Molecular Genetic Analysis of a Brown-headed Cowbird (*Molothrus ater*) Population.

Molecular Genetic Analysis of a Brown-headed Cowbird (*Molothrus ater*) Population.

Paul Miller, B.Sc.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Master of Science
McMaster University

MASTER OF SCIENCE (1993) McMASTER UNIVERSITY

(Biology) Hamilton, Ontario

TITLE: Molecular Genetic Analysis of a Brown-headed Cowbird (*Molothrus ater*) Population.

AUTHOR: Paul Christopher Miller, B. Sc. (McMaster University)

SUPERVISOR: Dr. H. Lisle Gibbs

NUMBER OF PAGES: vi, 88

ACKNOWLEDGEMENTS

Many thanks to the numerous people who provided guidance, assistance, constructive criticism and friendship over the past two years. Dr. Sealey and his students Glen, Paula, Marleen, Dianne and Sharon and all the people at the University of Manitoba Field Station, Russ, Dr. Robinson, Mosquito Bob and Little Skeeter but especially Doreen and the rest of the cooks, were exceedingly gracious hosts and knowledgeable collaborators. Thanks also to my colleagues at MAC, especially my lab mate Cindy, Dan 'RAPIDDIPDOTDAT' Fieldhouse and my supervisor Dr. H. Lisle Gibbs for his enthusiasm and attentiveness (it's been a learning experience for both of us). I would like to thank my family for putting up with my infrequent appearances at home and supporting all of my academic interests. And finally to Emma, thanks for choosing McMaster!

TABLE OF CONTENTS

Chapter Two: Genetic Analysis of Host Specificity in the Brown-headed Cowbird (*Molothrus ater*): Results from MtDNA and RAPD Nuclear Markers.

LIST OF TABLES and FIGURES

Chapter One

Chapter 2

CHAPTER 1

Inter- and intraspecific variation in the mitochondrial DNA (mtDNA) control region of a Passerine bird: Implications of the deletion of the left hypervariable domain.

ABSTRACT

The mtDNA control region of the Brown-headed Cowbird (Molothrus ater) was sequenced and comparisons made at the inter- and intraspecific level. Comparison of the control region with that of another Passerine, Darwin's Finch (Geospiza scandens), revealed a high degree of both gross and fine scale structural similarity. At the nucleotide level, this comparison confirmed the presence of a hypervariable domain which evolves at rate approximately 5 times faster than coding mtDNA as well as a relatively conserved central domain which evolves at rate comparable to coding mtDNA. Both species displayed the typical avian mtDNA gene organisation previously described by Desjardins and Morais (1990, 1991) and Quinn and Wilson (in press). However, the most notable structural feature in common was

the apparent deletion of the entire left hypervariable domain (CR1). At a finer scale, Conserved Sequence Block (CSB1) was perfectly conserved between cowbird and finch and Conserved Sequence Block 2 (CSB2) was 78% similar. The hypervariable right domain showed the largest degree of sequence divergence between species, 22.7%, while the central domain and phe-tRNA showed much less divergence, 6.47 and 4.41% respectively. At an intraspecific level, in 524 bases of sequence from 31 nestling cowbirds from a population at Delta, Manitoba, only 3 variable sites were detected which defined a total of 4 haplotypes. The average percent sequence divergence for this population was 0.27%. This level of variation within the cowbird population is low compared to other vertebrate populations. This relative lack of variation is largely attributable to the loss of the left hypervariable domain (CR1). The loss of CR1 will limit the control region's usefulness for high resolution population level studies but may make it a useful marker for phylogenetic studies within the class Aves.

INTRODUCTION

Mitochondrial DNA (mtDNA) is emerging as the molecule of choice in studies of phylogeographic variation (Ball et al 1988, Thomas et al 1990, Wenink et al 1992, Edwards 1993) and for the construction of phylogenies of closely related taxa (Meyer et al 1990, Smith and Patton 1991, Ball et al 1992, Shedlock et al 1992). Several reviews (Wilson et al 1985, Avise et al. 1987, Moritz et al 1987, Harrison 1989, Kocher et al 1989, Avise 1991) have highlighted the features of mtDNA that make it appropriate for such studies: (i) the mtDNA molecule is haploid, uniparentally inherited and essentially non-recombining. Therefore, mutations arising within the molecule are, for the most part, faithfully recorded in history and consequently allow researchers to build phylogenies of related molecules; (ii) on the whole, the mtDNA genome evolves at a rate approximately 5 to 10 times faster than single copy nuclear DNA (Brown et al 1979) and hence is more likely to detect variation between closely related taxa. However, rates of sequence evolution within the mtDNA molecule are not homogeneous. The non-coding control region in humans, for example, is estimated to evolve at a rate approximately 5 times faster than the rest of the mtDNA genome (Greenberg et al. 1983, Cann et al. 1987) and

therefore promises an even finer scale of resolution, capable of detecting variation within species and possibly within populations.

The control region, or D-loop, is the major non-coding region in the mtDNA genome of vertebrates and shows little sequence homology across broad taxa except for a few small, conserved blocks (Walberg and Clayton 1981, Saccone et al. 1987). Such conserved sequence blocks (CSBs) within the control region, are believed to possess the sequences which control initiation (Anderson et al 1981) and termination (Doda et al 1981) of mtDNA synthesis as well as transcriptional promoter sequences for both heavy and light strands (Cantatore and Attardi 1980).

Like the mtDNA molecule, the control region is itself heterogeneous and can be divided into three distinct regions with respect to nucleotide base composition and rate of sequence evolution. Two hypervariable domains, Control Region 1 (CR1) and Control Region 2 (CR2) are located at the 5′ and 3′ ends of the light strand, respectively, and flank a central, guanine rich (in the light strand) conserved region (Greenberg et al 1983).

The advent of rapid sequencing protocols using PCR technology with 'universal' primers (Kocher et al 1989) has produced a wealth of information about sequence variation in the control region of a variety of vertebrate taxa. The two hypervariable regions, especially CR1, have proven useful in phylogeographic studies of mammals (Thomas et al 1990, Vigilant et al 1991, 1992), fish (Bernatchez et al 1992, Brown et al 1993), and birds (Quinn 1992, Wenink et al 1993, Edwards 1993) and also in the production of phylogenies between closely related fish species (Meyer et al 1990, Shedlock et al 1992).

Unfortunately, the available universal primers have failed to amplify control region sequence from many Passerine bird species (Gelter et al 1993). Upon cloning and sequencing of the control region in three Darwin's Finch species (Geospiza sp.), Gelter et al (1993) discovered that virtually all of the left domain was missing. This purported deletion has reduced the size of the finch control region to 646 bases, which is small compared to the control regions of most other vertebrates which generally exceed 1 kb in size (Brown 1985, Desjardins and Morais 1991 and 1992, Saccone et al 1991).

Subsequent amplification of several other Passerine species' control region, with external primers, yielded amplification products varying in size from 700 to 800 bases which suggests that the deletion maybe widespread among this large, diverse group of birds (Gelter et al 1993). To date the only other Passerine control region sequenced, that of the Grey-crowned Babbler (Pomatostomas temporalus), possesses a typical vertebrate organization, including both hypervariable regions (Edwards 1993). Therefore, establishing the generality of the "finch" control region structure awaits sequence information from a wider range of Passerines.

Gelter et al (1993) speculated that the loss of such a major source of variation in the control region of Passerines would likely limit it's usefulness for high resolution population level studies within these birds. However, whether this is the case is unknown because information about intraspecific levels of variation in bird species with small control regions is currently unavailable.

To assess this possibility I have surveyed control region variation in a population of brood parasitic Brown-headed Cowbirds (Molothrus ater) as part of a larger study

investigating the potential for host specificity to contribute to the long-term genetic structure of this population (Miller and Gibbs, in prep.). I have used the sequence generated by this study to: (i) assess the similarity of structural features between the control region of the Brown-headed cowbird and Darwin's Finch (Geospiza scandens); and (ii) investigate the usefulness of the control region, as a marker for studies of intraspecific variation, in a species missing the left hypervariable region (CR1).

My results suggest that the structural features highlighted by Gelter et al (1993) in Darwin's Finch are highly conserved in the Brown-headed Cowbird and also that there is a reduction in control region variation, within this population, compared with other populations of vertebrate species with an intact CR1.

MATERIALS AND METHODS

Samples: Brown-headed Cowbirds are brood parasites which lay their eggs in other 'host' species nests. Host parents that accept cowbird eggs then incubate them and raise the young parasite. Cowbird nestlings were sampled from two of their most abundant host species, the Yellow Warbler (Dendroica petechia) (YW) and the Red-winged Blackbird (Agelaius phoeniceus) (RWB), at the University of Manitoba Field Station at Delta, Manitoba from May-July 1992. The majority of nestling's sampled came from host nests separated by at least 200 m. However, in one case two samples came from the same host nest and in a few others the nests were less than 100 m apart. For a more complete description of the sample locations and a discussion of the implications for sample independance see Miller and Gibbs (in prep.).

In total, blood and/or tissue samples were collected from 19 nestlings from YW nests and 13 from RWB nests. Blood samples (15-75 ul) were stored in 800 ul of 1 x lysis buffer (4 M urea, 0.2 M NaCl, 0.1 M Tris-HCl [pH 8.0], 0.5 % n-laurylsarcosine, 10 mM EDTA, [Seutin et al 1991]) at 4 °C. Tissue samples (approx. 1 g), from unhatched cowbird eggs and dead nestlings, were stored in 15 ml sterile tubes at -20 °C until further processing.

DNA Extraction: Half of the volume of each blood sample was suspended in 15 ml sterile tubes with 1 x lysis buffer to a total volume of 4 ml. Frozen muscle samples (approximately 0.1 g), along with 4 ml of 1 x lysis buffer, were mechanically ground for approximately 15 minutes using a mortar and pestle with liquid nitrogen. The resultant powdered muscle samples were collected into sterile 15 ml tubes. The samples were rocked gently at 37 °C overnight or until the solution was translucent and homogeneous. The samples were then digested with 72 units of proteinase-K at 37 °C for at least 12 hours. DNA was extracted twice with phenol and chloroform (70:30) and once with chloroform. The DNA was precipitated by adding sodium acetate to a final concentration of 0.15 M and two volumes of 95 % cold ethanol. The DNA was spooled out of the solution using a sterile pasteur pipette, washed in 70 % ethanol and allowed to air dry for five minutes. Once dried, the DNA was dissolved in 0.3-1.0 ml TNE2 (10 mM Tris-HCl, 10 mM NaCl, 2 mM EDTA, pH 8.0) and gently rocked overnight at 37 °C and subsequently stored at -20 °C.

D-loop Amplification: The primers GSL-glu and GSH-12S (Gelter et al 1993), designed from Darwin's Finch mtDNA, were used in PCR amplifications of the Brown-headed Cowbird control region (see Fig.1). In total, 31 nestlings from the

Delta site were amplified. Amplifications were carried out in 50 ul total volumes with 50 ng of whole genomic DNA template and in the presence of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2 mM MgCl2, 0.001 % gelatin, 200 uM each dNTP, 0.5 uM of each primer, and 2.5 units of Taq polymerase. Two drops of mineral oil were added to the top of the reaction mixture which was then heated to 94 °C for five minutes prior to 35 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and extension at 72 °C for 1.5 min. Negative controls (all reagents except distilled water was substituted for template DNA) were included in each amplification to evaluate the possibility of contamination. Ten microlitres of the product was then run on a 2 % low-melt agarose gel, and visualized by staining with ethidium bromide. The entire amplified band was excised and then dissolved in 1 ml of distilled water by heating at 65 °C for 5 min. and then stored at -20 °C.

RFLP survey: Prior to sequencing, an RFLP survey was carried out with nine restriction enzymes (Alu I, Cla I, Dra I, Hae III, Hinf I, Mbo I, Msp I, Rsa I and Taq I) to assess their potential for detecting intraspecific variation within the control region. Five to ten microlitres of PCR amplified D-loop products were digested with three units of enzyme, following the manufacturer's suggestions. Restriction digest

products were run on 1.0 % agarose 0.5 x TBE gels until the Orange-G loading dye had run at least 10 cm. The gels were then stained with Ethidium Bromide, visualised under uv light and photographed. Only one polymorphic restriction site, a Taq I site, was found within the population surveyed. I therefore decided to sequence the control region for each individual in an attempt to detect greater levels of intraspecific variation in this region.

Sequencing

Reference Sequence from mtDNA: I amplified the control region, as above, from a purified mitochondrial sample extracted from a cowbird from California (supplied by R.C. Fleischer) to obtain a reference sequence that could be used to check for the presence of a nuclear homologue of this region (cf. Quinn 1992). This fragment was then cloned into a plasmid vector following the procedure outlined by Marchuk et al (1991) and sequenced using standard double-stranded sequencing protocols (Sambrook et al 1989). I later confirmed most of the cloned sequence by direct sequencing of the control region PCR product amplified from purified mtDNA using the procedures described below (see below for areas not directly sequenced).

