Double-strand conformation polymorphism (DSCP) analysis of the mitochondrial control region generates highly variable markers for population studies in a social insect

L. Atkinson and E. S. Adams

Department of Biology, University of Rochester, Rochester, New York, USA

Abstract

Genetic markers were obtained for the termite Nasutitermes corniger by DSCP (double-strand conformation polymorphism) analysis of PCR-amplified mitochondrial control region DNA. This procedure revealed twenty-one haplotypes in forty-four colonies, whereas a restriction fragment length polymorphism analysis detected only nine haplotypes. Sequence analysis of DSCP fragments of contrasting mobilities suggests that the electrophoretic haplotypes are caused by DNA curvature in this highly AT-rich region. DSCP markers showed that some termite colonies contained maternally unrelated queens, each of which produced worker offspring. This pattern is consistent with nest founding by unrelated queens. Due to the availability of conserved primers for the mtDNA control region, DSCP analysis has the additional advantage of fast and inexpensive implementation.

In this paper we demonstrate that the DSCP technique is useful for examining the breeding structure of termite colonies. In the termite Nasutitermes corniger, colonies often contain multiple queens and kings (polygamy). Two different origins of these multiple reproductives have been hypothesized. Multiple reproductives can be reared within the colony following the death of the original founders or during the formation of nest buds. Kings and queens produced in this manner are termed ‘replacement’ or ‘supplementary’ reproductives. Alternatively, colonies may acquire multiple reproductives by cooperative nest founding, by adoption, or by colony fusion (Roisin, 1993; Leponce et al., 1996).

The two alternative causes of polygamy produce contrasting genetic structures, with different evolutionary implications. It is usually not possible to distinguish between the two classes of multiple reproductives by morphological differences, but genetic analyses employing highly variable mtDNA markers can readily discern these routes to polygamy. If multiple queens and kings arise only by replacement reproduction, then they will share the same mitochondrial DNA haplotype. If multiple reproductives are acquired by other routes, they will often have different haplotypes. Despite numerous studies on the social Hymenoptera, genetic...
evidence characterizing colony foundation processes is lacking in termites.

The application of genetic markers to assess relatedness in social insect colonies is particularly crucial because of the importance of kin selection theory in models of the evolution and maintenance of social behaviour (Crozier & Pamilo, 1996). Mitochondrial markers can be used to distinguish between alternative origins of queens and kings, to determine maternal relatedness, or to quantify reproductive skew (Reeve & Ratnieks, 1993). Although mtDNA markers do not reveal paternity, they can be used in conjunction with other markers to estimate coefficients of relatedness and other aspects of colony genetic structure.

In this paper we describe the markers used for determining maternal relatedness in multiple-queen colonies of *N. corniger* (Order Isoptera). By sequencing DNA from individuals of various haplotypes we establish the nucleotide basis for different electrophoretic haplotypes and show that the differences in electrophoretic mobility probably result from sequence-directed changes in DNA curvature (Hagerman, 1990). The variation detected by DSCP analysis is compared to that found by restriction fragment length polymorphism (RFLP) analysis. In addition, we demonstrate that DSCP analysis can be readily applied to other species, with examples from two taxonomically diverse orders (Diptera and Hymenoptera).

**Results and Discussion**

**Characterization of *N. corniger* haplotypes**

We analysed 271 queens and 147 kings from forty-four natural colonies and six additional colonies that contained replacement reproductives induced by artificial orphaning. Because reproductives within social insect colonies can be related, individual queens and kings may not always represent different matrilines. To obtain an estimate of the number of independent matrilines in our sample, we assumed that the reproductives in monogamous colonies derived from different matrilines because, within pairs, the queens and kings almost always had different haplotypes. However, multiple occurrences of the same haplotype within a polygamous colony’s reproductives can result from replacement reproduction. Therefore different haplotypes were counted only once in each polygamous colony and additional individuals that shared these haplotypes were assumed to have a common origin in the colony. Based on these assumptions, we sampled eighty-three independent matrilines.

A ∼2.1 kb mtDNA region, comprising the tRNA genes for methionine, isoleucine and glutamine, the ∼1500 bp control region, and 367 bases of the 5'-end of the 12s rRNA gene, was PCR-amplified with conserved primers in the 12s rRNA and methionine tRNA genes. Cytochrome c oxidase subunit I and II (COI/II) primers amplified ∼1.5 kb of the cytochrome c oxidase I and II genes. A restriction fragment length polymorphism (RFLP) analysis of the mtDNA control region and COI/II gene PCR products from the total sample of reproductives revealed nine haplotypes, including two common haplotypes and seven rare haplotypes, each found in only one or two colonies. The two common haplotypes (*HinfI/Ddel* 'A' and 'B') differed by three *HinfI* restriction sites in the control region and one *Ddel* restriction site in the COI/II genes. The seven rare haplotypes differed from either *HinfI/Ddel* 'A' or 'B' by one additional restriction site.

