra*, no. 34,35; *Loxigilla noctis*, no. 8,31), three lineages from one other (*Allenia fusca*, no. 4,5,20), and four lineages from yet another (*Acrocephalus arundinaceus*, no. 10,11,59,63). Of the pairs of parasite lineages co-occurring within host species, fewer than half are closely related within the parasite phylogenetic tree. This suggests either, that parasites commonly switch hosts across large taxonomic distances or, that the larger clades of parasites arose before the origin of modern families of birds and that the distribution gaps resulted from extinction or inadequate sampling (Paterson *et al.* 1993).
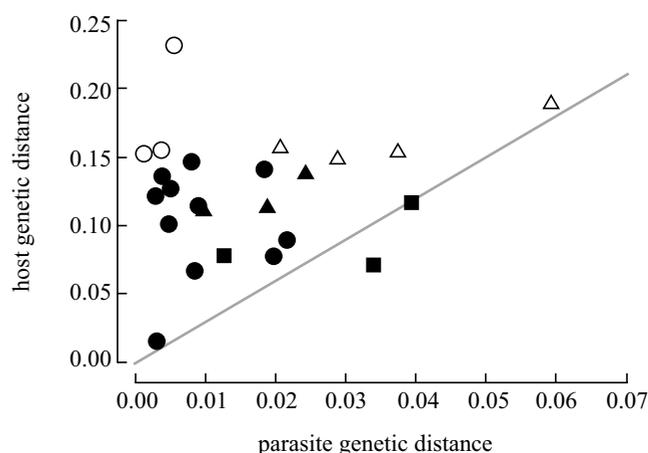
Figure 3. The relationship between cytochrome *b* sequence divergence (LogDet distances) between pairs of closely related parasites and sequence divergence between their hosts, based on 24 subterminal nodes indicated in figure 1 (data in electronic Appendix C). Parasites were the focus of the phylogenetic analysis and therefore the selection of sister parasite lineages establishes the horizontal axis against which host divergence is compared. Avian cytochrome *b* sequences were obtained from GenBank or generated from blood samples used in this study (see electronic Appendix D). The line with a slope of 3 d_{ij}(host)/d_{ij}(parasite) was drawn by eye assuming that parasite divergence in this sample represents cospeciation or switching between closely related hosts. This line is supported by lower edge regression. Filled symbols, within host–family comparisons; open symbols, between host–family comparisons. Circles, more than 80% bootstrap support; squares, between 60–80% bootstrap support; triangles, less than 60% bootstrap support.