Sequence alignment: The cowbird reference sequence was aligned with the Darwin's Finch control region using the alignment subroutine of Seqaid II v3.5 (Rhoads and Roufa 1989) with a k-tuple of 5. The k-tuple values, which range from 1-6, specify the number of contiguous bases which must match between the two sequences in order to be scored as a hit. Large k-tuples produce longer stretches of continuous alignment with fewer small gaps but require further editing when large deletions are present. Most of the right domain required further manual alignment with ESEE (Cabot and Beckenbach 1989) due to the large number of insertion\deletion events. When the two sequences were misaligned and gaps were introduced, preference was given to fewer large gaps over more small gaps. Each gap, regardless of size, was treated as one mutational event in later analyses.

Template Preparation for Genomic DNA Samples: To produce single stranded template for sequencing from the genomic samples, asymmetric PCR amplifications (Gyllensten and Erlich 1988) were carried out as descrbed above with the following modifications: the reaction volume was raised to 100 ul, 2 ul of the gel purified double stranded D-loop solution was used as the template and the light strand primer, GSL-glu, was diluted one hundred fold to 5 pM. After

amplification, asymmetrical PCR products were centrifugally dialysed three times for eight minutes at 7000 rpm using Ultra-Free-MC (Millipore) centrifuge dialysis tubes rinsing the pellet each time with 300 ul of distilled deionized water. The purified products were then resuspended in 25 ul of distilled, deionized, H_2O and stored at -20 °C. After three attempts, no single strand template was produced for one individual (Nestling 21). Therefore, the total number of sequences analysed was 31 instead of 32, with 18 from YW nests and 13 from RWB nests.

Sequencing: Dideoxy sequencing (Sanger, Nicklen & Coulson 1977) with Sequenase 2.0 kits (USB) was carried out following the manufacturer's recommendations with the following modifications: the labelling mix was diluted to 1:20 and the labelling reaction was incubated for less than 1 min. Sequencing reaction products were resolved on 8 % polyacrylamide gels, vacuum dried with an aspirator for 2 h at 80 °C and then autoradiographed on Kodak X-\Omat AR film for at least 48 h.

To sequence the entire control region, three separate sequencing reactions were carried out per individual. These involved sequencing same single stranded template with three light strand primers, the terminal

primer GSL-glu and two internal primers MAL-dl1 and MAL-dl2 (Fig. 1). Initially, three hundred and fifty-one bases were read beginning 40 bases into the control region's central domain and extending 50 bases into the right domain to base 390. This sequence was followed by a 76 base gap and then a further 241 bases encompassing the remainder of the right domain, and all of the phe-tRNA (Fig. 1). Based on comparison of the cloned cowbird sequence to that of Darwin's Finch (Fig. 2), the control region was determined to be 639 bases long from glu-tRNA to phe-tRNA. Thus, I sequenced 301 of 340 bases in the central domain, 223 of the 299 base right domain and all of the 68 base phe-tRNA for all individuals (Fig. 2).

Shortly after entering the right domain the sequence became unreadable following a long poly(T) region (13 bases in total in the cloned cowbird sequence) beginning at base 391 (Fig. 2). I hypothesized that the Taq polymerase may have been slipping in this area, thereby producing multiple PCR products differing in size by one to a few bases (Chen, Kuang and Lee 1991). To test this hypothesis and to survey for potential size variation in this region, a 162 base fragment spanning bases 283 to 444 which contained this region was amplified from 20 individuals as described above except, the light strand internal primer, MAL-dl1 was end

labelled with P-32 and an internal heavy strand primer, MAH-d13, was used in place of GSH-12S (cf. Ellegren 1992). The PCR products were resolved on 8 % polyacrylamide gels, using sequenced M13 as a size standard, vacuum dried with an aspirator for 2 h at 80 °C and then autoradiographed on Kodak X-\Omat AR film overnight.

Each individual's profile displayed multiple bands each differing in size by a single base. The largest and presumably 'correct' band (Ellegren 1992), in all of the individuals surveyed, was 162 bases. Together these observations suggest that (i) the sequencing problem I experienced was likely due to an artefact during the initial amplification and also (ii) size variation in this region is unlikely.

RESULTS AND DISCUSSION

Control Region Structure

I compared the structural features of two Passerine control region sequences currently available, the Brown-headed Cowbird and Darwin's Finch. Alignment of the cowbird reference sequence with the finch control region revealed a high degree of similarity with respect to both gross and fine scale structure (Fig. 2).

The cowbird and finch sequence share the characteristic gene order found within the class Aves; i.e. from the 5' end of the light strand: glu-tRNA, control region, phe-tRNA, 12s-rRNA (Desjardins and Morais 1990). The two control regions were also comparable in size (639 bases in the cowbird versus 646 in the finch) with the reduction in size resulting from an apparent large scale deletion of the entire left domain. However, the left domain is present in Galliform birds, such as the Common Chicken (Gallus gallus) (Desjardins and Morais 1990) and the Japanese Quail (Coturnix japonicus) (Desjardins and Morais 1991), an Anseriform, the Lesser Snow Goose (Anser caerulescens caerulescens) (Quinn and Wilson 1993) and a Passeriform, the Grey-crowned Babbler (Edwards 1993).

Such large scale mutations in the mtDNA genome are potentially useful as taxonomic markers. For example, the mtDNA gene rearrangement in birds (Desjardins and Morais 1990) and tRNA rearrangement in marsupials (Paabo et al 1991), characterise their taxa and differentiate them from other vertebrates. Our result confirms, through sequence analysis, the suggestion of Gelter et al (1993) that the left domain of the control region is absent in other Passerine birds. However, at least one Passerine bird, the Grey-crowned Babbler does not share this deletion (Edwards 1993). The deletion of CR1 is therefore unlikely to have occured in a common ancestor to all Passerines and thus may prove to be a valuable marker for delineating genealogies within this group.

At a finer scale of resolution, I looked for the presence of sequence blocks identified as being conserved across several taxa. Such conserved sequence blocks (CSBs) are believed to possess the control sequences for initiation (Anderson et al 1981) and termination (Doda et al 1981) of mtDNA synthesis as well as transcriptional promoter sequences for both heavy and light strands (Cantatore and Attardi 1980). Of the three CSBs identified in mammals by Walberg and Clayton (1981), only CSB-1 and CSB-3 are present in birds (Gelter et al 1993). CSB-1 shows a high degree of

sequence similarity (85%) between Darwin's Finch and the common Chicken and Japanese Quail (Gelter et al 1993) and is perfectly conserved between the cowbird and finch. CSB-3 is present as only a remnant in birds and is located near the 3' of the right domain (Fig. 2). This region shows only 76% similarity between cowbird and finch. The extreme conservation of CSB-1, across a wide range of taxa, argues in favour of the suggestion that it possesses important control functions and in addition, make it an ideal candidate for building a universal internal primer for amplification and sequencing of this region (Gelter et al 1993).

The termination associated sequences (TAS) defined by Doda et al (1981) are normally located in the left domain of the control region and hence are missing in both the cowbird and the finch. However, Gelter et al (1993) report the presence of homologous sequences within the glu-tRNA which may serve the same function as the missing TAS sequences.

Nucleotide Variation

Interspecific Variation: Percent sequence divergence (%sd) between the cowbird and the finch was 14.1% for the entire control region, 6.47% for the central domain, 22.7% for the

right domain and only 4.41% for the phe-tRNA. If the %sd of the phe-tRNA is assumed to be characteristic of the remainder of the mtDNA genome, then it can be used to calibrate the rate of divergence of the cowbird control region in this population. Sheilds and Wilson (1987) have estimated the rate of divergence of the goose mtDNA genome to be approximately 1.3 %sd / 10^6 years. Since the control region is 3.2 times as variable as the phe-tRNA, it's rate of evolution is approximately 4.16 %sd / 10^6 years. Similarly, the hypervariable right domain is evolving at a rate of 6.7 %sd / 10^6 years or approximately 5.15 times as fast as the rest of the cowbird mtDNA genome.

The central domain of Passerines, like that of other vertebrates (Greenberg 1983), evolves at about the same rate as mtDNA coding genes, suggesting that this region is under strong primary structural constraint (Brown et al 1986). The reason for this constraint is unknown since no definite function has been attributed to this region.

The pattern of nucleotide change was very different between these three regions. A strong bias towards transitional base substitutions over transversions was observed within both the central domain (14 of 18) and the phe-tRNA (3 of 3). Four small insertion\deletion events also

occured within the central domain, three of which were in the first 23 bases of the cowbird sequence, none were observed in the phe-tRNA. These trends toward transitional bias and a relative lack of insertion\deletion events (compared to the right domain) have also been documented in interspecific comparisons of mtDNA protein genes in rodents and primates (Brown et al 1986).

The right domain showed no such transitional bias; in fact, nearly equal numbers of transitions (25) and transversions (29) were observed. Insertion\deletion events were much more frequent in the right domain (14), most of which were 1 to 2 bases in length. These trends were also observed in the control region of rodents (Brown et al 1986).

Taken together, the high rate of sequence divergence, frequent insertion\deletion events, and lack of a transitional substitution bias in the right domain, both point toward a lack of constraint on the primary structure of this region as a whole.

Intraspecific Variation: Very little sequence variability was detected among the 31 cowbird nestlings surveyed. Of the 524 bases of control region that was sequenced, only three

variable nucleotide sites were detected which defined four haplotypes (Table 1). There was one transition mutation within the central domain at position 187, a second transition and an insertion\deletion event were located within the right domain at positions 526 and 600 respectively.

The average, pair-wise, percent sequence divergence for the central and right domains were 0.09% and 0.27% respectively and 0.17% for the entire control region. The magnitude of the difference in within population variablity detected in the central and hypervariable right domains, approximately 1:3, was roughly equivalent to the interspecfic difference in variability between these two regions (6.47% in the central domain vs 22.7% in the right domain).

A comparison of this cowbird population's control region variation with that of other vertebrate populations (see Table 2) yielded the following insights. (i) The level of variation detected by the control region was substantially less in the cowbird population. The most obvious reason for this reduced variability is that the hypervariable left domain, CR1, is missing. (ii) CR2 is consistently less variable than CR1, in studies where both

hypervariable domains are compared. On average CR1 is almost twice as variable, within populations, as is CR2. This inequity rises to over 6-fold in the Dunlin (Wenink et al 1993), the only bird species for which both regions have been surveyed at the population level. It is therefore not suprising that studies which sequence only part of the control region invariably choose the left domain. (iii) The level of variation detected by the remaining hypervariable domain, CR2, in this cowbird population (%sd= 0.27) is comparable to that found in the Dunlin (%sd= 0.0-0.29) but is less than that found in other vertebrate populations such as humans (%sd= 1.22, 0.66) and Brown Trout (%sd= 0.27-0.65). This suggests that the CR2 may be less variable in the class Aves than in other vertebrates.

The sampling regime used in this study had several features which may have contributed to the low level of control region variation detected, compared to other vertebrate populations. First, the vertebrate populations listed in Table 2 are not strictly comparable with respect to the geographic scale of sampling. Some individual populations, like the Dunlin, Lesser Snow Goose, Brown Trout and Human, were surveyed over entire continents, whereas other populations, Brown-headed Cowbird, Grey-crowned Babbler, Kangaroo Rat and !Kung, were much more restricted.

Sampling from a limited geographic distribution would tend to underestimate the control region variation present within a larger, or possibly continuous population. For example, if the three Kangaroo Rat populations are considered as one large population then the %sd increases to 2.20 from a range of 0.33 to 1.66.

This cowbird population, having been sampled from a limited geographic distribution (see Miller and Gibbs in prep.), may not be representative of the total control region variation present. However, it is interesting to note that the cowbird sample from California was indistinguishable from the Manitoba cowbirds, based on control region sequence.

Second, the potential exists for at least some of the nestlings sampled to be half or full siblings, especially those from host nests in close proximity. Such a lack of independance of samples would tend to produce a more conservative estimate of within population variation. In fact, this bias has relatively little effect on the estimate of within population variation. The %sd calculated for the minimum number of females represented by this data (4, ie one for each haplotype detected) is still only 0.32%

compared to 0.17% when all 31 individuals are considered independant samples.

Third, I sampled only nestlings from two abundant, heavily parasitized host species (YW and RWB). If genetic structure correlating with female host use was present within this population, then the full range of host species would have to be included to ensure a representative sample was taken. However, no such structure has been found (Miller and Gibbs in prep.).

CONCLUSION

My results suggest that the CR1-deficient control region will be of limited use in intra-populational studies requiring high resolution but will be an effective marker for inter-specific comparisons. Its usefulness for phylogeography and phylogenetics of closely related Passerine species remains unclear and awaits a wider scale survey of control region variation.