Coincidentally, DSCP analysis of the uncut 2.1 kb mtDNA control region PCR products also revealed nine haplotypes. Haplotypes could be classified by electrophoretic mobility into two distinct groups corresponding to the *HinfI/Ddel* 'A' or 'B' RFLP lineages, defined above. Within these two groups, additional small mobility differences were observed; however, most of the DSCP haplotypes obtained from analysis of uncut PCR product did not correspond to the rare RFLP haplotypes.

We obtained better resolution of the 2.1 kb control region PCR product by digesting it with a restriction enzyme prior to DSCP analysis, which allowed greater separation of the resulting small fragments. Restriction enzymes that cut the PCR product into ∼200–800 bp fragments produced the greatest mobility differences. The *Apol* digest gave the best resolution in *N. corniger*. DSCP analysis using other enzymes besides *Apol* revealed no additional haplotypes; therefore we used *Apol* for all our analyses. DSCP analysis of the mtDNA control region PCR products digested with *Apol* revealed a total of twenty-one haplotypes among the reproductives in 44 colonies (Fig. 1). All RFLP haplotypes appeared as distinct DSCP haplotypes; DSCP analysis distinguished twelve additional haplotypes that were not detected by RFLP analysis. The most common haplotype occurred in twenty-one of the eighty-three matrilines; ten haplotypes were seen in only one colony each. All differences between haplotypes were consistent and reproducible. MtDNA population structure will be described in future work.

The *Apol* digest of the control region PCR product yielded five restriction fragments. Although the *Apol* RFLP pattern from all *N. corniger* individuals appeared identical on agarose gels, restriction fragments varied in mobility when electrophoresed through acrylamide gels for the DSCP procedure, allowing the identification of haplotypes. For most haplotypes, the largest...
mobility differences were found in the \(\approx 558\) bp Apol restriction fragment (Fig. 1).

Sequencing the \(\approx 558\) bp Apol restriction fragment from five queens that displayed varying degrees of mobility difference (see Fig. 1) revealed that the electrophoretic mobility haplotypes were due to conformation rather than length differences. The fragment varied by nought to fourteen nucleotide substitutions and one to six insertions or deletions (indels) of 1–2 bp over a total length of 557–558 bp (Table 1 and Fig. 2). Double digests allowed us to locate more precisely the regions within the \(\approx 558\) bp Apol fragment that were responsible for the observed mobility differences. The 237 bp EcoRV/Apol fragment (Fig. 2, nucleotides 1–237) showed the same degree of mobility variation as the \(\approx 558\) bp Apol fragment. All other double-digest restriction fragments that included this region also displayed mobility differences similar in magnitude to those in the larger \(\approx 558\) bp Apol restriction fragment. The Haelll and XbaI digests further pinpointed the region responsible for mobility differences. By cutting next to an insertion site at nucleotide 91 (Fig. 2), these digests eliminated the mobility difference between queens 4 and 5, whereas the mobility patterns of queens 1–3 remained unchanged.

The largest mobility differences were among queens 1–3, which are HinfI/Ddel RFLP haplotype ‘B’, and between these queens and queens 4 and 5, which are HinfI/Ddel RFLP haplotype ‘A’ (see Fig. 1). These differences are caused by single base substitutions or indels at three sites (positions 182, 186 and 196; double underlined in Fig. 2) within a series of six A or T tracts (shown in bold, Fig. 2). A smaller mobility difference that differentiates queens 4 and 5 is caused by an insertion at nucleotide 91 (Fig. 2).

The electrophoretic mobility haplotypes in \(N.\) corniger appear to be caused by sequence-directed curvature of double-stranded DNA. As seen in the \(N.\) corniger DNA fragments, the electrophoretic mobilities of some double-stranded DNA fragments in polyacrylamide gels cannot be predicted solely on the basis of their size. This phenomenon has long attracted the attention of biochemists (reviewed in Trifonov, 1985; Hagerman, 1990). Often, fragments with anomalous rates of migration contain a series of oligo-(d(A),d(T)) tracts (‘A-tracts’) aligned approximately in phase with the DNA helix repeat (reviewed in Trifonov, 1985; Hagerman, 1990). These A-tracts cause intrinsic curving of the DNA helix; nucleotides flanking or interrupting A-tracts alter this curvature and affect the electrophoretic mobility of the DNA molecules in polyacrylamide gels (Koo et al., 1986; Nagaich et al., 1994). As shown above, a series of A-tracts spaced approximately every ten bases is found in the \(N.\) corniger sequence that is responsible for the greatest mobility differences, and mutations in these A-tracts account for the differences among haplotypes (Fig. 2). A-tracts and loci of curvature have been found in the mtDNA origins of replication of several other organisms (e.g. Marini et al., 1982; Singh et al., 1987; Welter et al., 1989). In addition to A-tracts, other dinucleotide ‘wedges’ that occur with a period close to the helical repeat of DNA can also produce DNA curvature (Boisshoy et al., 1991).

