Diagnosing genetically diverse avian malarial infections using mixed-sequence analysis and TA-cloning

J. PÉREZ-TRIS* *and* S. BENSCH

Department of Animal Ecology, Lund University, Ecology Building, SE-22362 Lund, Sweden

(*Received 10 August 2004; revised 13 December 2004; accepted 17 December 2004*)

SUMMARY

Birds harbouring several malarial parasites are common in the wild, and resolving such multiple infections is important for our understanding of host–parasite relationships. We propose a simple and reasonably accurate method for detecting and resolving multiple infections, based on the analysis of parasite cytochrome *b* DNA sequences: genetically mixed infections are first identified by double nucleotide peaks on sequence electropherograms, and later retrieved by TA-cloning. We applied this method to wild birds, and to experimentally created mixes with varying proportion of two parasites (*Plasmodium* spp. and *Haemoproteus* spp.). In general, the method was very efficient in detecting and resolving multiple infections, but some problems were encountered. Several multiple infections were erroneously scored as simple, either because one of the parasite lineages was a better target for the primers used, or because it was much more abundant in the mix. On the other hand, single nucleotide substitutions and template switching during PCR produced artificial sequences in some clones. We discuss the utility of the method, and propose a framework for its use when screening for genetically diverse avian malarial parasites.

Key words: cytochrome *b*, host–parasite relationships, within-host interactions, multiple parasite infections, nested PCR.

INTRODUCTION

Understanding the ecology and evolution of host–parasite relationships is complicated because parasitized individuals commonly harbour several different parasites. Coexisting parasites, particularly if they are closely related, are expected to establish a variety of interactions, from competition for limited host resources to mutualistic associations (Bruce *et al*. 2000; Paul *et al*. 2002; Lello *et al*. 2004). The outcome of such in-host processes may determine various aspects of host–parasite relationships, such as the structuring of parasite communities or the evolution of parasite virulence (van Baalen & Sabelis, 1995; Poulin, 1997; Read & Taylor, 2001).

The advent of PCR-based techniques for parasite detection and identification has significantly improved our understanding of host–parasite interactions, enabling differentiation between closely related, often cryptic, parasite species (Perkins, 2000; Bensch *et al*. 2004; Ricklefs, Fallon & Bermingham, 2004; Yamasaki *et al*. 2004). A striking example is the recent advance in the knowledge of the diversity and patterns of evolution of avian blood parasites of the genera *Plasmodium* and *Haemoproteus*, which have both ecological and economical importance as agents of infectious disease, including avian malaria, in wild and domestic birds (Atkinson & van Riper, 1991). Molecular methods have allowed, among other achievements, a re-interpretation of the ''one host-one parasite'' idea, revealing a high degree of cospeciation between hosts and parasites, frequent host shifting, and multiple instances of cryptic speciation (Bensch *et al*. 2000; Waldenström *et al*. 2002; Bensch *et al*. 2004; Ricklefs *et al*. 2004).

However, these analyses have broadly overlooked multiply infected birds, although these may be frequent in natural populations (e.g., van Riper, Atkinson & Seed, 1993). With ordinary PCR techniques, multiple infections produce mixes of amplified products distinguishable by a few bases, and the samples are often discarded in further analyses to avoid analytical problems intrinsic to the handling of mixed sequences (but see Waldenström *et al*. 2004). Genetically diverse infections can reflect the coexistence of different genotypes of one parasite species within the same host, or can reveal co-infection by different parasite species or genera. Therefore, retrieving the different parasites present in a single host would allow us to approach important issues in host–parasite interactions, such as how many different parasite lineages can infect one individual host, the structure of within-host parasite associations, or how coexistence may affect evolution of virulence of different parasites, both within and among parasite species (van Baalen & Sabelis, 1995; Poulin, 1997; Read & Taylor, 2001; de Roode *et al*. 2004; Lello *et al*. 2004).

Recently, molecular methods have been developed to screen for multiple parasite infections, such as multiplex PCR or real-time PCR techniques

* Corresponding author: Tel.: +46 46 222 37 96. Fax: +46 46 222 47 16. E-mail: javier.perez-tris@zooekol.lu.se