LITERATURE CITED

Avise, J.C. 1991. Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. Annu. Rev. Genet. 25: 45-69

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T. et al. 1987. Intraspecific phylogeography: the mitochondrial bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18: 489-522.

Ball, R.M., Freeman, S., James, F.C., Bermingham, E. and Avise J.C. 1988. Phylogeographic population structure of Red-winged Blackbirds assessed by mitochondrial DNA. Proc. Natl. Acad. Sci. USA. 85: 1558-1562.

Ball, R.M. and Avise, J.C. 1992. Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. Auk 109: 626-636.

Bernatchez, L., Guyomard, R., and Bonhomme, F. 1992. DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout Salmo trutta populations. Mol. Ecol. 1: 161-173.

Brown, W.M., George Jr., M. and Wilson, A.C. 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA. 76: 1967-1971.

Brown, J.R., Gadaleta, G., Pepe, G., Saccone, C.and Sbisa, E. 1986. Structural conservation and variation in the D-loop containing region of vertebrate mitochondrial DNA. J. Mol. Biol. 192: 503-511.

Brown, J.R., Beckenbach, A.T. and Smith, M.J. 1993. Intraspecific DNA sequence variation of the mitochondrial control region of White Sturgeon (Acipenser transmontanus). Mol. Biol. Evol. 10: 326-341.

Cabot, E.L. and Beckenbach, A.T. 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput. Appl. Biosci. 5: 233-234.

Cann, R.L., Stoneking, M. and Wilson, A.C. 1987. Mitochondrial DNA and human evolution. Nature 325: 31-36.

Cantatore, P. and Attardi, G. 1980. Mapping of nascent light and heavy strand transcriots on the physical map of HeLa cell mitochondrial DNA. Nucleic Acids Res. 8: 2605-2625.

Chen, E.Y., Kuang, W-J. and Lee, A. 1991. Overview of manual and automated DNA sequencing by the dideoxy chain termination method. Methods: A Companion to Methods in Enzymology 3: 3-19.

Desjardins, P. and Morais, R. 1990. Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. J. Mol. Biol. 212: 599-634.

Desjardins, P. and Morais, R. 1991. Nucleotide sequence and evolution of coding and non-coding regions of a quail mitochondrial genome. J. Mol. Evol. 32: 153-161.

Digby, T.J., Gray, M.W. and Lazier, C.B. 1992. Rainbow trout mitochondrial DNA: sequence and structural characteristics of the non-coding control region and flanking tRNA genes. Gene 113:197-204.

Doda, J.N., Wright, C.T. and Clayton, D.A. 1981. Elongation of displacement loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc. Natl. Acad. Sci. USA. 78: 6116-6120.

Edwards, S.V. 1993. Long-distance gene flow in a cooperative breeder detected in genealogies of mitochondrial DNA sequences. Proc. R. Soc. Lond. Series B. 252: 177-185.

Ellegren, H. 1992. Polymerase-Chain-Reaction (PCR) analysis of microsatellites - A new approach to the study of genetic relationships in birds. Auk 109: 886-895.

Gelter, H.P., Gibbs, H.L. and Boag, P.T. 1993. Large deletions in the control region of Darwin Finch mitochondrial DNA: evolutionary and functional implications. Proc. Natl. Acad. Sci. USA. (in press).

Greenberg, B.D., Newbold, J.E. and Sugino, A. 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in the human mitochondrial DNA. Gene 21: 33-49.

Gyllensten, U.B. and Erlich, H.A. 1988. Generation of single-stranded DNA by the polymerase chain reaction and it's application to direct sequencing of the HLA-DQα locus. Proc. Natl. Acad. Sci. USA. 85: 7652-7656.

Harrison, R.G. 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. 4: 6-11.

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

Marchuk, D., Drumm, M., Saulino, A. and Collins, F.S. 1991. Constructon of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res. 19: 1154.

Meyer, A., Kocher, T.D., Basasibwaki, P. and Wilson, A.C. 1990. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. Nature 347: 550-553.

Moritz, C., Dowling, T.E. and Brown, W.M. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Ann. Rev. Ecol. Sys. 18: 269-292.

Pääbo, S., Thomas, W.K., Whitfield, K.M., Kumazawa, Y. and Wilson, A.C. 1991. Rearrangements of mitochondrial transfer RNA genes in marsupials. J. Mol. Evol. 33: 426-430.

Quinn, T.W. 1992. The genetic legacy of Mother Goose- phylogeographic patterns of lesser snow goose Chen caerulescens caerulescens maternal lineages. Mol. Ecol. 1: 105-117.

Quinn, T.W. and Wilson, A.C. 1993. Sequence evolution in and around the mitochondrial control region in birds. J. Mol. Evol. (in press)

Rhoads, D.D. and Roufa, D.J. 1989. Seqaid II version 3.5. Molecular Genetics Laboratory, Kansas State University.

Saccone, C., Attimonelli, M. and Sbisa, E. 1987. Structural elements highly preserved during the evolution of the D-loop containing region in the vertebrate mitochondrial DNA. J. Mol. Evol. 26: 205-211.

Saccone, C., Pesole, G. and Sbisa, E., 1991. The main regulatory region of mitochondrial DNA: Structure-function model and evolutionary pattern. J. Mol. Evol. 33: 83-91.

Sanger, F., Nicklen, S. and Coulson, A.R. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74: 5463-5467.

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y.

Seutin, G., White, B.N. and Boag, P.T. 1991. Preservation of avian blood and tissue samples for DNA analyses. Can. J. Zool. 69: 82-90.

Shedlock, A.M., Parker, J.D., Crispin, D.A., Pietsch, T.W. and Burmer, G.C. 1992. Evolution of the Salmonid mitochondrial control region. Mol. Phylog. and Evol. 1: 179-192.

Sheilds, G.F. and Wilson, A.C. 1987. Calibration of mitochondrial DNA evolution in geese. J. Mol. Evol. 24: 212-217.

Smith, M.F. and Patton, J.L. 1991. Variation in mitochondrial Cytochrome b sequences in natural populations of South American Akodontine rodents (Muridae: Sigmodontinae). Mol. Biol. Evol. 8: 85-103.

Stoneking, M., Hedgecock, D., Higuchi, R.G., Vigilant, L., and Erlich, H.A. 1991. Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. Am. J. Genet. 48: 370-382.

Thomas, W.K., Paabo, S., Villablanca, F.X. and Wilson, A.C. 1990. Spatial and temporal continuity of kangaroo rat populations shown by sequencing mitochondrial DNA from museum samples. J. Mol. Evol. 31: 101-112.

Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. and Wilson, A.C. 1991. African populations and the evolution of human mitochondrial DNA. Science, Wash. 253: 1503-1507.

Vigilant, L., Pennington, R., Harpending, H., Kocher, T.D. and Wilson, A.C. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Proc. Natl. Acad. Sci. USA. 86: 9350-9354.

Walberg, M.W. and Clayton, D.A. 1981. Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res. 9: 5411-5421.

Wenink, P.W., Baker, A.J. and Tilanus, M.G.J. 1993. Hypervariable control region sequences reveal population structuring in a long-distance migrant shorebird, the Dunlin (Calidris alpina). Proc. Natl. Acad. Sci. USA. 90: 94-98.

Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D. and Stoneking, M. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26: 375-400.

Table 1 Variation in 524 bases of control region sequence among 31 nestling cowbirds. Base positions are relative to the cloned cowbird sequence and are highlighted in Fig. 2.

Haplotype	No. of individuals	Bases at variable positions		
		187	526	600
1	17	C	C	-
2	9	C	C	A
3	2	T	C	-
4	3	T	T	-

Note: The reference cowbird from California did not vary from the Manitoba cowbirds at any additional nucleotide positions and was of haplotype 3.

Table 2 Intra-populational variation in the mt DNA control region across vertebrate taxa. Levels of sequence divergence between the hypervariable regions CR1 and CR2 are compared. The measure of divergence calculated was percent sequence divergence defined as the average pairwise divergence among individuals. The column TOT represents the combination of CR1 and CR2 together excluding the rarely published central domain.

TAXA	Population	n	CR1	CR2	TOT	Study
Passerine birds						
Brown-headed Cowbird (*Molothrus ater*)	Delta, Manitoba	31	n\a	0.27	n\a	
Grey-crowned Babbler (*Pomatostomus temporalus*)	*P.t. temporalus*	5(9)	1.7	n\a	n\a	Edwards (1993)[1]
	P.t. rubeculus	7(6)	1.0	n\a	n\a	
Non-passerine birds						
Dunlin (*Calidris alpina*)	Alaska	14	0.37	0	0.18	Wenink et al (1993)
	West Coast N.A.	3	1.83	0	0.89	
	Taymyr peninsula	14	0.96	0.12	0.53	
	Western Europe	23	0.95	0.29	0.61	
	Gulf Coast N.A.	16	0.37	0.17	0.27	

Table 2 (continued)

TAXA	Population	n	CR1	CR2	TOT	Study
Lesser Snow Goose	Clade 'A'	30	0.43	n\a	n\a	Quinn (1992)[2]
(Chen caerulescens)	Clade 'B'	51	1.21	n\a	n\a	
Mammals						
Kangaroo Rat	'A'	34	1.66	n\a	n\a	Thomas et al
(Dipodomys	'B'	40	0.33	n\a	n\a	(1990)
Panamintinus)	'C'	32	0.79	n\a	n\a	
	combined	108	2.20	n\a	n\a	
Humans	African origin	9	2.31	1.22	1.70	Stoneking et al
(Homo sapiens)						(1991)
	Kung!	15	1.06	0.66	0.86	Vigilant et al (1989)[3]
Fish						
Brown Trout	Atlantic Basin	53	0	0	0	Bernatchez
(Salmo trutta)	Danube Basin	12	0.55	0.27	0.41	et al (1992)
	Mediterranean	36	0.89	0.65	0.77	
	Adriatic	50	0.59	0.57	0.58	

Table 2 (continued)

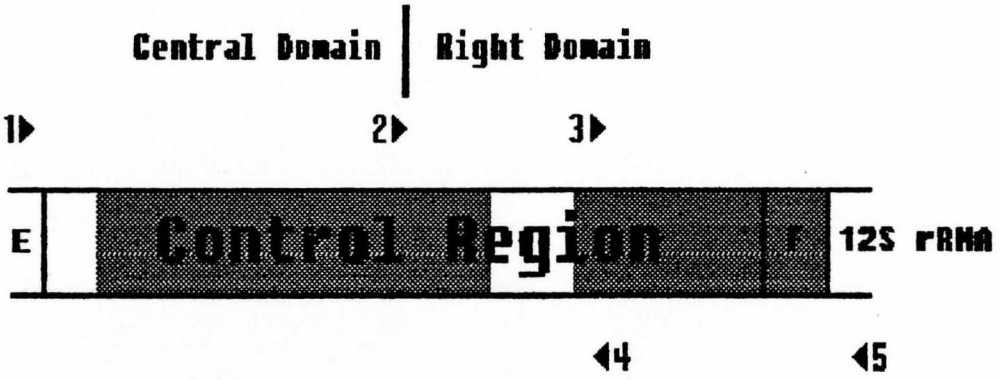
TAXA	Population	n	CR1	CR2	TOT	Study
White Sturgeon	Columbia river	16	1.13	n\a	n\a	Brown et al
(Acipenser	Fraser river	11	0.75	n\a	n\a	(1993)
transmontanus)	Combined	27	1.20	n\a	n\a	

[1]Twelve populations were surveyed in this study, 5 from the subspecies P.t. temporalis and 7 from P.t. rubeculus. The average percent sequence divergence within populations of each subspecies is given, with the average population size in parentheses.

[2]The two clades identified by Quinn (1992) were used in comparisons of within population variation in the control region, rather than the three populations actually surveyed. I felt that this was appropriate since the large difference between clades (6.7% sequence divergence) would artificially inflate measures of divergence within populations containing individuals from both clades.

[3]The absolute size of CR1 and CR2 were reported in Stoneking et al (1991). My calculations underestimate actual percent sequence divergence calculated by Vigilant et al (1989) since only, approximately 740 bases of a possible 756 were sequenced. For example, Vigilant et al (1989) calculated a TOT percent sequence divergence of 0.88% for 15 !Kung compared to my value of 0.86%.