Besides the few indels or substitutions that affect the electrophoretic mobility of fragments in \(N.\) corniger, sequencing revealed other nucleotide substitutions and indels associated with each haplotype that appear to have no effect on conformation. Although it is therefore likely that DSCP analysis will miss some mutations, sequencing \(\approx 350\) bp of the 12s rRNA gene and control region from two to three termites from each of six DSCP haplotypes collected in different geographic

---

**Table 1.** Sequence differences in the 557–558 bp Apol restriction fragment from the mtDNA control region PCR product for the five \(N.\) corniger queens shown in Figs 1 and 2. The number of base substitutions between each queen pair are given, with the number of insertions and deletions (all limited to 1–2 bp in length) shown in parentheses. The greatest differences in base composition and mobility are between members of the ‘A’ and ‘B’ HinfI/Ddel RFLP lineages (queens 4 and 5, and queens 1–3, respectively).

<table>
<thead>
<tr>
<th>Queen</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>–</td>
<td>2 (1)</td>
<td>5 (2)</td>
<td>12 (4)</td>
<td>12 (6)</td>
</tr>
<tr>
<td>2</td>
<td>–</td>
<td>5 (1)</td>
<td>14 (3)</td>
<td>14 (5)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>–</td>
<td>14 (4)</td>
<td>14 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>–</td>
<td>–</td>
<td>0 (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

---

*Insect Molecular Biology* 6: 369–376
Figure 2. Alignment of the 557–558 bp ApoI restriction fragment DNA sequence from five *N. corniger* queens. Sequences encompass the 5' end of the 12s rRNA gene and the adjacent portion of the control region. A dot indicates a base identical to that of queen 1; a dash indicates that the nucleotide is absent. Restriction sites are underlined. The A and T tracts are in bold. Double underlining indicates the nucleotide positions responsible for mobility differences among haplotypes. The RFLP haplotypes of queens 1–3 are *HindIII* (B/E); quee 4 and 5 are *HindIII* (A). Differences between sequences are summarised in Table 1 and DSCP haplotypes are shown in Fig. 1, lanes 2–6. GenBank accession numbers for the sequences of queens 1–5 are U71289–U71292.
areas (5–60 km apart) failed to distinguish additional haplotypes.

Double-strand conformation polymorphisms (DSCPs) as genetic markers

The high haplotype diversity revealed in *N. corniger* by DSCP analysis is useful for examining this termite's breeding structure and the mechanisms by which multiple reproductives arise in colonies. In six of forty-four naturally occurring colonies, we obtained clear evidence of maternally unrelated queens within the same colony. In several of the colonies the queens were similar in size and coloration, indicating they were probably the same age. In at least these cases it appears that colonies were founded by two or more unrelated queens. Each queen produced offspring. The DSCP haplotypes of the queens and some of the workers from a two-queen colony are shown in Fig. 3. Detailed description of the termite study and the colony structures revealed by DSCP haplotype patterns are given elsewhere (Atkinson & Adams, 1997). The apportionment of reproduction among queens will be described in future work.

In our study DSCP analysis detected as many haplotypes as were obtained by sequencing several hundred bp per individual, at a much lower cost. DSCP markers were 2.3 times more variable than conventional RFLP markers. Furthermore, DSCP analysis offers a fast way to obtain markers. Its utility can be quickly evaluated because DSCP analysis can be performed on PCR products of any length (1.6–4 kb in this study; see below). This versatility allows the use of PCR products amplified with conserved primers. To obtain markers with other equally sensitive DNA screening methods, it would be necessary to identify, and often to sequence and design primers for, a variable DNA region of <500 bp (Lessa & Applebaum, 1993). SSCP analysis (following Hiss et al., 1994) of the large PCR products used in this study did not yield any detectable polymorphism; digestion of these PCR products to obtain smaller fragments resulted in numerous poorly resolved bands.