doi:10.1017/S003118200500733X Printed in the United Kingdom

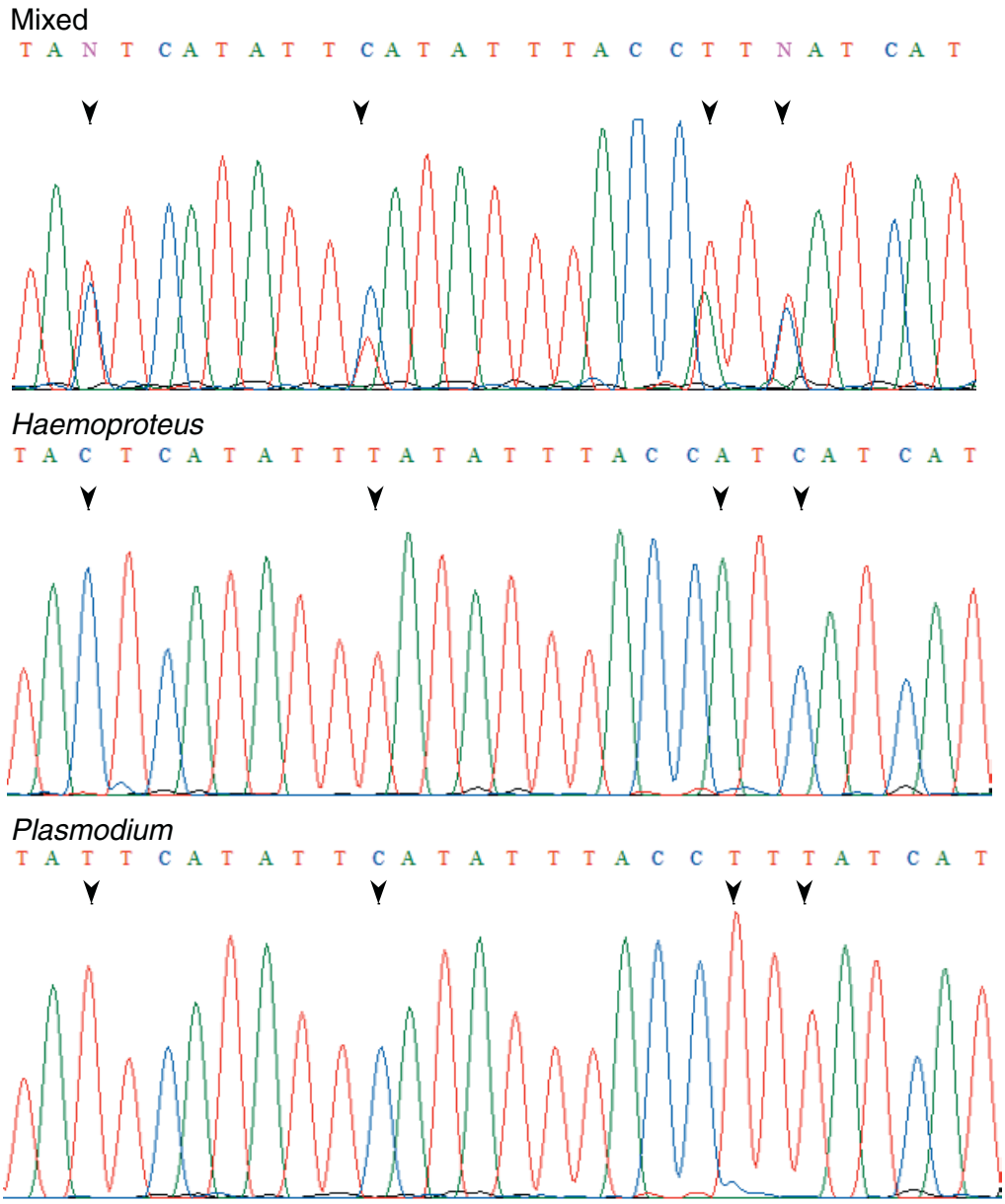

Fig. 1. Segment of the sequence electropherograms corresponding to a natural mixed infection from a blackcap. The graphs show the sequencing signal of the original (mixed) PCR product, and of two clones obtained from a second PCR product revealing parasites in the mix (*Plasmodium* sp. and *Haemoproteus* sp. in this case). Note the correspondence between the variable sites distinguishing the cloned sequences and the double peaks observed in the original sequence (arrows). The sequences on top of each graph were automatically produced by sequencing analysis software (unresolved bases are coded N).

(Cheesman *et al*. 2003; Perandin *et al*. 2004; Yamasaki *et al*. 2004). These methods make use of known sequence variation to specifically detect different parasite lineages (Cheesman *et al*. 2003; Perandin *et al*. 2004). However, such specific assays cannot be used for diagnosing multiple infections when mixing lineages have not yet been described. Inopportunely, many avian malarial parasites might be first discovered in a mixed infection, and since differentiated lineages often are genetically very similar (some cryptic species may differ by just 0·4% in the cytochrome *b* sequence; Bensch *et al*. 2004), they may be overlooked.

We used PCR-based methods to screen for *Plasmodium* spp. and *Haemoproteus* spp. of blackcaps (*Sylvia atricapilla*) and garden warblers (*S. borin*), during a study of the spatio-temporal structuring of parasite communities of these host species (unpublished data). During this survey, several birds were suspected to be simultaneously infected by more than one parasite lineage. Putative multiple infections were revealed by superimposed double nucleotide peaks on the sequence electropherograms (Fig. 1). In principle, any mixed PCR products can be separated by cloning (e.g., Waldenström *et al*. 2004), but this method involves some potential

https://doi.org/10.1017/S003118200500733X Published online by Cambridge University Press