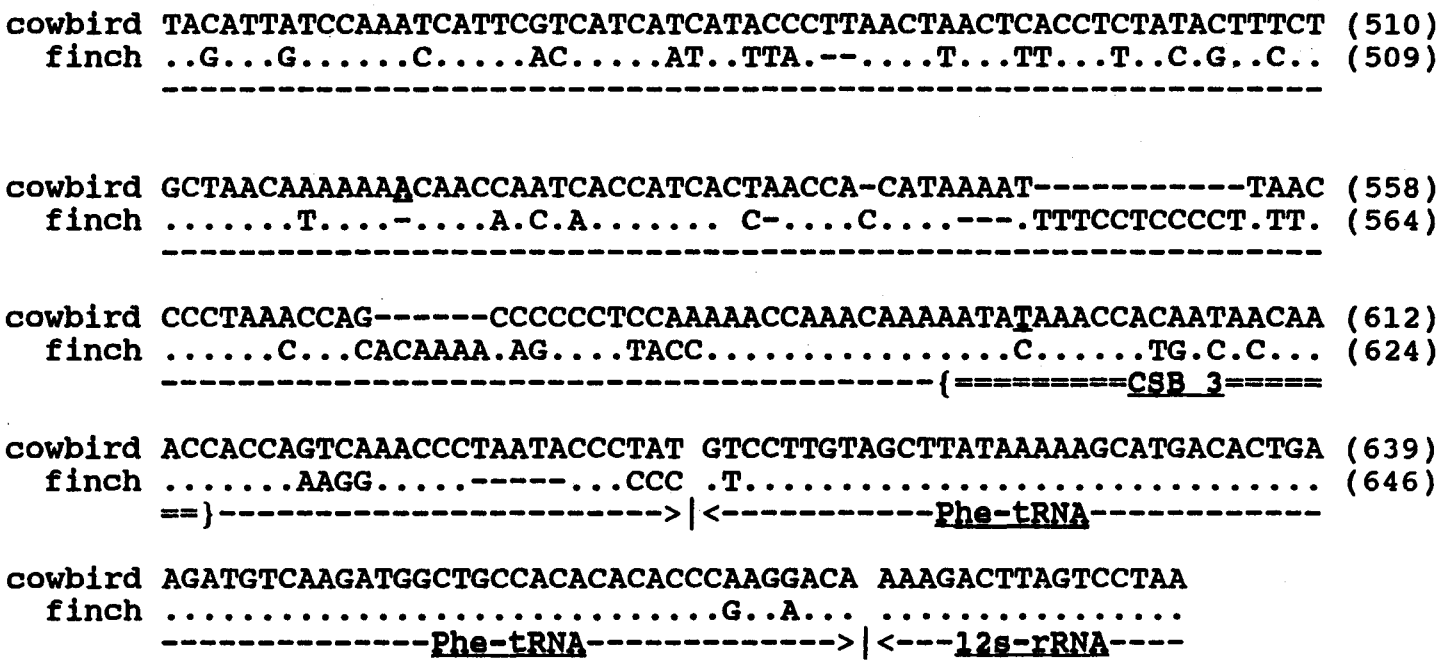
Figure 1 Primer locations and sequences used in the sequencing of the cowbird control region. The orientation shown is for the light strand, 5′ to 3′, beginning with the Glu-tRNA. The directionality of the primers are indicated by the arrowheads; light strand primers above and heavy strand primers below the figure. The primer sequences are given in the 5′ to 3′ orientation. The complete sequence of the 639 base cowbird control region and the 68 base, 5′ flanking Phe-tRNA(F) are shown, as are partial sequences of the 3′ flanking Glu-tRNA(E) and the 12s rRNA gene. This sequence was obtained from a cloned mtDNA fragment amplified with universal Passerine primers from Gelter et al.(1993) (see text). The shaded areas represent parts of the cowbird control region sequenced in this study.



1	GSL-Glu	TTGGTTGTAACTTCAGGAAC
2	MAL-CB1	AGGGCGCGCTACCTACCTCT
3	MAL-CB2	TTCATCAAACAATAAACCCA
4	MAH-CB3	TATGGGTTTATTGTTTGATG
5	GSH-12s	AAGGTTAGGACTAAGTCTTT

Figure 2 Brown-headed Cowbird (Molothrus ater) control region (top), aligned with Darwin's finch (Geospiza scandens) control region. The sequence reported is for the light strand and is shown in the 5′ to 3′ orientation. Dots indicate identical bases between cowbird and finch sequence while dashes indicate gaps. Numbers in parentheses refer to the base position of the control region sequence, with base 1 defined as the first base downstream of the Glu-tRNA. Highlighted bases in the cowbird sequence denote intra-populational variable sites (see Table 1). Darwin's finch sequence and labelled structural features are from Gelter et al (1993).

```
                                        1
cowbird TTGGTTGTAACTTCAGGAAC ATAAACCCCTCCAAT--CCTTCTCACTTGCTCTTCACAG (37)
  finch ..................... ...--..TGA.G...TC...C...-.............. (36)
        ------Glu-tRNA----->|<-----------Central Domain-----------

cowbird ATACAAGTGGTCGGTTGAATATnCCTCCCTACTCTCATTACCTCGGCATACCGACCTCCT (97)
  finch .....................CT........T........TT................T. (96)
        ------------------------------------------------------------

cowbird ACACTTGTTTTTTTTTAGCGTCTCTTCAATAAGCCCCTCAAGTGCAGAGCAGGTGTTATC (157)
  finch .............C.....................A.....................A.... (156)
        ------------------------------------------------------------

cowbird TTCCTCTTGACATGTCCATCACATGACCGTCGAGCATATGAATCCCCTAACACCCAGAAT (217)
  finch ............................C...A.......................... (216)
        ------------------------------------------------------------

cowbird GTCATGGTCTGACGGATAAGGTC-GTCGCAAACTTGGCACTGATGCACTTTGACCCCATT (277)
  finch .......A..A............A..................................... (276)
        ------------------------------------------------------------

cowbird CATGGAGGGCGCGCTACCTACCTCTAGACAACAGATAGTGTAATGGTTGCCGGACATATC (337)
  finch ......................................................... (336)
        ----------------------------{=========CSB-1==========}----

cowbird AAT TATTTTATCATTTACTAGGGACTGTCATTTAAATCCCATTTTACGCATC-ATTTTT (395)
  finch -.. ..........AA.T.....A.T.TA.........T..............T...... (395)
        -->|<------------------Right Domain-----------------------

cowbird TTTTTTTTATCTTGA--TTTTTA--TTTTTTTTCATCAAACAAT-AAACCCATAAATTCC (450)
  finch ...----C.......CA......AAA.......G.......... .....-...T..... (450)
        ------------------------------------------------------------
```


cowbird TACATTATCCAAATCATTCGTCATCATCATACCCTTAACTAACTCACCTCTATACTTTCT (510)
finch ..G...G.....C.....AC.....AT..TTA.--....T...TT...T..C.G..C.. (509)
cowbird GCTAACAAAAAAACAACCAATCACCATCACTAACCA-CATAAAAT----------TAAC (558)
finch .....T....-....A.C.A..... C-....C....---.TTTCCTCCCCT.TT. (564)
cowbird CCCTAAACCAG------CCCCCCTCCAAAAACCAAACAAAAATATAAACCACAATAACAA (612)
finch ......C...CACAAAA.AG....TACC.....C.....TG.C.C... (624)
CSB 3
cowbird ACCACCAGTCAAACCCTAATACCCTAT GTCCTTGTAGCTTATAAAAAGCATGACACTGA (639)
finch .....AAGG.....-----...CCC .T..... (646)
Phe-tRNA
cowbird AGATGTCAAGATGGCTGCCACACACACCCAAGGACA AAAGACTTAGTCCTAA
finch .....G..A... .....
Phe-tRNA
12s-rRNA

CHAPTER 2

Genetic Analysis of Host Specificity in the Brown-headed Cowbird (*Molothrus ater*): Results from mtDNA and RAPD markers.

ABSTRACT

Genetic analyses of Brown-headed Cowbird (*Molothrus ater*) chicks reared in two different host species' nests, the Red-winged Blackbird (*Agelaius phoeniceus*) and the Yellow Warbler (*Dendroica petechia*) were done to test for the presence of host specific female linneages (Hypothesis 1) or cryptic species (Hypothesis 2) among a population of cowbirds from Delta Manitoba. The results showed that: 1) There were no detectable differences in mtDNA control region haplotypes between cowbird chicks from different host nests 2) RAPD nuclear markers suggested some differentiation between the putative host races. However, the analysis of the RAPDs was severely affected by the presence of outliers in the data set. An unambiguous interpretation of RAPD marker variation in this population would require further

analysis using a larger sample size. Thus, I provisionally conclude that the hypothesis which is best supported by the data is that no female-specific host races or cryptic species exist in this population of Brown-headed Cowbirds.

INTRODUCTION

Obligate interspecific brood parasitism is a reproductive tactic employed by approximately 1 % of bird species, roughly 80 in total (Payne 1977). Female parasites lay their eggs in the nests of 'host' species who, should they accept the egg, then raise the parasite young.

When a host accepts a parasitic egg, the likely outcome is a reduction in the hosts' reproductive success, either directly by ejection of unhatched eggs and nestlings or indirectly via competition between host and parasite nestlings for food resources (Payne 1977). This loss of fitness results in potent selection pressure for the development of host defenses. Host rejection of the parasitic egg has a direct impact on the parasite's fitness and should therefore prompt an evolutionary response such as host switching, if other suitable hosts are present or alternatively, the parasite may evolve appropriate counter-adaptations to circumvent host defenses, especially if host choice is limited. The result of this dynamic process of adaptation and counter-adaptation between parasite and host would lead to a parasite specialized for that particular host.

Some parasites, like the Screaming Cowbird (Molothrus rufoaxillaris), are extreme specialists which parasitize a single host, the Baywinged Cowbird (M. badius) and display co-evolved adaptations to host defenses, such as egg mimicry (Friedmann 1929). Other parasites appear to adopt more of a generalist approach by parasitizing multiple host species within their range. For example, the Common Cuckoo (Cuculus canorus) uses many host species throughout it's extensive range, yet only a few species are parasitized in any one region (Southern 1954). Finally there are parasites such as the Brown-headed (M. ater) and Shiny (M. boneriensis) Cowbirds which are considered extreme generalists, parasitizing nearly every passerine with which they are sympatric (Friedmann et al 1929).

The Brown-headed Cowbird's designation as an extreme generalist has been largely based on it's long host species list and apparent lack of co-evolved adaptations to any one host. However, the direct observation of cowbird laying patterns is logistically very difficult due to the cryptic behaviour of females during egg laying and hence, little direct evidence exists to support the contention that cowbirds are generalists.

Most studies which have addressed the question of host specificity in this species have provided only circumstantial evidence based on assigning eggs laid in various nests to individual females using egg markings and morphology. Of these studies, some have found no evidence for individual specialization (Friedmann 1929, Friedmann et al 1977, Jones 1941) whereas others found that at least some females were using predominantly one host (Walkinshaw 1949, McGeen and McGeen 1968). The only direct study of individual female cowbird laying patterns, which used protein electrophoresis to match eggs laid in various host species' nests to individual females, found no evidence of host specificity (Fleischer 1985). However, the number of female's surveyed in this study was small and the host species used by the cowbirds were ecologically very similar. In addition, this sort of analyses has not been repeated for other populations of cowbirds.

In this study, I have taken a different approach to address the question of host-specificity in Brown-headed Cowbirds. Rather than concentrate on individual females (cf. Fleischer 1985), I have chosen to examine this issue at a population level. In particular I assume that long term host-specificity would have a significant and detectable effect on the genetic structure of cowbird populations. For

example, within a population of host specific female cowbird lineages, different mutations would accumulate in each lineage over time. This would result in a genetically structured population with respect to host use, assuming that gene flow between lineages is limited and that the markers used are sensitive enough to detect differences between lineages.

Different hypotheses about the kind of structure present in cowbird populations can be assessed using a combination of two types of markers which have different modes of inheritance. These include mitochondrial DNA (mtDNA), which is inherited maternally, and nuclear DNA, which segregates through both sexes. By comparing the patterns of structure found in each of these markers I can test the following hypotheses about host choice in female cowbirds: (1) There are female host specific races within the population whereas males do not mate assortatively with females from specific races. (2) There are actually cryptic species, hence positive assortative mating, within the population ans each 'species' specializes on a different host. (3) There has been no long term host specificity in a population, or it is not detectable with these particular markers.

If (1) is true, an analysis of mtDNA variation would reveal distinct races, since mtDNA is inherited maternally. In contrast nuclear DNA would still be exchanged between races through male gene flow and hence no genetic structure would be found for this marker. If (2) were true then neither mitochondrial nor nuclear DNA would be exchanged between races and hence the genetic analysis of both types of DNA would reveal structure in each. Finally, if neither type of DNA reveal genetic structure then (3) would be supported.

Current data describing the host-use habits of the Brown-headed Cowbirds (Payne 1977, Rothstein 1990) appear to support hypothesis 3. However, individual females may still specialize to some extent (cf Walkinshaw 1949, McGeen and McGeen 1968). Other brood parasitic birds fall into the two remaining categories. For example,the Common Cuckoo parasitizes multiple hosts in any one locale but there is evidence, based on host egg mimicry, for host specific female lineages or gens (Southern 1954). It is not clear whether males mate assortatively within their gentes; however, the lack of any morphological or behavioral differences between gens (Southern 1954) suggests that hypothesis (1) best describes this system. In at least one Viduine finch the potentials exists for cryptic species to

be present (hypothesis 2) since females are host specific and preferentially mate with males who mimic the song of their host species (Payne 1973).