The DSCP technique can be applied to a wide range of species. The 12s rRNA gene and control region of the mitochondrial genome harbour sequence variation detectable by DSCP analysis. Other DNA regions also display polymorphic DNA curvature (reviewed in Hagerman, 1990), but the one other region that we tested, the mtDNA cytochrome *c* oxidase subunit I and II genes in *N. corniger* (both uncut and digested with restriction enzymes), did not generate DSCP markers, even among PCR products with known restriction site differences. DSCP analysis is most likely to be useful in sequences that have *A*-tracts. Fortunately, the noncoding control region is highly variable and AT-rich in most insect species, and conserved PCR primers in regions flanking the control region are readily available (Simon et al., 1994). We applied the DSCP technique to two other species and detected seven haplotypes among seven *Drosophila falleni* flies and four haplotypes in five colonies of the ant *Pogonomyrmex barbatus*.

As with all methods of detecting genetic variation, there are shortcomings to DSCP analysis. It is clear that some mutations do not affect DNA curving and cannot be distinguished by DSCP analysis. Therefore this technique will not work with all PCR products. In addition, the large number of haplotypes that are defined by small mobility differences requires electrophoresing some PCR products more than once to place fragments with similar mobility next to one another. Furthermore, although it is possible to make a reasonable prediction of curvature from known sequences (e.g. Bolshoy et al., 1991), there are a large number of sequences that can produce any given electrophoretic pattern. Therefore, because the underlying sequence differences cannot be estimated, it is not possible to order lineages for phylogenetic purposes. However, even in phylogenetic studies, DSCP analysis may be useful in identifying haplotypes for sequencing. Also, there are potential difficulties inherent to any study that uses mtDNA markers, including biparental inheritance (Avise, 1991), heteroplasmy (Harrison, 1989; Rand,
and nuclear copies of mtDNA (Zhang & Hewitt, 1996a, b).

In summary, DSCP analysis is a rapid and efficient method to detect variation in the control region of mtDNA. The level of variation detected by DSCP analysis of control region mtDNA can provide markers for examining genetic identity, maternity, kinship, family structure, and other aspects of population genetics that require variation from a rapidly evolving region.

**Experimental procedures**

**DNA extraction and amplification**

Reproductives of the termite *Nasutitermes corniger* were collected from forty-four colonies near Colon, Panama City and Portobelo in the Republic of Panama (Atkinson & Adams, 1997). Six of these colonies were resampled 1 year after removal of the original reproductives (artificial orphaning). DNA was extracted from a 1–2 mg portion of the abdominal cuticle of each termite using the Puregene DNA isolation kit (Genta Systems Inc.). The mtDNA control region, with parts of the flanking 12s rRNA and tRNA genes, was PCR-amplified with conserved primers (Kocher et al., 1989; Simon et al., 1994) in the methionine tRNA and 12s rRNA genes (primers from Taylor et al., 1993). PCR reactions contained 20 nM Tris-HCl (pH 8.4), 50 mM KCl, 2 mM MgCl₂, 200 μM of each dNTP, 100 nM of each primer, 2.0 units of *Taq* DNA polymerase (Gibco BRL, Life Technologies Inc.) and 0.1–1 μg of template DNA in a total volume of 100 μl. Amplification conditions were twenty-five cycles of 1 min at 94°C, 1 min at 55°C, and 2 min at 72°C. In addition, portions of the cytochrome c oxidase subunit I and II genes were amplified using the same PCR conditions except that the annealing temperature was 37°C and the PCR was conducted for thirty-five cycles (primers C1-J-2195 and C2-N-3661, minus the 3'-base; Simon et al., 1994).

**RFLP analysis**

To screen for restriction fragment length polymorphisms, we digested pooled PCR products (4 μl PCR product from each of five individuals) with each of twenty-one restriction enzymes (Apol, Asel, BamH, Bsr, Ddel, DraI, DraII, EcoRV, HaeIII, HhaI, HindIII, HinflI, Msel, MspI, PstI, PvuII, Rsal, SapI, XbaI and XbaI) following the supplier's instructions (New England Biolabs). In total, we screened mtDNA control region and COI/II gene PCR products from 271 queens and 147 kings. The digested PCR products were electrophoresed through 12% agarose gels, and visualized under UV light after ethidium bromide staining. Restriction enzymes that cut at polymorphic sites were subsequently used to digest PCR products from queens and kings individually.