problems. Firstly, parasite DNA might be amplified from very low quantities in a background of host DNA (some PCR methods are sensitive enough to detect infections with concentrations of parasite DNA corresponding to 1 infected erythrocyte per 100 000; Fallon *et al*. 2003; Waldenström *et al*. 2004). This would increase the risk to clone artificially mutated sequences, resulting from accumulation of base substitutions or template switching (the so called jumping PCR), particularly if such PCR errors occur during the initial amplification cycles (Hofreiter *et al*. 2001). Secondly, we do not know whether double peaks on electropherograms can reliably reveal multiple infections if the proportion of the mixing parasites changes. For example, if two parasites coexist at very different intensities, the most common parasite could dominate the PCR reaction preventing the amplification of the scarce one. This problem might be increased when the mix is formed by *Plasmodium* spp. and *Haemoproteus* spp. parasites, because most PCR methods were developed from conserved *Plasmodium* sequences (e.g. Richard *et al*. 2002), and hence might preferentially amplify this genus. Unfortunately, known avian *Plasmodium* and *Haemoproteus* sequences are too similar to develop primers that can amplify each genus separately, without loosing some lineages of the focal genus (Hellgren *et al*. 2004). These potential problems would lead to an underestimation of the frequency of multiple infections in natural populations.

In this paper, we examine both natural multiple infections and experimentally created parasite mixes, to assess the applicability of the analysis of mixing DNA sequences to the study of multiple infections of avian malarial parasites. Our goal is to identify the pros and cons of this technique, and to provide a framework to its future application in studies of the ecology and evolution of host–parasite relationships.

MATERIALS AND METHODS

Detection of multiple infections in wild birds

We analysed 361 blackcaps and 15 garden warblers captured either in summer or in winter at different locations in Spain, Belgium, France and Sweden. We obtained blood samples by venipuncture, stored them in preservation buffer, and released birds at the site of capture. Total DNA was extracted from the samples using standard phenol/chloroform or NH_4Ac methods, and diluted to a working DNA concentration of 25 ng/μl. We did not consider host species, geographical origin or time of capture in further analyses.

We detected parasite infections using a nested PCR method (Waldenström *et al*. 2004) designed to amplify 479 bp of the cytochrome *b* gene of *Plasmodium* and *Haemoproteus* parasites from avian total blood DNA. The method started with a pre-amplification (20 cycles) using primers located outside the target fragment (HaemFN and HaemR2N), followed by a final amplification (35 cycles) with the primers HaemF and HaemR2 (Waldenström *et al*. 2004). In both steps, PCR reactions were set up in 25 μl total volumes including 25 ng of template DNA (1 μl of pre-amplified PCR product in the second reaction), 1 × PCR buffer (Perkin Elmer), 0·125 mM of each nucleotide, 0·4 μM of each primer, 1·1 mM $MgCl_2$, and 0·5 units of AmpliTaq DNA polymerase (Perkin Elmer). Reactions started with 3 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 50 °C and 45 s at 72 °C, and they were terminated by a 10-min extension at 72 °C. We evaluated 2·5 μl of each final reaction on 2% agarose gels stained with ethidium bromide and using 0·5 × TBE buffer. A total of 196 infected birds were thus identified.

Samples with positive PCR reactions were precipitated by adding 11 μl of 8 M NH_4Ac and 33 μl of ethanol, and resuspended in 10–15 μl water. We used 2 μl for direct sequencing using a dye-terminator AmpliCycle® sequencing kit and an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems), following manufacturer's recommendations. Reading data were processed with the ABI PRISM® Sequencing Analysis Software v3·7 (Applied Biosystems), and output files were visualized, and sequences edited, using BioEdit (Hall, 1999). Sequence electropherograms were printed and carefully inspected for double nucleotide peaks. Low-quality sequences (with high background or too weak signal) were discarded, and the corresponding samples were amplified again.

TA-cloning of natural multiple infections

Samples showing at least one double peak on the sequence electropherogram were subjected to a second, independent amplification, using the same procedure as described above, and the resulting PCR products were cloned using a TOPO TA-Cloning® kit (Invitrogen). Within 24 h of amplification, PCR products were inserted into vectors (pCR® 2.1-TOPO®), which were used to transform One Shot® chemically competent *Escherichia coli* cells by heat shock at 42 °C, according to manufacturer's indications. Cells were cultured on S.O.C. medium (Invitrogen) at 37 °C for 1 h in a shaking incubator, and then plated on LB-medium supplemented with 50 μg/ml Ampicilin and 50 μl of X-gal (40 mg/ml). Plates were incubated overnight at 37 °C. White colonies (containing transformant cells) were picked and diluted in 100 μl water. DNA was extracted by heating at 100 °C during 3 min.

We amplified a minimum of 8 DNA inserts from each plate using the primers M13 forward and M13 reverse included in the kit, and the PCR

https://doi.org/10.1017/S003118200500733X Published online by Cambridge University Press

Table 1. Relative concentration of *Plasmodium* and *Haemoproteus* in experimentally created mixture series, measured as percentage of red blood cells (RBC) infected

(DNA concentration of the test parasite was lowered 5-fold at each step in the series, while DNA of a competing parasite was kept constant (0·58% for *Plasmodium* and 1·875% for *Haemoproteus*). Control lines did not have DNA from the competing parasite. Results show amplification from the test lines of *Plasmodium* (P), *Haemoproteus* (H) or a mix of both (lowercases indicate that double peaks were small for the corresponding sequences), and results of PCR reactions (positive or negative amplification) from the corresponding samples in the control lines. Each column of results corresponds to an independent PCR reaction.)