Molecular techniques

To assess mtDNA variation I have used the Polymerase Chain Reaction (PCR) to amplify and directly sequence part of the cowbird control region. The control region was selected because it has been shown to evolve at a rate approximately 5 times faster than the rest of the mtDNA molecule in humans (Greenberg et al. 1983, Cann et al. 1987), therefore making it the most likely region to possess within population variation. Recent studies have shown that variation in this region can be successfully used to detect genetic structure at the population level in birds (Quinn 1992, Wenink et al. 1993, Edwards 1993).

To survey nuclear DNA variation, Random Amplified Polymorphic DNA (RAPD) markers were generated. RAPD is a technique which uses the polymerase chain reaction (PCR), in combination with a single, short oligonucleotide primer (usually 10 bp), to amplify small fragments of genomic DNA, less than 5kb in size (Williams et al. 1990, Welsh and McClelland 1990). Polymorphims are detected by the presence or absence of individual bands, where absence of a band is

most often due to the loss of a priming site. Most bands therefore behave as dominant markers and can be analysed in the same way as restriction fragment length polymorphisms (RFLPs), with a correction for the inability to detect heterozygotes (Clark and Lanigan, in press).

The purpose of the study was to use these two genetic markers to analyze variation in Brown-headed Cowbird chicks collected from the nests of two hosts species in Southern Manitoba to test the three hypotheses outlined above. This type of population level analysis can be applied to the study of host-specificity in many other parasitic species where multiple hosts are parasitized.

MATERIALS AND METHODS

Study Site: Samples were collected at the University of Manitoba Field Station, Delta Manitoba between May and July, 1992, in collaboration with Dr. Spencer Sealey. The Delta site provided a unique opportunity to investigate cowbird host preference for two important reasons: First, cowbird parasitism rates are extremely high, thereby increasing the likelihood that adequate samples could be obtained (Weatherhead 1989, Sealey pers. comm.). Second, within this site there are two heavily parasitised cowbird hosts, the Red-winged Blackbird (Agelaius phoeniceus) (RWB) and the Yellow Warbler (Dendroica petechia) (YW) (Weatherhead 1989, Sealey pers. comm.) which occupy strikingly different (marsh versus forest) yet contiguous habitats which have a sharp boundary between them. This unique habitat extends along most of the southern shore of Lake Manitoba.

I felt that a comparison of cowbird chicks from two such ecologically different habitats would be the most likely one to detect cowbird host specificity, if it exists. The strategies that a female cowbird uses to parasitize one woodland host species's nests are likely transferable to other woodland species. However, these strategies may not be optimal for parasitizing a host from a distinctly different

habitat, such as a marsh, and may require a different set of skills, thereby leading to specialization on hosts from one or the other habitat.

Sample Collection: To obtain tissue samples for genetic analyses, Dr. Sealey's students and I searched for parasitized Red-winged Blackbird and Yellow Warbler nests in the habitat surrounding the station and also at the Bell Estate, located approximately 5 km to the east of the station. Parasitized nests were marked and monitored each day. Blood samples (15-100 ul) were taken from the jugular of nestlings, no sooner than four days post-hatching, and stored in 800 ul of 1 x lysis buffer (4 M urea, 0.2 M NaCl, 0.1 M Tris-HCl [pH 8.0], 0.5 % n-laurylsarcosine, 10 mM EDTA, [Seutin et al 1991]) at 4 °C. In cases where the nest was abandoned, the unhatched cowbird egg or dead nestling was collected and imediately stored at -20 °C in a sterile 15 ml tube. In total, tissue samples were collected from 19 cowbird nestlings from Yellow Warbler nests and 13 from Red-winged Blackbird nests.

The majority of nestling's sampled either came from host nests seperated by at least 200 m or were shown to have been from different females because they had different mtDNA haplotypes. In one case however, two samples with the same

haplotype came from the same host nest and five other pairs of samples came from nests that were less than 100 m apart. Therefore, the potential exists for at least some of these nestlings to be half or full siblings. Most of these potential sibling groups were laid in nests of the same host species, with only one pair laid in nests of different host species. This potential lack of independence of samples would tend to reduce the estimates of within host-population variation and artificially inflate divergence between populations.

To assess the possible effect of relatedness among chicks sampled on the results of my analysis, I have adopted an alternate approach to an evaluation of genetic relationship (cf. Westneat 1991) based on DNA fingerprinting (cf. Burke and Bruford 1987, Wetton et al. 1987): For all genetic analyses I consider both (i) the entire data set and (ii) a group of restricted data sets wherein the samples used were limited to those obtained from nests further apart than an estimate of the size of female cowbirds' egg laying territories. The only study to track individual female egg laying (Fleisher 1985) found that most females laid there eggs within 300 m of each other and that there was considerable overlap in female egg laying territory. Banded females at the Delta site were observed to have a more

restricted territory, approximately 200 m, most likely due to the extremely high density of cowbirds at this site (Sealey pers. comm.). The high density of females and observed overlap of female egg laying territories (Fleischer 1985) make eggs laid in host nest further than 200 m apart unlikely to be siblings. Therefore individuals which shared the same mtDNA haplotype and came from foster nests within this 200 m radius were first grouped and then a randomly chosen representative from each group was used to make up a new restricted data set. In the one case where two nestlings were sampled from different host nests that were within 200 m apart, I chose the cowbird from the RWB nest to make up the data set in a effort to increase the sample size of the RWB cowbirds. Twenty restricted data sets were produced in total, which consisted of 14 YW and 9 RWB reared nestlings.

Sequencing

The primers GSL-glu and GSH-12S (Gelter et al 1993), designed from Darwin's Finch mtDNA, were used in PCR amplifications of the control region from Brown-headed Cowbird genomic DNA. Asymmetric PCR (cf. Gyllensten and Erlich 1988) of the double stranded control region template was used to produce single stranded template for sequencing. Dideoxy sequencing (Sanger, Nicklen & Coulson 1977) with Sequenase 2.0 kits (USB) was then carried out on the single

stranded control region template. The protocols used for DNA extraction, PCR amplification and sequencing have been described in detail elsewhere (Miller and Gibbs in prep.).

In total, 524 bases of the cowbird control region were determined for 31 nestlings, 18 from YW nests and 13 from RWB nests. I was unable to produce single stranded DNA template for one individual, laid in a YW nest (see Miller and Gibbs in prep.). Three variable nucleotide positions were detected within the 524 bases sequenced, defining 4 haplotypes (Table 1).

Sequence analysis: I used a X^2 test of association on the distribution of the 4 haplotypes to test whether the frequency of haplotypes differed among cowbirds chicks sampled from RWB versus YW reared nests (see Table 1). However, the small sample sizes of 4 of the 8 cells made the validity of this statistic suspect. To avoid combining data to increase cell size, I used a Monte Carlo solution to this problem suggested by Roff and Bentzen (1989) and later expanded and programmed by Zaykin and Pudovkin (in press). The general approach is to generate the χ^2 distribution expected under the null hypothesis of homogeneity, compute the observed value of X_o^2 and determine where it lies within this randomly generated distribution.

The computer program generates random RxC tables, with the restriction that marginal row and column totals remain constant. It then calculates the X_r^2 value for each table and compares it to the value from the original table, X_o^2. The ratio of the number of tables generated, where $X_r^2 \geq X_o^2$, to the total number of tables generated is given as the probability of homogeneity. If the haplotypes are distributed non-randomly between YW and RWB fostered nestlings then the probability of homogeneity would be low and the interpretation would be that the population was genetically structured with respect to host preference. Conversely, failure to reject the null hypothesis of homogeneity would suggest that no mtDNA stucture is present between host races.

RAPDS

Amplification conditions: Amplifications were performed in a Perkin Elmer Cetus thermal cycler under the conditions described by Williams et al (1990) with the following modifications: 50 ng of genomic DNA was used as the template, and in the last twenty cycles of the PCR program the extension time was incremented by ten seconds for each successive cycle. A total of ten different, ten base pair, random sequence primers from Operon Technologies (kit A, 1-10) were used to amplify DNA from thirty-two individual

cowbird nestlings. Negative controls (distilled water was substituted for template DNA) were included in each amplification to evaluate the possibility of contamination. Each amplification was repeated at least once. Amplification products were run on 1.2 % agarose 0.5 x TBE gels until the Orange-G loading dye had run at least 14cm. The gels were then stained with Ethidium Bromide, destained in water, visualised under shortwave UV light and photographed.

Occasionally amplification products, which did not correspond in size to any other amplified fragment, were present in the negative control lanes. This appears to be a common phenomenon when generating RAPDs in a wide variety of taxa (Hadrys, Balick and Schierwater 1992) and is probably the result of low level contamination by other DNA sources which are outcompeted by the template DNA in experimental reactions and hence are visible only in the control lane.

Scoring: For each amplification, all bands were identified the name of the primer used and their approximate size in kilobases. All bands were assumed to be of nuclear DNA origin due to the large size disparity between the mitochondrial and nuclear DNA genome. The bands were then scored for presence versus absence in each individual. Following the suggestions of Hadrys, Balick and Schierwater

(1992), only those bands that could be unambiguously scored in each individual and were consistantly present in different amplifications of the same individuals were used to make up the data set. The darker bands tended to be both: (i) more reliably amplified and (ii) easier to score as present versus absent than faint bands (Figure 1).

RAPD analysis: Most population level analyses of RAPDs have failed to address this marker's inability to discriminate between homozygotes and heterozygotes, a shortcoming which would tend to underestimate nucleotide diversity in traditional bandsharing analyses (Clark and Lanigan, in press). The majority of studies have used Nei and Li's (1979) bandsharing calculation to generate similarity matrices on which some form of cluster analysis is done to look for population substructuring (cf. Kazan et al 1993, Russel et al 1993). Other studies have addressed the same problem by comparing a measure of genotypic diversity within and between populations (cf. Russel et al 1993). Neither method attempts to correct for the inability of RAPDs to detect heterozygotes.

To address this problem I have chosen to analyze RAPDs using a computer program, RAPDDIP, developed by Clark and Lanigan (in press) which estimates nucleotide diversity

'π' within and divergence 'd' between populations. The program calculates a modified bandsharing statistic, similar to the one used in RFLP studies (Nei and Li 1979), which corrects for the inability of RAPDs to detect heterozygous loci in diploid individuals. The modified bandsharing values are then used to estimate nucleotide diversity and divergence as suggested by Nei (1987).

To test the statistical significance of the estimates of sequence divergence and to further explore patterns in the data I used three seperate randomization techniques: (1) a permutation test, (2) the jacknife and (3) bootstrapping. Randomization techniques, such as these, obviate the necessity of satisfying the restrictive condition of normality required by parametric statistics by randomly resampling the data multiple times and calculating the test statistic each time to produce a null or random distribution of the test statistic. Hypotheses can then be tested by comparing the observed value of the test statistic to the null distribution of the test statistic. The more traditional way of addressing the problem of non-normal data, ie non-parametric statistics, lacks the power to accept alternative hypotheses that randomization techniques offer (Manly 1991).

The program RAPDDIP was modified to produce a specified number of either permuted or bootstrapped data sets and calculate both d and π for each. Permuted data sets were created by random assignment of individuals to one of the two purported host populations, while bootstrapped data sets were created by random sampling of all individuals, with replacement, until both purported host races had attained their original population size. The histogram of these random d's represents the null distribution of the parameter if no structure were present in the data. The 95th percentile of this distribution was used as the cutoff point to test the significance of the observed d from the original data set.

I used the jacknife procedure to assess each individual's contribution to the overall d and determine if outliers were present in the data set. An outlier is an unusually large or small data point that exerts a disproportionately large effect on the calculation of many test statistics. The jacknife sequentially removes single individuals from the data set and recalculates the parameter producing a set of 'pseudo values' which represent individuals' contribution to the calculation of the parameter. Potential outliers in this set of pseudo d's were identified using Minitab 5.1.1 which labelled as 'possible'

outliers all observations lying outside of the 'inner fence' of data, described as the range $\{Q_1-3/2(Q_3-Q_1),\ Q_3+3/2(Q_3-Q_1)\}$ and labelled as 'probable' outliers all observations lying outside of the outer fence of data, described as the range $\{Q_1-3(Q_3-Q_1),\ Q_3+3(Q_3-Q_1)\}$, where Q_1 and Q_3 are the first and third quartiles respectively (Tukey 1977). This method of identifying outliers was used since the distribution of pseudo d's was non-normal.