(b) *Relative rates of host and parasite nucleotide substitution*

When parasites cospeciate with their hosts or switch among closely related hosts, parasite sister lineages cannot be older than host sister lineages. Accordingly, cospeciation should place a lower edge on the distribution of host taxon divergence with respect to genetic distances between their parasites (Page & Hafner 1996). This edge should pass through the origin and have a slope equal to the ratio of rates at which host sequences and parasite sequences diverge [d_{ij}(host)/d_{ij}(parasite)]. Comparisons of host and parasite cytochrome *b* sequence divergences (see electronic Appendix C) among closely related parasite lineages (figure 3) suggest an edge to the distribution with a slope of about 3, indicating that host sequences diverge about three times faster than parasite sequences. We used a lower bound regression to estimate the approximate slope of the line in figure 3 (Blackburn *et al.* 1992). Parasite genetic divergence was divided into intervals of 0.005 sequence divergence and the lowest value for host genetic divergence within each interval ($n = 8$) was used in a linear regression of host distance on parasite distance. When an intercept was estimated, the slope of the regression was 2.61 ± 0.56 ($t = 4.6$, $p = 0.0024$). When the regression was forced through the origin, the slope was 3.48 ± 0.35 ($t = 9.8$, $p < 10^{-4}$). The major axis regression slope (Box 15.2; Sokal & Rohlf 1995), which assumes that both parasite and host divergence are estimated with error (type II regression), was 3.91 (95% confidence limits, 2.0–22.6).
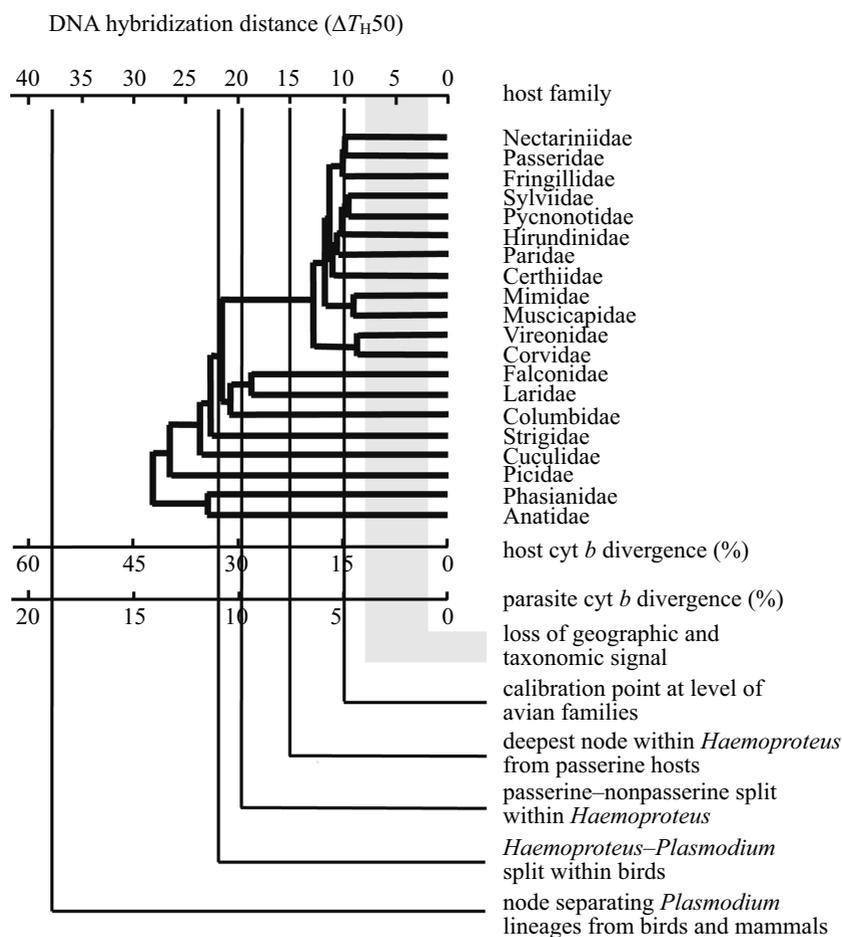
Figure 4. Phylogenetic tree for families of birds included in this study based on the DNA–DNA hybridization studies of Sibley & Ahlquist (1990). The scale of DNA melting point differences ($\Delta T_H 50$, °C) is indicated at the top. Host and parasite cytochrome *b* divergence (below) are scaled to each other in a 3 : 1 ratio. Family-level distinctions are designated within a range $\Delta T_H 50 = 9–11$ °C. Cytochrome *b* distances between taxa in sister families of passerine birds average about $d_{xy} = 0.15$. Thus, we have anchored the parasite and host phylogenetic trees at a distance of d_{xy}(parasite) = 0.05, $\Delta T_H 50$(host) = 10 °C, and d_{xy}(host) = 0.15.

None of these estimates differ significantly from a slope of 3 and all are significantly greater than 1. Although such slow parasite sequence evolution is unexpected, it is consistent with ages estimated for lineages of mammalian *Plasmodium* from other gene sequences (Escalante & Ayala 1994; Escalante *et al.* 1998).

4. DISCUSSION

The potential for malaria parasites to switch from avian to mammalian hosts, including humans, is very remote. The absence of parasite lineages shared between avian and mammalian hosts is not surprising, however, considering differences in the host environment, including nucleated avian erythrocytes. Accordingly, we should expect that the *Plasmodium* parasites of reptiles, which also have nucleated erythrocytes, would be more closely related to those of birds. Escalante & Ayala (1994) determined that the reptilian *P. mexicanum* (host: western fence lizard *Sceloporus occidentalis*) was more closely allied with the avian parasites *P. gallinaceum* and *P. lophurae* than it was with mammalian (primate and rodent) lineages. Perkins (2000) reported a cytochrome *b* divergence of *ca.* 8.0–8.5% between *P. gallinaceum* (figure 1, no. 58) and *P. azurophilum*

recovered from West Indian *Anolis* lizards. This would place the node connecting these lineages near the base of the avian parasite phylogeny, but well removed from mammalian *Plasmodium*.