	Plasmodium dilution (*Haemoproteus* constant)			*Haemoproteus* dilution (*Plasmodium* constant)			Controls	
Dilution	% RBC	PCR results		% RBC	PCR results		*Plasm.*	*Haem.*
1	0·58	P	P	1·875	P	P	(+) (+)	(+) (+)
2	0·116	P	P	0·375	P	P	(+) (+)	(+) (+)
3	0·023	P+h	P+H	0·075	P	P	(+) (+)	(+) (+)
4	0·0046	P+H	P+H	0·015	P	P	(+) (+)	(+) (+)
5	0·0009	p+H	p+H	0·003	P	P	(+) (+)	(+) (−)
6	0·00019	H	H	0·0006	P	P	(+) (+)	(+) (+)
7	0·000037	H	H	0·00012	P	P	(−) (−)	(−) (−)
8	$7{\cdot}42\times10^{-6}$	H	H	$2{\cdot}4\times10^{-5}$	P	P	(+) (−)	(−) (−)
9	$1{\cdot}48\times10^{-6}$	H	H	$4{\cdot}8\times10^{-6}$	P	P	(−) (−)	(−) (−)
10	$2{\cdot}97\times10^{-7}$	H	H	$9{\cdot}6\times10^{-7}$	P	P	(−) (−)	(−) (−)

conditions described above. From each individual, we sequenced 3 to 14 clones to identify the parasite lineages present in the mix.

Experimental multiple infections

We created experimental multiple infections by mixing DNA extracts obtained from two great reed warblers (*Acrocephalus arundinaceus*) for which parasite intensity and identity had been determined by microscope inspection of blood smears and DNA sequencing, respectively. We wanted to test (1) whether *Plasmodium* is preferentially amplified when mixing with *Haemoproteus* in a multiple infection, and (2) if varying proportions of the mixing parasites affect the risk of not scoring multiple infections due to PCR competition. Normally, *Plasmodium* infections have much lower intensities than *Haemoproteus* infections (Atkinson & van Riper, 1991; Waldenström *et al*. 2004; personal observations). To extend the range of possible intensities of infection of *Plasmodium* upwards, we selected a great reed warbler with an unusually high intensity of *Plasmodium* (1·16% parasitaemia, strain GRW2 of *P. nucleophilum*, GenBank AF254962; Bensch *et al*. 2000). As a representative of a *Haemoproteus* infection, we selected a great reed warbler infected by this genus at a normal intensity (3·75% parasitaemia, strain GRW1 of *H. sylvae*, GenBank AF254964; Bensch *et al*. 2000). It is worth mentioning that, while *Haemoproteus* only occurs in the bloodstream at the gametocytic stage, *Plasmodium* undergoes schizogony in blood, which should increase the number of copies of parasite DNA relative to number of infected erythrocytes (Fallis & Desser, 1977; van Riper *et al*. 1993). Given that we chose one sample with extraordinarily high intensity of *Plasmodium* and one with normal intensity of *Haemoproteus*, the ratio of original DNA copies should favour *Plasmodium* in the mix. Both original samples had an equal concentration of DNA of 25 ng/μl.

We set our artificial mixes in two 10-step dilution series (Table 1). At each step in the series, the concentration of one parasite (the test parasite) was diluted 5 times in a DNA extract (25 ng/μl) from a non-infected great reed warbler. Great reed warbler DNA was used instead of water to keep the same DNA concentration throughout the dilution series. Both dilution series, one with *Plasmodium* and the other with *Haemoproteus* as test parasites, were split into experimental and control lines. In the experimental lines, the concentration of the competing parasite was kept constant at a half of its natural intensity, by mixing equal volumes of its DNA extract and the corresponding DNA dilutions in the series. The control lines had a decreasing concentration of the test parasite parallel to the experimental lines; however, they did not include DNA from the competing parasite, but were mixed with an equivalent amount of DNA from a non-infected bird. When the test parasite was no longer amplified at a given step in the experimental dilution series, the control series allowed us to distinguish between competition and template shortage as the cause.

We used the same nested PCR method described above (Waldenström *et al*. 2004) to determine positive and negative amplifications from each step in the control and experimental dilution series. PCR products were sequenced and the electropherograms

https://doi.org/10.1017/S003118200500733X Published online by Cambridge University Press

were inspected for double peaks. All this procedure was done twice to check for reproducibility of the results.

RESULTS

Screening for natural multiple infections

Among 196 infected birds, 40 were suspected to harbour multiple infections according to the occurrence of double peaks on the sequence electropherograms (Fig. 1). Depending on the identity of parasites (from putative intraspecific strains to different genera) and the number of parasite lineages present in the mix, the number of double peaks on the electropherograms varied between one and more than 60. Among the parasites found, *Haemoproteus* lineages showed between 0·2% and 7% sequence divergence, and *Haemoproteus* lineages differed from *Plasmodium* lineages by 13% on average. Remarkably, double peaks had to be interpreted visually, as most often they were not distinguished from the background base signals, but were scored as resolved bases by automated sequence analysis (Fig. 1). Nine suspected mixes could not be corroborated by cloning of a second PCR product. In all these cases, the double peaks on the original electropherogram were small, and the main sequence was the only one retrieved by cloning of a second PCR product. Of the 31 samples that were satisfactorily resolved by cloning, only 10 samples showed small double peaks on the original electropherogram. The relationship between the size of double peaks and consistency of detection of multiple infections was statistically significant (Fisher exact test: $P=0{\cdot}0003$).