It is important to identify outliers in a data set since they may exert more influence on the calculation of the parameter than other data points and can therefore bias the interpretation of results, especially when sample sizes are small as was the case in this study.

RESULTS

mtDNA: Visual inspection of the frequency of haplotypes among YW- and RWB-fostered cowbirds in Table 1 suggests that they are distributed randomly among the two groups. In support of this observation, the Monte Carlo simulation failed to reject the null hypothesis of homogeneity (X^2 = 1.13819, P = 0.8234, 10000 randomizations). Similar results were obtained when each of the twenty restricted data sets (see Methods, Table 2) were analysed (all P >> 0.05, 10000 randomizations).

RAPDS: Of the 10 primers surveyed 8 produced at least one scorable band. Thirty bands were scored in total, varying in size from 0.7 to 3.0 kilobases. Fourteen of these bands were totally monomorphic and another 7 were present in at least 75 % of the individuals surveyed. Three bands were relatively rare, ie. present in less than 25 % of individuals, while the remaining 6 were moderately variable, present in between 25 % to 75 % of individuals (see Appendix 1 for raw data).

The program RAPDDIP (Clark and Lanigan in press) estimated the nucleotide diversity within the YW and RWB

purported host races to be 0.179 % and 0.134 % respectively, with a nucleotide divergence between them of 0.204 %.

The null distribution of permuted and bootstrapped ds, based on 10,000 randomizations, are displayed in Figure 2. The cutoffs for the 95th percentile were 0.198 % and 0.205 % for the permutted and bootstrapped null distributions respectively. The observed d (0.204 %) was within the 99.5th percentile of the permutted null distribution and within the 94th percentile of the bootstrapped distribution, ie the observed d was greater than 9953 of the permutted d's and 9402 of the bootstrapped d's. Thus, the presence of nuclear DNA structure is strongly supported by the permutation test ($p < 0.01$) but less so by the bootstrap analysis ($p = 0.06$).

The analyses of the restricted data sets are presented in Table 2. The permutation test resulted in nine of 20 data set's d values being considered significant at the $p = 0.05$ level, whereas the bootstrap analysis resulted in only four of these being significant.

The individual jacknife estimates of d revealed that most nestlings, 22 of 32, had a relatively small impact on the observed d, falling within the inner fence of

observations (0.187% to 0.219 %). Six nestlings, 19, 29, 41, 48 and 49 from YW nests and 32 from a RWB nest were identified as possible outliers, falling between the inner and outer fences (0.187 % to 0.175% and 0.219 % to 0.231%). The remaining 4 nestlings, 35 from a YW nest and 31, 38 and 46 from RWB nests (see Appendix 1) were identified as probable outliers, lying outside of the outer fences (0.175% to 0.231%).

DISCUSSION

Interpretation of molecular data

mtDNA: Brown-headed cowbird mtDNA haplotypes were randomly distributed between chicks from YW and RWB nests suggesting that this population is homogeneous for mtDNA variation and that no population detectable structure along host species lines exists.

Limitations: The control region was chosen for the survey of mtDNA variation because it is the most rapidly evolving region in the molecule. However, little within population sequence variation was detected: only 3 variable sites were observed in the 524 bases sequenced for an average percent sequence divergence of 0.17 % (Miller and Gibbs in prep.). Comparable studies of within population variation in other bird species have found percent sequence divergence estimates far exceeding that found in this cowbird population, some by as much as much as six times (summarised in Miller and Gibbs in prep.). The comparitively low level of variation observed in this study was in part due to the deletion of the left hypervariable domain in the cowbird control region (Miller and Gibbs in prep.). The variation detected in the remainder of the cowbird control region is

however, typical of that found in this region in other birds (Wenink et al 1993).

The lack of variation in the control region raises concern about the sensitivity of this marker to detect within population genetic structure. In particular, recently diverged female specific lineages may not have had time to accumulate different mutations in this region and would therefore not be detected. For example, the minimum detectable difference in two host race's control region, e.g. one base in the 524 bases sequenced or a %sd of 0.19, would require the lineages to have been seperated for more than 4.57×10^4 years, assuming that the cowbird control region evolves at a rate of 4.16 %sd / 10^6 years (Miller and Gibbs in prep.). Unfortunately, no other mtDNA markers evolve as quickly as the control region and therefore the prospects for detecting recently evolved female specific lineages in this population, if they exist, are not very promising.

RAPDs: Interpretation of these results proved difficult for two reasons, both of which are related to the presence of outliers in the data set. First, there was a lack of agreement between the results of the two randomization procedures used to test for the significance of observed

d's. In addition, analysis of restricted data sets revealed further discrepancy between these two randomization procedures. The permutation test resulted in 9 of 20 data sets with significant d's while only 4 of these gave a significant result using the bootstrap analysis, although the remaining 5 approached significance ($0.05 < p < 0.1$) (Table 2). Secondly, the observation that some restricted data sets have significant d's while others do not is puzzling and supports the results of the jacknife analysis which suggested the presence of outliers in the data set.

The discrepancy between the two randomization procedures was likely due to the different resampling regimes used in each which resulted in a differential sensitivity to outliers. The permutation test treats each observation as an equally likely event and randomly assigns them to a treatment level, or in this case to one of two populations, such that the total number of individuals in each population remains constant. Therefore each individual is present only once in any given permuted data set. Bootstrapping regards the sampled genotypes as the best representation of the actual population distribution of genotypes. The true population can therefore be approximated by an infinite population wherein the sampled genotypes are equally likely (Manly 1991). In short, the observations are

pooled and randomly sampled with replacement until each population has reached it's original size, thereby producing bootstrapped data sets which may contain either multiple cases of a particular individual or no representation of that individual at all.

The significance of this difference in data resampling can be illustrated by comparison of the null distribution of d's produced by each randomization procedure (Figure 2). The bootstrapped distribution is much broader and has longer tails than the permuted distribution and it is therefore not suprising that it would be less likely to produce significant d's. The wider range of d's produced by the bootstrap analysis can be attributed to the fact that resampling with replacement allows outliers to be present multiple times or not at all in any given data set. For example, if an individual whose presence exerts a large effect on the calculation of d is sampled more than once in a given bootstrapped data set then the estimate of d will be higher than could have been achieved if the individual could only be present once in the data set, as is the case in the permutation test.

Thus overall, the bootstrap procedure shows a greater sensitivity to outliers than the permutation test.

Bootstrapping would therefore result in a more conservative test of statistical significance of estimates of d than would a permutation test and therefore may be preferable when small data sets are used.

The restricted data sets were originally produced to control for the effect of non-independant samples. Nestlings originating from host nests in close proximity, less than 200m, were treated as a potential sibling group and a single representative was randomly chosen to complete the data set. If nestlings in the potential sibling groups really were more similar to each other than the rest of the population then the choice of the group's representative should not have mattered. However, I have already shown that certain individuals influenced the value of d much more so than others and hence their presence or absence in the restricted data sets may dictate whether the estimates of d are significant.

In fact, only nestling 19's presence in the restricted data sets appeared to be highly correlated with significant d's (Table 2). It was present in a total of 10 data sets, 9 of which had d's significant at $p < 0.05$ in the permutation analysis, the 10th was nearly so ($p = 0.06$) and no other data sets approached significance ($p > 0.1$).

Similarly, nestling 19 was present in all 4 data sets with significant d's in the bootstrap analyses, the remaining 6 data sets with nestling 19 approached significance ($p < 0.1$) and again no other data set approached significance ($p > 0.1$).

Removal of nestling 19 from the remaining restricted data sets resulted in a loss of significance at the $p < 0.05$ level in each case. Its removal from the full data set caused the estimate of d to drop from 0.204 % to 0.179 % with a corresponding increase in the estimate of p for both the bootstrap (0.060 to 0.342) and permutation analyses (0.005 to 0.125) with the latter case resulting in a loss of significance.

The considerable influence that nestling 19 had on the calculation of d was most likely a result of it's unique genotype. It possessed one of three rare bands and was missing several common bands. The last observation is significant because the inability of RAPDs to detect heterozygotes means that the population frequency of the null allele must be estimated from q^2, or the proportion of individuals without a band, which is a biased estimator when the null allele is rare (Lynch and Milligan 1993), especially when sample sizes are small. Since nestling 19

lacks several common bands and the study's sample size was small, the finding that estimates of d were so dependant on this particular individual is therefore not suprising.

The bootstrap and permutation analyses of RAPD nuclear markers both suggest some degree of differentiation between YW and RWB host races. However, the results are severely compromised due to the effect of outliers on the analyses. Ultimately, the solution to this problem would be to increase the sample size, while attempting to ensure sample independance, in the hopes of gaining a more representative sample from the population and hence greater accuracy in the estimation of population allele frequncies.

Host Specificity ?

Based on the variation detected in the mtDNA control region there was no population structure, with respect to host use, in the population of Brown-headed cowbirds studied. Hence, female host-specific lineages are not likely present, or if present, are too recently evolved to be detected with mtDNA markers.

Interpretation of the analysis of nuclear DNA variation remains unclear. While both randomization procedures detected some degree of population structure, the

finding that one individual had such a profound effect on the significance of estimates of d makes this data set somewhat suspect.

Although the nuclear DNA evidence was inconclusive, it is likely the case that no structure is present here either. It is difficult to imagine a scenario in which mtDNA could be homogeneous while nuclear DNA was structured since females which contributed their mtDNA to different 'host races' also contributed nuclear genes and therefore any nuclear variation present should tend to become homogeneous. However, one possibility is that if RAPD markers evolve at a faster rate than the mtDNA control region, then they might detect existing genetic structure that the control region failed to.

To test this possibility I compared estimates of percent sequence divergence d from the two DNA markers used in this population. Within-population genetic divergence can be calculated directly from sequence data by taking the average pairwise divergence among indivduals. The portion of the control region sequenced in this study yielded a 0.17 % average sequence divergence. Clark and Lanigan's program RAPDDIP (in press) estimates the same parameter for nuclear DNA based on modified bandsharing of RAPD markers. A

jacknifed estimate of the average %sd thus calculated was 0.19 % with a 95 % confidence interval of (0.16 %, 0.22 %). While the RAPD estimate is biased by the presence of outliers, it suggests that the rates of evolution of these two DNA markers in this species are at least comparable. A larger survey of unrelated adult cowbirds would be necessary to confirm this conclusion.

These results lend support to the commonly held belief that Brown-headed cowbirds are, as a species, host generalists (Payne 1977, Rothstein 1990). To address the possibility of individual female specificity, hinted at in studies by Walkinshaw (1949) and McGeen and McGeen (1969), an approach centered on tracking individual female laying patterns (cf. Fleischer 1985) but using the greater resolving power of DNA fingerprinting, would be more appropriate. Individual host specificity detected at this level might signal the begining of host specific lineages not yet detectable with population level analyses.

CONCLUSION

My genetic analysis of Brown-headed Cowbird chicks reared in two different host nests yielded the following results: 1) There was no detectable difference in mtDNA haplotypes between cowbird chicks from different host nests. 2) Results from RAPD nuclear markers suggest weak differentiation but the results are severely compromised due to the effect of outliers on the analyses. Thus, on the basis of the mtDNA data, I provisionally conclude that hypothesis 1, female specific host races and hypothesis 2, cryptic species, are unlikely to be true. However, additional analysis of the variation in RAPD markers using a larger sample size is required to firmly support the remaining hypothesis, ie lack of host specificity.

LITERATURE CITED

Burke, T. and Bruford, M. 1987. DNA fingerprinting in birds. Nature 327: 149-152.

Cann, R.L., Stoneking, M. and Wilson, A.C. 1987. Mitochondrial DNA and human evolution. Nature 325: 31-36.

Clark, K.L. and Robertson, R.J. 1981. Cowbird parasitism and evolution of anti-parasite strategies in the Yellow Warbler. Wilson Bull. 93: 249-258.

Clark, A.G. and Lanigan, C.M.S. 1993. Prospects for estimating nucleotide divergence with RAPDs. Mol. Biol. Evol. (in press).

Darley, J.A. 1983. Territorial behavior of the female Brown-headed Cowbird (Molothrus ater). Can. J. Zool. 61: 65-69.

Davies, N.B. and Brooke, M.D. 1989. An experimental-study of co-evolution between the cuckoo, (Cuculus canorus), and its hosts. 2. Host egg markings, chick discrimination and general discussion. J. Anim. Ecol. 58: 225-236.

Dufty, A.M. 1982. Movements and activities of radio-tracked Brown-headed Cowbirds. Auk 99: 316-327.

Edwards, S.V. 1993. Long-distance gene flow in a cooperative breeder detected in genealogies of mitochondrial DNA sequences. Proc. R. Soc. Lond. Series B. 252: 177-185.

Fleischer, R.C. 1985. A new technique to identify and assess the dispersion of eggs of individual brood parasites. Behav. Ecol. Sociobiol. 17: 91-99.