**DSCP analysis**

For DSCP analysis, we used 0.4 mm, 16 x 18 cm, 6% polyacrylamide vertical gels (49:1 acrylamide:bisacrylamide) with 5% glycerol. We also tested 5% and 7% polyacrylamide gels, each with both 49:1 and 19:1 acrylamide:bisacrylamide concentrations, but these resulted in lower resolution. Protocols followed those described for SSCP analysis (Hiss et al., 1994), except that we applied double-stranded PCR product to gels of a 6% rather than 5% acrylamide concentration. Because we developed the DSCP method independently, protocols differ from those described in Saad et al. (1994). Uncut samples (2 μl PCR product with 1.5 μl 6 x loading buffer I; Sambrook et al., 1989) or previously digested samples (2 μl PCR product, 0.4 μl enzyme buffer, 0.07 μl restriction enzyme (0.28 units; New England Biolabs), 0.07 μl BSA and 2 μl ddH₂O, with 1.5 μl 6 x loading buffer I) were loaded onto gels using a 32-well sharktooth comb. Electrophoresis was carried out in Gibco BRL model V-16 gel boxes (Life Technologies Inc.) at room temperature for 24 h at 300 V (uncut product) or 85 V (digested product). Gels were then silver stained using either a kit (Bio-Rad Laboratories Inc.) or the protocol described in Merril et al. (1981). To distinguish similar haplotypes, it was sometimes necessary to electrophorese PCR products through more than one gel, so that fragments of similar mobility could be placed in adjacent lanes for clarification of small differences.

Several tests were performed to examine the inheritance patterns of the mtDNA markers. Strict maternal inheritance of haplotypes was demonstrated by analysing mothers and offspring from ten single-queen colonies. Consistent inheritance over at least three generations was confirmed by analysing queens removed from field colonies, the replacement reproductives reared following the removal, and the offspring of the replacement reproductives. PCR products amplified from DNA extracted separately from termite legs and abdomens produced the same haplotype patterns, verifying that the amplified DNA did not derive from gut endosymbionts.

**Sequence and fragment analysis**

To examine the nucleotide basis for the DSCP haplotype differences, we sequenced the part of the control region that displayed the greatest mobility differences, the ± 558 bp *Apol* fragment. PCR products were gel purified and cloned into an mp18 derivative designed for PCR-amplified DNA (Burke et al., 1995). The cloned PCR products were sequenced by the dideoxy chain termination method (Sanger et al., 1977) using the Sequenase kit (United States Biochemical) with the −40 universal primer (United States Biochemical) and internal primer 5′-CACAAGTATAATACGCG-3′ (forward). The opposite strand was sequenced using an automated sequencer with primers 5′-TAAGGATGTAGTTAGGCG-3′ (reverse) and 5′- TTTGTTCATCTG/A/GTTACAC-3′ (reverse). In addition, to examine other individuals with the same haplotypes as the five queens that were sequenced as described above, we sequenced approximately 350 bases of the same mitochondrial region from ten more queens and kings using the primer 5′-CACAAGTATAATACGCG-3′. Sequencing reactions were performed in a Perkin-Elmer thermal cycler following the standard protocol for the ABI PRISM sequencing kit (Perkin-Elmer Corp.). Excess dye terminators were removed by extraction in Centri-sep spin columns (Princeton Separations) prior to loading on an ABI 373 automated sequencer (Applied Biosystems). All sequences were aligned by eye using the computer program ESEE (Cebot & Beckenbach, 1989).

To further isolate the region responsible for mobility differences, we sequenced mtDNA control region PCR product with *Apol* and each of the following restriction enzymes: *Haell*, *EcoRV*, *MspI*, *XbaI*, *Hinf*I and *XbaI*, according to supplier's instructions (New England Biolabs Inc.). Each additional enzyme digested the
Apo fragment into two or three fragments, and did not cut the other Apo fragments. Products were resolved using the DSCP procedure.

### DSCP analysis of other species

We performed DSCP analysis on digested mtDNA control region PCR product from seven individuals of the fly *Drosophila fallensi* (Drosophilidae) and two workers from each of five colonies of the ant *Pogonomyrmex barbatus* (Formicidae). The protocols were as described above, except that PCR products were digested with *SspI* prior to electrophoresis, and the ant DNA was amplified with primers SR-J-14612 and C1-N-2191 (Simon et al., 1994). Also, PCR reactions for the amplification of the 4 kb region of ant DNA contained 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 3.25 mM MgCl, 300 μM of each dNTP, 200 nM of each primer, 0.75 units of *Taq* DNA polymerase (Gibco BRL, Life Technologies Inc.) and 0.1–1 μg of template DNA in a total volume of 25 μl. Amplification conditions were 1 min at 94°C, 1 min at 45°C, and 4.5 min at 68°C, with 5 additional minutes at 68°C after the thirty cycles. Amplification cycles.

### References

- **Chapman and Hall, New York.**
- **Cold Spring Harbor Laboratory Press, New York.**