Most multiply infected birds (21 individuals) harboured two parasite strains, but there were 8 birds with 3, and 2 with 4 strains. Infections by different *Haemoproteus* parasites were the most common (22 birds), and mixes of *Plasmodium* strains were very rare (1 single case). Mixes involving the two genera were relatively common (8 birds). These frequencies fitted (Fig. 2; $\chi^2_{(2)}=4{\cdot}24$, $P=0{\cdot}12$) the proportions expected from random pairwise combinations of all the parasite strains found in the sample studied, given their frequencies in the infected population (we found 22 *Haemoproteus* and 5 *Plasmodium* strains in our sample, present in 0·05% to 40% of infected birds; unpublished data).

Normally, the parasite strains forming a multiple infection could be determined after sequencing 6 clones. A mix was considered to be resolved when the sequences retrieved by cloning could explain all double peaks observed on the original electropherogram. We sequenced more than 6 clones if required to resolve the mixed signal (1 sample with 3 parasites needed 14 clones to be fully resolved). However, in most cases this was not necessary; for example, of 8 mixes formed by 3 parasites, 6 could be resolved with 6 clones, and the two 4-parasite infections could be resolved after sequencing 6 and 8 clones, respectively.

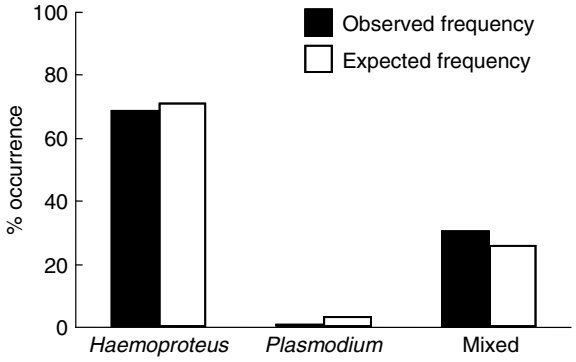

Fig. 2. Frequency of multiple infections involving 2 *Haemoproteus* strains, 2 *Plasmodium* strains, or a mix of both genera among 196 infected warblers. Expected frequencies were obtained by simulating 5000 random pairwise combinations from all lineages in the parasite community, taking into account their prevalences in the sample (see text for details).

Problems with TA-cloning

We encountered the two expected types of PCR errors: single nucleotide mutations and jumping PCR artefacts. However, these cases were the exception: of 241 clones obtained in total, 194 (80·5%) seemed free of PCR derived errors; they corresponded to known strains found in singly infected birds, or to strains that were found in multiple infections in different birds, or in different clones of the same bird. All these sequences were always necessary to explain the double peaks seen on the original electropherogram, an observation that confirmed their reliability.

PCR mutations were found in 37 clones (15·4%), but they were very easily identified as they occurred on random positions along the sequence, and did not match to the previously observed double peaks on the original electropherogram. Three clones (1·2%) seemed to mix true and mutated sequences, as their electropherograms had double peaks absent from the original electropherograms. Finally, 7 clones (2·9%) seemed to be a result of jumping PCR, arising from partially elongated fragments that were extended on the wrong template in the next PCR cycle, resulting in the combination of parts of two different true sequences into a new, chimerical sequence.

The 37 clones with mutations included a total of 49 substitutions (5 clones showed more than one substitution), of which 35 (71·5%) were transitions, 12 (24·5%) were transversions, and 2 (4%) were indels. None of the cloned PCR products showed consistent mutations (repetitions of the same substitution in different clones). The average mutation rate was 0·00042 substitutions per site per PCR reaction.

https://doi.org/10.1017/S003118200500733X Published online by Cambridge University Press

Artificial sequences resulting from jumping PCR were more problematic, as they sometimes matched to double peaks on the original electropherograms, and therefore were hard to distinguish from true parasite sequences. Faced with this problem, we decided to consider sequences to represent true parasite lineages only if their clones were necessary to explain previously observed double peaks, or if they had been obtained independently from different birds.

Detectability of multiple infections

Our analysis of experimental multiple infections produced similar results in both sets of experiments, apart from showing some inconsistent PCR amplifications from samples with low concentration of parasite DNA (Table 1). The sequences obtained showed double peaks at the variable sites distinguishing the two mixing parasite lineages. However, we were unable to detect all expected multiple infections (Table 1). When *Plasmodium* was the test parasite, we found no signal from *Haemoproteus* in the first two steps along the dilution series, despite a significant proportion of initial DNA copies belonged to this parasite (Table 1). Multiple infections were revealed by double peaks in the following 3 steps in the dilution series, when *Plasmodium* intensity was lowered from 0·023% infected erythrocytes to nearly 1 infected erythrocyte per 100 000 (Table 1). At this last step, close to the limit of detection, double peaks were small, and some of them were absent. Finally, *Haemoproteus* out-competed *Plasmodium* when the intensity of the latter was lowered to just 1 infected erythrocyte in a million (PCR failure was discarded by positive amplifications from control samples with the same concentration of parasite DNA; Table 1). Below this threshold, the test parasite was hardly amplified, as shown by frequent negative PCR for the control series (Table 1).