Friedmann, H. 1929. *The cowbirds, a study in the biology of social parasitism*. Thomas Springfield, Ill.

Friedmann, H., Kiff, L.F. and Rothstein, S.I. 1977. A further contribution to knowledge of host relations in the parasitic cowbirds. Smithson. Contrib. Zool. 235: 1-75.

Greenberg, B.D., Newbold, J.E. and Sugino, A. 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in the human mitochondrial DNA. Gene 21: 33-49.

Gelter, H.P., Gibbs, H.L. and Boag, P.T. 1993. Large deletions in the control region of Darwin Finch mitochondrial DNA: evolutionary and functional implications. Proc. Natl. Acad. Sci. USA. (in press).

Gyllensten, U.B. and Erlich, H.A. 1988. Generation of single-stranded DNA by the polymerase chain reaction and it's application to direct sequencing of the HLA-DQα locus. Proc. Natl. Acad. Sci. USA. 85: 7652-7656.

Hadrys, H., Balick, M. and Scierwater, B. 1992. Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. Mol. Ecol. 1: 55-63.

Kazan, K., Manners, J.M. and Cameron, D.F. 1993. Genetic relationships and variation in the Stylosanthes guianensis species complex assessed by random amplified polymorphic DNA. Genome 36: 43-49.

Lynch, M. and Milligan, B.G. 1993. Analysis of population genetic structure with RAPD markers. Mol. Biol. Evol. (in press).

Manly, B.F.J. 1991. Randomization and Monte Carlo methods in biology. Chapman and Hall, New York.

McGeen, D.S., McGeen, J.J. 1968. Cowbirds of Otter lake. Wilson Bull. 80: 84-93.

Nei, M. and Li, W.-H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA. 76: 5269-5273.

Payne, R.B. 1977. The ecology of brood parasitism in birds. Annu. Rev. Ecol. Syst. 8: 1-28.

Payne, R.B. 1973. Vocal mimicry of the paradise whydahs (Vidua) and response of female whydahs to the songs of their hosts (Pytilia) and their mimics. Anim. Behav. 21: 762-771.

Quinn, T.W. 1992. The genetic legacy of Mother Goose-phylogeographic patterns of lesser snow goose Chen caerulescens caerulescens maternal lineages. Mol. Ecol. 1: 105-117.

Roff, D.A. and P. Bentzen. 1989. The Statistical Analysis of Mitochondrial DNA Polymorphisms: X^2 and the Problem of Small Samples. Mol. Biol. Evol. 6:539-545.

Rothstein, S.I. 1990. A model system for coevolution: Avian brood parasitism. Annu. Rev. Ecol. Syst. 21: 481-508.

Russell, J.R., Hosein, F., Johnson, E., Waugh, R. and Powell, W. 1993. Genetic differentiation of cocoa (Theobroma cacao L.)populations revealed by RAPD analysis. Mol. Ecol. 2: 89-87.

Sanger, F., Nicklen, S. and Coulson, A.R. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74: 5463-5467.

Seutin, G., White, B.N., Boag, P.T. 1991. Preservation of avian blood and tissue samples for DNA analyses. Can. J. Zool. 69: 82-90.

Southern, H.N. 1954. Mimicry in cuckoo eggs. In *Evolution as a Process*. ed. J. Huxley, A.C. HArdy, E.B. Ford, pp. 219-232. London: George Allan and Unwin.

Teather, K.L. and Robertson, R.J. 1985. Female spacing patterns in Brown-headed Cowbirds. Can. J. Zool. 63: 218-222.

Tukey, J.W. 1977. *Exploratory Data Analysis*. Addison-Wesley; MA. 500 pp.

Walkinshaw, L.H. 1949. Twenty-five eggs apparently laid by a cowbird. Wilson Bull. 61: 82-85.

Weatherhead, P.J. 1989. Sex-ratios, host-specific reproductive success, and impact of Brown-headed Cowbirds. Auk 106:358-366.

Welsh, J., Petersen, C. and McClelland, M. 1990. Polymorphisms generated by arbitrarily primed PCR in the mouse: application to strain identification and genetic mapping. Nucleic Acids Res. 19: 303-306.

Wenink, P.W., Baker, A.J. and Tilanus, M.G.J. 1993. Hypervariable control region sequences reveal population structuring in a long-distance migrant shorebird, the Dunlin (Calidris alpina). Proc. Natl. Acad. Sci. USA. 90: 94-98.

Westneat, D.F. 1990. Genetic parentage in the indigo bunting: a study using DNA fingerprinting. Behav. Ecol. Sociobiol. 27: 67-76.

Wetton, J.H., Carter, R.E., Parkin, D.T., Walters, D. 1987. Demographic study of a wild house sparrow population by DNA fingerprinting. Nature 327: 147-149.

Williams, J.G.K., Kubelik, A.R., Kenneth, J.L., Rafalski, J.A., Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18: 6531-6535.

Zaykin, D.V. and A.I. Pudovkin. (in press) Two programs to estimate Chi-square values using pseudo-probability test. J.Hered.

Table 1 Variation in control region sequence among 31 nestling cowbirds from two host species, Red-winged Balckbird (RWB) and Yellow Warbler (YW). Base positions are relative to the reference cowbird sequence from Miller and Gibbs (in prep.).

Haplotype	Bases at 3 variable positions			Number of Individuals	
	187	526	600	RWB	YW
1	C	C	-	6(4)	11(7)
2	C	C	A	5(3)	4(4)
3	T	C	-	1(1)	1(1)
4	T	T	-	1(0)	2(1)

Note: The numbers in parentheses represent individuals sharing a haplotype in the restricted data sets.

Table 2. Divergence estimates for twenty restricted data sets including P-values for both bootstrap and permutation analyses. The data sets were restricted by selecting only one nestling from each potential sibling group, where potential siblings are defined as groups of nestlings originally laid in nests less than 200m apart.

	Sibling Groups														
Data set	35	42	12	40	49	33	34	19	44	25	37	46	<d>	P(boot)	P(perm)
1		X	X			X			X	X			0.224%	0.393	0.535
2	X				X		X	X			X		0.194%	0.076	0.036 *
3		X		X		X		X				X	0.198%	0.084	0.004 *
4	X			X			X		X	X			0.189%	0.324	0.374
5		X	X				X		X	X			0.179%	0.168	0.157
6	X		X			X			X		X		0.164%	0.458	0.648
7		X			X	X			X			X	0.166%	0.534	0.786
8	X				X		X	X			X		0.205%	0.163	0.060
9		X		X			X		X			X	0.190%	0.119	0.133
10	X				X	X		X				X	0.215%	0.085	0.015 *
11		X	X			X		X		X			0.232%	0.012 *	0.002 *
12	X		X				X		X		X		0.165%	0.162	0.204

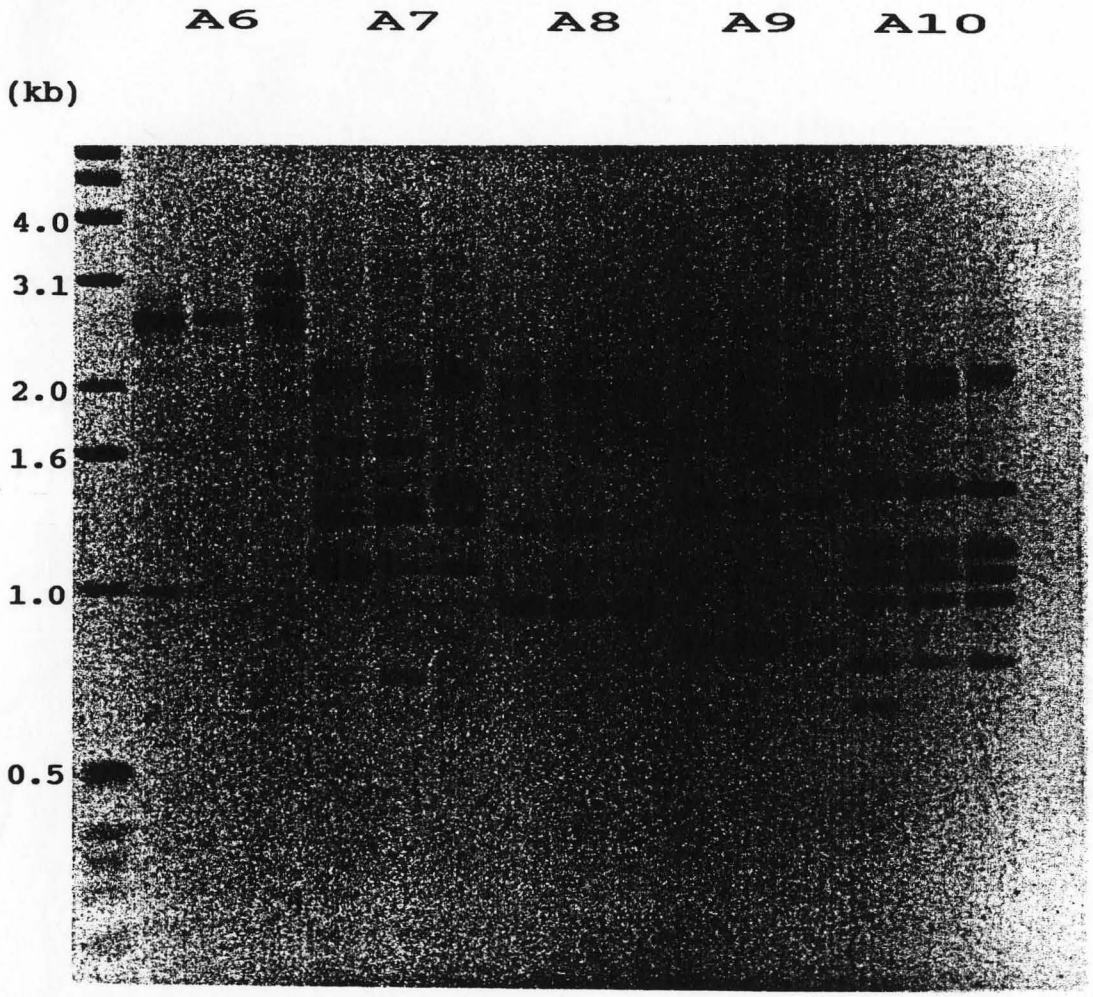
Table 2 (continued)

Data set	35|42	12|40|49	33|34	19|44	25|37|46	<d>	P(boot)	P(perm)
13	X	X	X	X	X	0.200%	0.040 *	0.018 *
14	X	X	X	X	X	0.193%	0.601	0.700
15	X	X	X	X	X	0.232%	0.012 *	0.000 *
16	X	X	X	X	X	0.212%	0.090	0.018 *
17	X	X	X	X	X	0.186%	0.294	0.375
18	X	X	X	X	X	0.202%	0.053	0.013 *
19	X	X	X	X	X	0.245%	0.018 *	0.000 *
20	X	X	X	X	X	0.189%	0.303	0.417

Note: The 'X's mark the individual, from each potential sibling group, which was selected to make up the data set. Individuals 29 and 30 were removed from all restricted data sets since the former's haplotype was unknown and the latter's nest of origin, a YW nest, was not mapped. There was only one potential sibling group which consisted of both a YW (18) and a RWB (32) reared nestling, in this case the RWB individual was chosen for the restricted data sets to increase the RWB sample size. The asterisks denote P-values which were significant at the 0.05 level. P-values for both analyses were based on 1000 randomizations.