Several studies comparing host and symbiont phylogenetic trees among more closely related taxa have shown a high level of cospeciation and host taxonomic fidelity with only infrequent host switching (Hafner *et al.* 1994; Moran *et al.* 1995; Page *et al.* 1998). These cases involve lice that are external parasites of rodents and swiftlets, and endosymbiotic bacteria of aphids, all of which are transmitted vertically from parent to offspring or horizontally by direct contact among individuals of the same species. Malaria parasites differ in being transmitted by dipteran vectors, many of which have broad host distributions, representing a higher potential for host switching. For example, the avian malaria parasite *P. relictum* was introduced to Hawaii and spread by the introduced mosquito *Culex quinquefasciatus* to many native and non-native Hawaiian birds (Van Riper *et al.* 1986). We presume that most mosquito and biting midge vectors have broad host preferences and that host specialization by malaria parasites within avian communities must represent virulence–resistance interactions between *Plasmodium* and *Haemoproteus* and their avian

hosts. Such interactions may also explain the relative conservatism of parasite evolution and diversification in avian malaria, but also allow for occasional switching between unrelated hosts.

The low level of cytochrome *b* nucleotide divergence in avian (and mammalian) malaria parasites compared with that of host taxa is surprising. Short generation times and large population sizes of parasites argue for rapid evolution, which is supported by the relatively greater sequence divergence among parasites than among hosts observed in several systems (Hafner *et al.* 1994; Moran *et al.* 1995; Page *et al.* 1998). The reverse situation in avian malaria is difficult to explain. However, apicomplexan and avian mitochondrial DNAs differ in several important ways that potentially could influence rates of mutation and nucleotide substitution. Avian mtDNA exists as a *ca.* 16 kb circular molecule typical of vertebrates. *Plasmodium* mtDNA is the most reduced form known (Feagin 1994; Gray *et al.* 1999), consisting of a 6 kb linear molecule arranged in several linear concatemers, typically of four units, and having an unusual replication mechanism associated with a recombinant process (Preiser *et al.* 1996).

Assuming that parasite diversification in *Plasmodium* and *Haemoproteus* occurred primarily by cospeciation or switching among closely related hosts (figure 3), the low level of genetic divergence among parasites would represent a slow rate of nucleotide substitution rather than a recent origin of avian malaria and its rapid spread across highly differentiated host taxa. Accepting a 3 : 1 ratio in the rate of nucleotide divergence in hosts and parasites for the purpose of discussion, we can compare the two phylogenies placed on the same relative time-scale. Using this reciprocal calibration, we overlaid landmarks from the parasite phylogeny in figure 1 onto a phylogeny of the families of birds included in this study (figure 4). This comparison implies the following relationships between parasite and host evolutionary diversification. (i) The loss of host and region signals between 1 and 4% parasite sequence divergence is subsequent to the origin of modern families of birds. Thus, the occurrence of a single parasite lineage, or two closely related lineages, on hosts in different families is clearly the result of switching between distantly related hosts. (ii) The deepest node within *Haemoproteus* lineages recovered from passerine hosts (node A in figure 1, average $d_{ij} = 0.075 \pm 0.018$ s.d.) occurs near the base of the passerine radiation. Because the basal lineages of *Haemoproteus* are non-passerine, we suggest that most of the evolutionary diversification of *Haemoproteus* portrayed in figure 1 was associated with the radiation of passerines, which were the primary taxa sampled in this study (Bennett 1993). The deepest node within *Haemoproteus* (B, 0.099 ± 0.028 s.d.), which approximates the split between passerine and non-passerine parasite lineages, is consistent with the split between passerine and non-passerine hosts. (iii) The avian *Plasmodium–Haemoproteus* split (C, 0.110 ± 0.028 s.d.) is close to the base of modern birds, suggesting that the current diversity of avian malaria parasites has co-occurred for the most part with the radiation of modern birds. (iv) The mammal–bird *Plasmodium* split (D, 0.189 ± 0.021 s.d.) occurs prior to the diversification of modern birds, suggesting that the ancestors of modern birds carried the ancestors of modern avian malaria infections. Thus,

occurrences of widely divergent parasite lineages in the same host are consistent with long-term co-occurrence of parasite lineages within host clades, combined with extinction or inadequate sampling of parasite lineages, creating apparent gaps in the distribution of parasites among hosts (Hoberg *et al.* 1997; Paterson & Gray 1997; Paterson *et al.* 2000).