Consistent with a preferential amplification of *Plasmodium* at high intensities, we could not obtain the expected multiple infections from the *Haemoproteus* dilution line. The results for the control line were similar to the results obtained for *Plasmodium*, with a threshold of possible amplification located at concentration around 1 infected erythrocyte in a million (Table 1).

DISCUSSION

Our results show the general utility of electropherogram inspections and TA-cloning when screening for multiple infections of avian malarial parasites. Given the high resolution of current sequencing analysis techniques, variable nucleotides (double peaks) can be easily distinguished from background on sequence electropherograms, even if they are small. On most occasions, putative multiple infections can be resolved by TA-cloning. In addition, the correspondence between the base differences between sequences of clones and the double peaks previously seen on original electropherograms acts as a control for sequence reliability. This is particularly important when new parasite strains are found in multiple infections. Remarkably, of the 27 parasite strains identified in our study of blackcaps and garden warblers, 13 (nearly 50%) were found in multiple infections only, even though most of them were retrieved from different individual hosts (this was a consequence of the occurrence of some common lineages at very high prevalences, unpublished data). In most cases, parasite mixes were satisfactorily resolved after sequencing 6 clones, even when 3 or 4 parasites coexisted in the same host. This seems a reasonable amount of time and resources required to implement the technique.

The technique presented here allows identification of mixes of parasites with any degree of relatedness, from intraspecific variants (such as parasite strains distinguished by just 1 nucleotide difference in the sequence of interest) to distantly related taxa, such as different parasite genera. The possibility to identify multiple parasite genotypes within the same host opens an important research area in bird-parasite interactions, including a much better understanding of the diversity and structure of parasite communities, as well as the possibility to study the fitness effects and parasite interactions derived from co-infections involving intraspecific parasite strains or very closely related parasite species (de Roode *et al*. 2004).

Diagnosis of genetically mixed infections

The detection of actual multiple infections was sometimes difficult due to two problems. First, we found evidence that, at least with the PCR method used (Waldenström *et al*. 2004) and with extremely high intensities of *Plasmodium*, this parasite might be preferentially amplified when coexisting with *Haemoproteus*. Our experimental dilution lines revealed this potential problem when the intensity of *Plasmodium* was high (0·116% parasitaemia or higher). In the great reed warbler population from which samples were obtained, *Plasmodium* infections are normally much milder: the clear majority have a parasitaemia lower than 0·01%, which lies within the range where *Plasmodium* and *Haemoproteus* were co-amplified in our experiment (0·001–0·02% parasitemia). In fact, in the natural bird populations studied, the frequency of multiple infections involving both parasite genera was similar to the frequency expected from the prevalence of the different lineages in the parasite community (Fig. 2). Therefore, the probability of missing such co-infections due to preferential amplification of

https://doi.org/10.1017/S003118200500733X Published online by Cambridge University Press

Plasmodium, is probably very low in most natural situations. The second expected problem was competition between parasite strains when they occurred at very different intensities. In our experiment, the test parasite was co-amplified even at an intensity of around 1 infected erythrocyte per 100 000. However, at such low intensities double peaks were small, and some of them disappeared. Eventually, multiple infections were no longer detected when intensity was lowered to around 1 infected erythrocyte in a million.

Taking these experimental results together, we conclude that different modes of PCR competition (higher primer affinity to *Plasmodium* or extremely imbalanced amounts of template) might prevent detection of multiple infections in a few natural situations. Based on both natural and experimentally created multiple infections, the presence of small double peaks on the electropherogram may indicate that one of the sequences was less favoured in the PCR. In some cases, inconsistent PCR results can be obtained, switching from positive (double peaks) to negative between tests, an incident encountered in 25% of multiple infections in our sample. In fact, extremely biased proportions of mixing parasites are expected, for example if one parasite is being cleared out from the bloodstream due to competition with other co-infecting parasites (de Roode *et al*. 2004). Therefore, special attention should be paid even to the smallest double peaks on electropherograms during parasite sequence analyses, for example checking their correspondence to variable sites distinguishing among already known parasite lineages. These cases are ecologically relevant if we bear in mind that the presence of one parasite in the bloodstream, whatever its abundance, indicates a successful infection event. In addition, it would be interesting to know whether the presence of a parasite at very low numbers has fitness consequences for a host that is more acutely infected by other, closely related parasites.