Figure 1 : RAPD profiles of Brown-headed Cowbirds. Figures (a) and (b) are duplicate RAPD amplification of the same 3 individuals (47, 48 and 49) with five (A6 to A10) of the 10 primers used in this study. Only those bands which were reproducible and could be scored unambiguously in each individual were used to make up the data set (see Appendix A for a list of scored bands and raw data). A 1 kb size marker (BRL) was run in the far left lane with labelled bands in kilo-base pairs (kb).

a)



b)

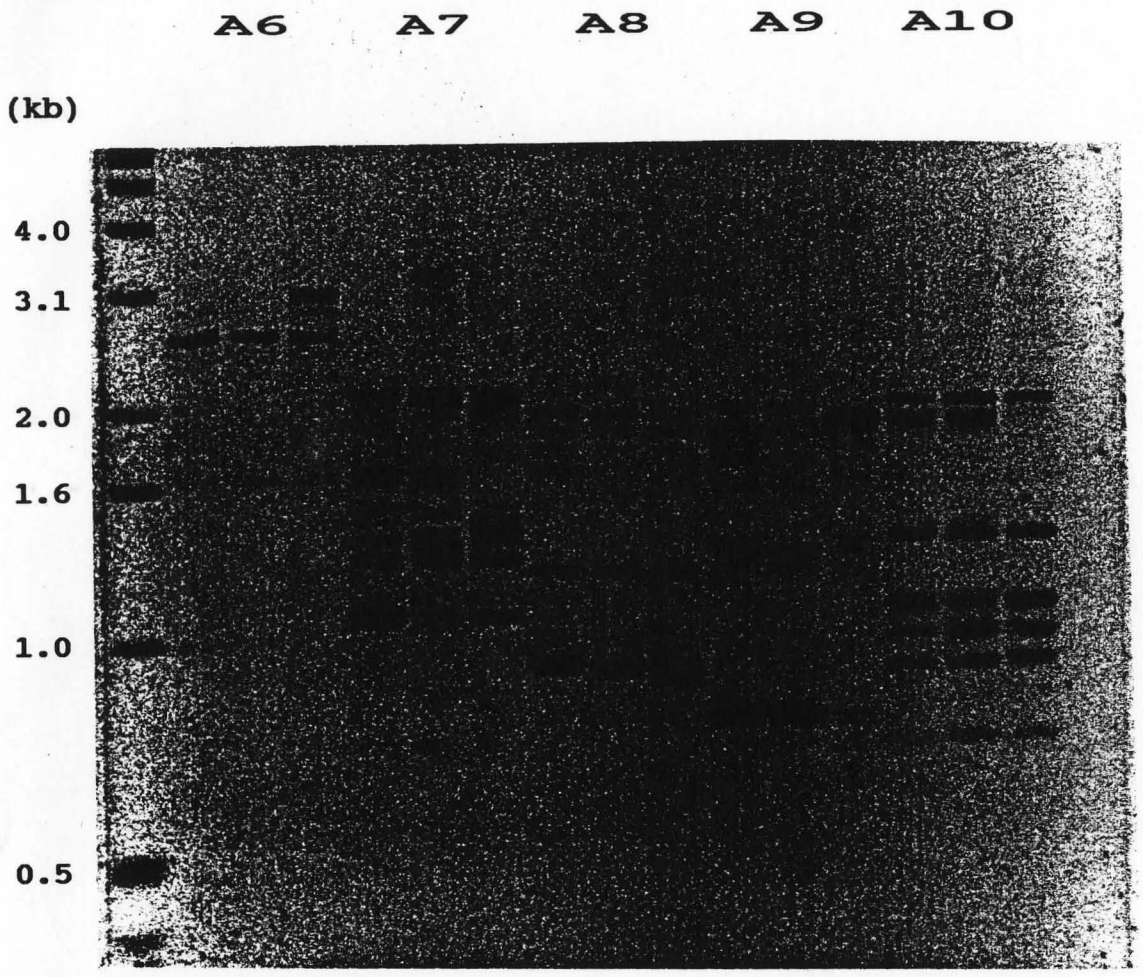
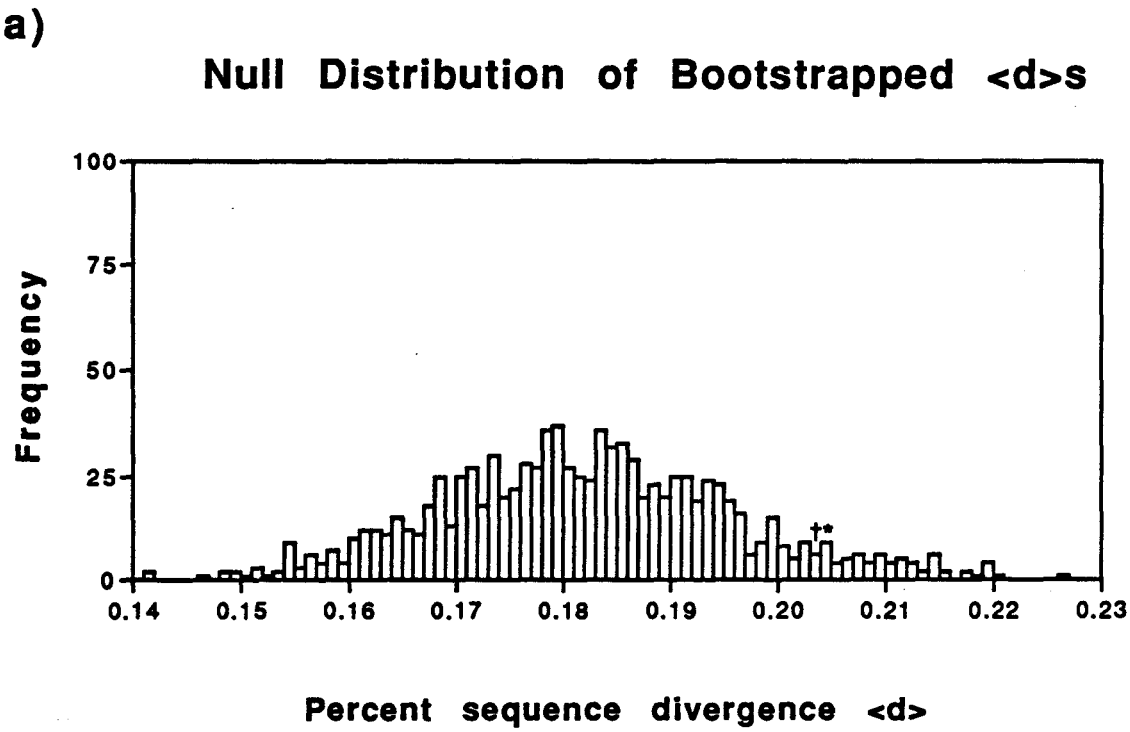
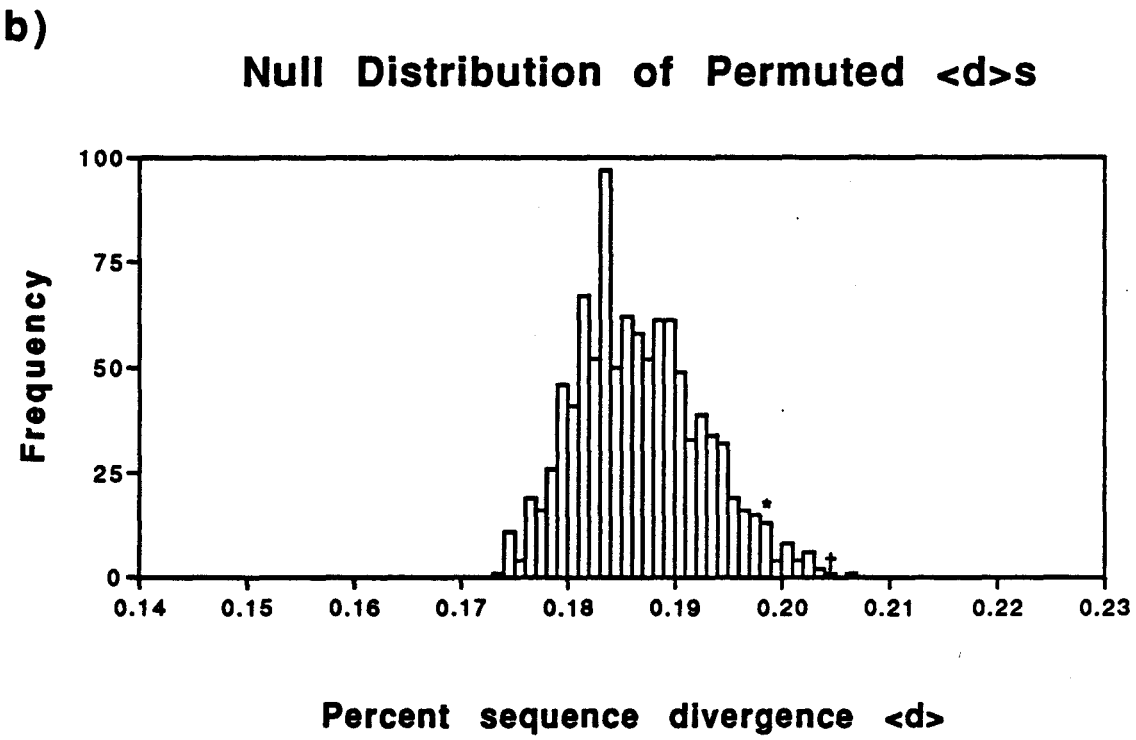


Figure 2: Frequency distributions of d values from (a) bootstrapped and (b) permuted data sets based on 1000 randomizations. The 95th percentiles are marked by astericks and correspond to d values of 0.205% and 0.198% for the bootstrapped and permuted null distributions respectively. The observed d of 0.204%, denoted by a cross, was within the 94th percentile of the bootstrapped null distribution and within the 99.5th percentile of the permutted null distribution.

a)


Null Distribution of Bootstrapped <d>s
Frequency
100
75
50
25
0
0.14
0.15
0.16
0.17
0.18
0.19
0.20
0.21
0.22
0.23
Percent sequence divergence <d>

b)


Null Distribution of Permuted <d>s
Frequency
100
75
50
25
0
0.14
0.15
0.16
0.17
0.18
0.19
0.20
0.21
0.22
0.23
Percent sequence divergence <d>

Appendix 1. RAPD raw data and mtDNA haplotypes. For the RAPD data 1/0 = presence/absence of a band. See Table 1 for mtDNA haplotype sequences.

Host: Yellow Warbler

Primer-Size(kb)	12	13	18	20	21	26	27	28	29	30	35	39	40	41	42	45	47	48	49
A01-0.6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A01-1.3	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0
A01-1.6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A02-1.1	0	1	1	1	1	0	1	0	0	0	0	0	1	1	1	0	1	1	1
A04-0.8	0	1	1	0	1	1	1	1	0	0	0	0	1	1	0	1	1	1	1
A04-1.3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A04-2.0	1	0	0	1	0	0	1	0	0	1	0	0	0	1	1	1	0	0	1
A06-1.0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
A06-2.8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A06-2.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A06-3.0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1
A07-1.2	1	1	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
A07-1.4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A07-1.5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A07-1.7	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	1	0
A07-2.1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A08-1.4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A08-2.0	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	0	0	0	0
A09-0.8	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0
A09-1.3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A09-2.0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
A10-0.5	0	0	0	0	1	1	0	0	0	0	1	0	1	0	0	0	0	0	0
A10-0.6	0	0	1	0	1	0	1	1	1	0	1	0	0	1	0	1	1	0	0
A10-0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-0.9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-1.0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-1.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-1.3	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
A10-2.0	1	1	0	1	1	1	0	1	1	1	1	1	0	1	0	1	1	1	0
A10-2.1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Haplotypes	1	2	4	4	2	1	2	2	?	1	1	1	1	1	1	1	1	3	1

Appendix 1 (continued)

Host: Red-winged Blackbird

Primer-Size(kb)	19	24	25	31	32	33	34	36	37	38	43	44	46
A01-0.6	1	1	1	1	1	1	1	1	1	1	1	1	1
A01-1.3	0	1	1	1	1	0	1	1	1	0	1	1	0
A01-1.6	1	1	1	1	1	1	1	1	1	1	1	1	1
A02-1.1	1	0	0	0	0	0	1	0	0	0	0	0	0
A04-0.8	1	1	1	1	1	1	1	1	1	1	1	1	1
A04-1.3	1	1	1	1	1	1	1	1	1	1	1	1	1
A04-2.0	0	0	0	1	0	0	1	0	0	1	0	0	1
A06-1.0	0	1	1	1	0	0	1	1	1	1	1	1	1
A06-2.8	1	1	1	1	1	1	1	1	1	1	1	1	1
A06-2.9	1	1	1	1	1	1	1	1	1	1	1	1	1
A06-3.0	0	0	0	0	0	1	1	1	0	0	0	0	0
A07-1.2	1	1	0	0	0	0	0	1	0	0	0	1	0
A07-1.4	1	1	1	1	1	1	1	1	1	1	1	1	1
A07-1.5	1	1	1	1	1	1	1	1	1	1	1	1	1
A07-1.7	1	0	0	0	0	0	0	0	0	0	0	0	0
A07-2.1	1	1	1	1	1	1	1	1	1	1	1	1	1
A08-1.4	1	1	1	1	1	1	1	1	1	1	1	1	1
A08-2.0	0	1	1	1	0	1	1	0	1	0	0	0	1
A09-0.8	1	1	1	0	1	1	1	1	1	1	1	1	1
A09-1.3	1	1	1	1	1	1	1	1	1	1	1	1	1
A09-2.0	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-0.5	0	0	0	0	0	1	0	1	0	0	0	1	1
A10-0.6	1	0	1	0	1	0	0	0	1	0	0	1	0
A10-0.7	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-0.9	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-1.0	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-1.1	1	1	1	1	1	1	1	1	1	1	1	1	1
A10-1.3	1	1	1	1	1	1	1	1	1	1	0	1	1
A10-2.0	0	1	1	1	1	1	0	1	1	1	0	0	0
A10-2.1	1	1	1	1	1	1	1	1	1	1	1	1	1
Haplotype	1	2	2	1	4	1	1	3	2	2	1	1	2