The evolution and distribution of avian blood parasites are conservative with respect to host phylogeny and geography. Clearly, cospeciation or switching between closely related hosts is the most common mode of parasite lineage proliferation, although host switching and long distance dispersal occur frequently enough to erase host taxonomic and geographical signatures at a fairly shallow depth within the parasite phylogenetic tree. Because vectors can potentially distribute blood parasites widely through an avian community, host specificity probably represents strong coevolution between parasite and host populations (Apanius *et al.* 2000). If this were the case, as our data suggest, then host switching across great host-taxonomic distance would depend more on serendipitous genetic changes in the parasite than on ecological changes in host habitat use or abundance of vectors. Thus, although habitat alteration and introduction of alien organisms will probably increase the frequency of emergence of new diseases, switching of disease organisms between distantly related hosts will be infrequent and essentially unpredictable.

REFERENCES

Anderson, J. F., Andreadis, T. G., Vossbrinck, C. R., Tirrell, S., Wakem, E. M., French, R. A., Garmendia, A. E. & Van Kruiningen, H. J. 1999 Isolation of West Nile virus from mosquitoes, crows, and a Cooper's hawk in Connecticut. *Science* **286**, 2331–2333.

Anderson, S. (and 13 others) 1981 Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.

Apanius, V., Yorinks, N., Bermingham, E. & Ricklefs, R. E. 2000 Island and taxon effects in the prevalence of blood parasites and activity of the immune system in Lesser Antillean birds. *Ecology* **81**, 1959–1969.

Atkinson, C. T. & Van Riper, C. I. 1991 Pathogenicity and epizootiology of avian haematozoa: *Plasmodium*, *Leucocytozoon*, and *Haemoproteus*. In *Bird–parasite interactions. Ecology, evolution, and behavior* (ed. J. E. Loye & M. Zuk), pp. 19–48. New York: Oxford University Press.

Bennett, G. F. 1993 Phylogenetic distribution and possible evolution of the avian species of the Haemoproteidae. *Syst. Parasitol.* **26**, 39–44.

Bennett, G. F., Bishop, M. A. & Peirce, M. A. 1993 Checklist of the avian species of *Plasmodium* Marchiafava and Celli, 1885 (Apicomplexa) and their distribution by avian family and Wallacean life zones. *Syst. Parasitol.* **26**, 171–179.

Bennett, G. F., Peirce, M. A. & Earlé, R. A. 1994 An annotated checklist of the valid avian species of *Haemoproteus*, *Leucocytozoon* (Apicomplexa, Haemosporida) and *Hepatozoon* (Apicomplexa, Haemogregarinidae). *Syst. Parasitol.* **29**, 61–73.

Bensch, S., Stjernman, M., Hasselquist, D., Ostman, O., Hansson, B., Westerdahl, H. & Pinheiro, R. T. 2000 Host specificity in avian blood parasites: a study of *Plasmodium* and *Haemoproteus* mitochondrial DNA amplified from birds. *Proc. R. Soc. Lond.* B **267**, 1583–1589. (DOI 10.1098/rspb. 2000.1181.)

Blackburn, T. M., Lawton, J. H. & Perry, J. N. 1992 A method of estimating the slope of upper bounds of plots of body size and abundance in natural animal assemblages. *Oikos* **65**, 107–112.

Daszak, P., Cunningham, A. A. & Hyatt, A. D. 2000 Emerging infectious diseases of wildlife—threats to biodiversity and human health. *Science* **287**, 443–449.

Escalante, A. A. & Ayala, F. J. 1994 Phylogeny of the malarial genus *Plasmodium*, derived from rRNA gene sequences. *Proc. Natl Acad. Sci. USA* **91**, 11 373–11 377.

Escalante, A. A., Barrio, E. & Ayala, F. J. 1995 Evolutionary origin of human and primate malarias—evidence from the circumsporozoite protein gene. *Mol. Biol. Evol.* **12**, 616–626.

Escalante, A. A., Freeland, D. E., Collins, W. E. & Lal, A. A. 1998 The evolution of primate malaria parasites based on the gene encoding cytochrome *b* from the linear mitochondrial genome. *Proc. Natl Acad. Sci. USA* **95**, 8124–8129.