Resolving multiple infections

Apart from the aforementioned problems to detect some multiple infections, the cases that were clearly revealed by double peaks presented further difficulties due to the presence of artificial sequences among cloned products. Single nucleotide mutations were relatively common, but in all cases they could be identified by a lack of correspondence between the variable site and the double peaks observed on the first, independently obtained sequence electropherogram. The pattern of substitutions, with a normal transition bias (Wakeley, 1996) and absence of consistent mutations (i.e., mutations incorporated to all clones obtained from 1 PCR product), suggests that the initial number of DNA copies was not too small, and the DNA template was not degraded (Hofreiter *et al*. 2001). Therefore, mutations most likely resulted from DNA polymerase errors during the course of an otherwise normal PCR reaction.

Artefacts resulting from jumping PCR were relatively rare, but created more difficulties as sometimes they matched to double peaks. Exceptionally, this prevented us to distinguish between sequences of parasite strains and artificial sequences. When faced with such a situation, we recommend that any new sequence should be trusted only after it has been found in more than one host, and repeat the cloning in search for candidate strains when any suspicious result is obtained.

An important feature of the technique presented here is that it will produce a conservative estimate of the proportion of multiple infections in a population, as a fraction of actual multiple infections (around 25% of the multiple infections initially scored in our sample) will inevitably be scored as single infections. The technique also provides a conservative estimate of the number of parasites forming a multiple infection. First, one should not assign new sequences as belonging to novel strains unless they are needed to explain double peaks on the original electropherogram. In some cases, however, more sequences than needed to explain double peaks are obtained by cloning, of which some may be artefacts and some may belong to parasite strains. For example, mixes can be composed of strains with overlapping variable sites, which may be hidden under the same double peaks but can be resolved by cloning. We did not find any evidence of this problem, but examples could be made from existing parasite lineages, and hence it should not be disregarded. Finally, our method introduces an asymmetry between scoring presence and absence of a multiple infection, or of a given parasite in a positively identified mix. While presences are always reliable, there is a chance that absences are due to failure in the detection method, a potential drawback of any study based on presence-absence data (Gu & Swihart, 2004).

Further recommendations

In summary, visual inspection of sequence electropherograms combined with TA-cloning allows screening for multiple infections with adequate resolution. Some parasite lineages may be first discovered in multiple infections, and methods like the one presented here are needed to add them up to the parasite community during population surveys. If mixed sequences are to be resolved by cloning, it is highly recommended that cloned products be obtained from an independent PCR, in order to allow distinguishing mutations and jumping PCR artefacts by comparison with the original sequence electropherogram. The incidence of PCR artefacts could be reduced by use of high-fidelity enzymes

https://doi.org/10.1017/S003118200500733X Published online by Cambridge University Press

(particularly for preventing nucleotide substitutions), or longer extension times in the PCR prior to cloning. However, if variable sites among cloned sequences are checked for correspondence with double peaks in the original electropherogram, standard *Taq* DNA polymerase shows a suitable performance and allows reduction of economic costs.

Special attention has to be paid to the correspondence between cloned sequences and double peaks on an independently obtained electropherogram. However, in some instances part of the variable sites may not be revealed by double peaks. This was shown by our experiment, when the concentration of the test parasite was near the detection threshold. We also found some such cases in natural multiple infections. In these cases, the electropherograms did not show double peaks at all the variable sites distinguishing the mixing parasites, but these were repeatedly cloned and identified as common strains in single infections. Missing double peaks are particularly (but not necessarily) frequent at the end of the sequence, so sequencing from both ends usually helps to circumvent this problem. When these uncommon situations involve strains that are already known in the parasite community, the sequences obtained can be safely accepted as belonging to different parasite lineages. However, if the cloning retrieves unknown sequences, repeating the analysis is essential to substantiate their reliability. Following these reasonably prudent criteria, we believe that this technique to screen for multiple infections will significantly contribute to our understanding of the ecological and evolutionary relationships between malarial parasites and their avian hosts.

We thank Dennis Hasselquist, Olof Hellgren and Jonas Waldenström for helpful ideas and discussion. J. Waldenström and two anonymous referees also improved an early version of the paper. J. P.-T. was supported by a Marie Curie Fellowship of the European Community, under contract number HPMF-CT-2002-02096.

REFERENCES

ATKINSON, C. T. & VAN RIPER, C. III (1991). Pathogenicity and epizootiology of avian haematozoa: *Plasmodium*, *Leucocytozoon* and *Haemoproteus*. In *Bird-Parasite Interactions: Ecology, Evolution and Behaviour* (ed. Loye, J. E. & Zuk, M.), pp. 19–48. Oxford University Press, Oxford.

BENSCH, S., STJERNMAN, M., HASSELQUIST, D., ÖSTMAN, Ö., 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. *Proceedings of the Royal Society of London, B* **267**, 1583–1589.

BENSCH, S., PÉREZ-TRIS, J., WALDENSTRÖM, J. & HELLGREN, O. (2004). Linkage between nuclear and mitochondrial DNA sequences in avian malaria parasites – multiple cases of cryptic speciation? *Evolution* **58**, 1617–1621.

BRUCE, M. C., DONNELLY, C. A., ALPERS, M. P., GALINSKI, M. R., BARNWELL, J. W., WALLIKER, D. & DAY, K. P. (2000). Cross-species interactions between malaria parasites in humans. *Science* **287**, 845–848.

CHEESMAN, S. J., DE ROODE, J. C., READ, A. F. & CARTER, R. (2003). Real-time quantitative PCR for analysis of genetically mixed infections of malaria parasites: technique validation and applications. *Molecular and Biochemical Parasitology* **131**, 83–91.

DE ROODE, J. C., CULLETON, R., CHEESMAN, S. J., CARTER, R. & READ, A. F. (2004). Host heterogeneity is a determinant of competitive exclusion or coexistence in genetically diverse malaria infections. *Proceedings of the Royal Society of London, B* **271**, 1073–1080.

FALLIS, A. M. & DESSER, S. (1977). On species of *Leucocytozoon, Haemoproteus*, and *Hepatocystis*. In *Parasitic Protozoa. Vol. 3* (ed. Kreier, J. P.), pp. 239–266. Academic Press, New York.

FALLON, S. M., RICKLEFS, R. E., SWANSON, B. L. & BERMINGHAM, E. (2003). Detecting avian malaria: an improved PCR diagnostic. *Journal of Parasitology* **89**, 1044–1047.

GU, W. D. & SWIHART, R. K. (2004). Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. *Biological Conservation* **116**, 195–203.

HALL, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleics Acids Symposium Series* **41**, 95–98.

HELLGREN, O., WALDENSTRÖM, J. & BENSCH, S. (2004). A new PCR assay for simultaneous studies of *Leucocytozoon, Plasmodium*, and *Haemoproteus* from avian blood. *Journal of Parasitology* **90**, 797–802.

HOFREITER, M., JAENICKE, V., SERRE, D., VON HAESELER, A. & PÄÄBO, S. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Research* **29**, 4793–4799.

LELLO, J., BOAG, B., FENTON, A., STEVENSON, I. R. & HUDSON, P. J. (2004). Competition and mutualism among the gut helminths of a mammalian host. *Nature, London* **428**, 840–844.

PAUL, R. E. L., NU, V. A. T., KRETTLI, A. U. & BREY, P. T. (2002). Interspecific competition during transmission of two sympatric malaria parasite species to the mosquito vector. *Proceedings of the Royal Society of London, B* **269**, 2551–2557.

PERANDIN, F., MANCA, N., CALDERARO, A., PICCOLO, G., GALATI, L., RICCI, L., MEDICI, M. C., ARCANGELETTI, M. C., SNOUNOU, G., DETTORI, G. & CHEZZI, C. (2004). Development of a Real-Time PCR assay for detection of *Plasmodium falciparum, Plasmodium vivax*, and *Plasmodium ovale* for routine clinical diagnosis. *Journal of Clinical Microbiology* **42**, 1214–1219.

PERKINS, S. L. (2000). Species concepts and malaria parasites: detecting a cryptic species of *Plasmodium*. *Proceedings of the Royal Society of London, B* **267**, 2345–2350.

POULIN, R. (1997). Species richness of parasite assemblages: evolution and patterns. *Annual Review of Ecology and Systematics* **28**, 341–358.

READ, A. F. & TAYLOR, L. H. (2001). The ecology of genetically diverse infections. *Science* **292**, 1099–1102.

RICHARD, F. A., SEHGAL, R. N. M., JONES, H. I. & SMITH, T. B. (2002). A comparative analysis of PCR-based detection methods for avian malaria. *Journal of Parasitology* **88**, 819–822.

RICKLEFS, R. E., FALLON, S. M. & BERMINGHAM, E. (2004). Evolutionary relationships, cospeciation, and host switching in avian malaria parasites. *Systematic Biology* **53**, 111–119.

VAN BAALEN, M. & SABELIS, M. W. (1995). The dynamics of multiple infection and the evolution of virulence. *American Naturalist* **146**, 881–910.

VAN RIPER, C. III, ATKINSON, C. T. & SEED, T. M. (1993). Plasmodia of birds. In *Parasitic protozoa*. *Vol*. 7. (ed. Kreier, J. P.), pp. 73–140. Academic Press, New York.

WAKELEY, J. (1996). The excess of transitions among nucleotide substitutions: new methods of estimation transition bias underscore its significance. *Trends in Ecology and Evolution* **11**, 158–163.

WALDENSTRÖM, J., BENSCH, S., KIBOI, S., HASSELQUIST, D. & OTTOSSON, U. (2002). Cross-species infection of blood parasites between resident and migratory songbirds in Africa. *Molecular Ecology* **11**, 1545–1554.

WALDENSTRÖM, J., BENSCH, S., HASSELQUIST, D. & ÖSTMAN, Ö. (2004). A new nested polymerase chain reaction method very efficient in detecting *Plasmodium* and *Haemoproteus* infections from avian blood. *Journal of Parasitology* **90**, 191–194.

YAMASAKI, H., ALLAN, J. C., SATO, M. O., NAKAO, M., SAKO, Y., NAKAYA, K., QIU, D., MAMUTI, W., CRAIG, P. S. & ITO, A. (2004). DNA differential diagnosis of taeniasis and cysticercosis by multiplex PCR. *Journal of Clinical Microbiology* **42**, 548–553.

https://doi.org/10.1017/S003118200500733X Published online by Cambridge University Press