Feagin, J. E. 1994 The extrachromosomal DNAs of Apicomplexan parasites. *A. Rev. Microbiol.* **48**, 81–104.

Gray, M. W., Burger, G. & Lang, B. F. 1999 Mitochondrial evolution. *Science* **283**, 1476–1481.

Hafner, M. S., Sudman, P. D., Villablanca, F. X., Spradling, T. A., Demastes, J. W. & Nadler, S. A. 1994 Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science* **265**, 1087–1090.

Hoberg, E. P., Brooks, D. R. & Siegel-Causey, D. 1997 Host–parasite co-speciation: history, principles, and prospects. In *Host–parasite evolution. General principles and avian models* (ed. D. H. Clayton & J. Moore), pp. 213–235. Oxford University Press.

Lanciotti, R. S. (and 23 others) 1999 Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* **286**, 2333–2337.

Lockhart, P. J., Steel, M. A., Hendy, M. D. & Penny, D. 1994 Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* **11**, 605–612.

Maddison, D. R. & Maddison, W. P. 2000 MACCLADE 4.0. Sunderland, MA: Sinauer.

Moran, N. A., Vondohlen, C. D. & Baumann, P. 1995 Faster evolutionary rates in endosymbiotic bacteria than in cospeciating insect hosts. *J. Mol. Evol.* **41**, 727–731.

Page, R. D. M. 1995 Parallel phylogenies: reconstructing the history of host–parasite assemblages. *Cladistics* **10**, 155–173.

Page, R. D. M. & Hafner, M. S. 1996 Molecular phylogenies and host–parasite cospeciation: gophers and lice as a model system. In *New uses for new phylogenies* (ed. P. H. Harvey, A. J. L. Brown, J. Maynard Smith & S. Nee), pp. 255–270. Oxford University Press.

Page, R. D. M., Lee, P. L. M., Becher, S. A., Griffiths, R. & Clayton, D. H. 1998 A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. *Mol. Phylogenet. Evol.* **9**, 276–293.

Paterson, A. M. & Gray, R. D. 1997 Host–parasite co-speciation, host switching, and missing the boat. In *Host–parasite evolution. General principles and avian models* (ed. D. H. Clayton & J. Moore), pp. 236–250. Oxford University Press.

Paterson, A. M., Gray, R. D. & Wallis, G. P. 1993 Parasites, petrels and penguins—does louse presence reflect seabird phylogeny? *Int. J. Parasitol.* **23**, 515–526.

Paterson, A. M., Wallis, G. P., Wallis, L. J. & Gray, R. D. 2000 Seabird and louse coevolution: complex histories revealed by 12S rRNA sequences and reconciliation analyses. *Syst. Biol.* **49**, 383–399.

Perkins, S. L. 2000 Species concepts and malaria parasites: detecting a cryptic species of *Plasmodium*. *Proc. R. Soc. Lond.* B **267**, 2345–2350. (DOI 10.1098/rspb.2000.1290.)

Preiser, P. R., Wilson, R. J. M., Moore, P. W., McCready, S., Hajibagheri, M. A. N., Blight, K. J., Strath, M. & Williamson, D. H. 1996 Recombination associated with replication of malarial mitochondrial DNA. *European Mol. Biol. Org. J.* **15**, 684–693.

Ronquist, F. 1998 Phylogenetic approaches in coevolution and biogeography. *Zool. Scripta* **26**, 313–322.

Sibley, C. G. & Ahlquist, J. E. 1990 *Phylogeny and classification of the birds of the world*. New Haven, CT: Yale University Press.

Sokal, R. R. & Rohlf, F. J. 1995 *Biometry*, 3rd edn. New York: Freeman.

Swofford, D. L. 1998 PAUP*. *Phylogenetic analysis using parsimony (*and other methods)*. Sunderland, MA: Sinauer.

Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. 1996 Phylogenetic inference. In *Molecular systematics* (ed. D. M. Hillis, C. Moritz & B. K. Mable), pp. 407–514. Sunderland, MA: Sinauer.

Van Riper III, C., Van Riper, S. G., Goff, M. L. & Laird, M. 1986 The epizootiology and ecological significance of malaria in Hawaiian land birds. *Ecol. Monogr.* **56**, 327–344.