SHORT COMMUNICATION

Exon-primed intron-crossing (EPIC) markers as a tool for ant phylogeography

Patrícia R. Ströher¹, Chenhong Li² & Marcio R. Pie¹

¹ Laboratório de Dinâmica Evolutiva e Sistemas Complexos, Departamento de Zoologia, Universidade Federal do Paraná, 81531–990 Curitiba-PR, Brazil. patricia.strroher@gmail.com; pie@ufpr.br
² College of Fisheries and Life Science, Shanghai Ocean University, 999 Hucheng Huan Road, Shanghai 201306, China.

ABSTRACT. Exon-primed intron-crossing (EPIC) markers as a tool for ant phylogeography. Due to their local abundance, diversity of adaptations and worldwide distribution, ants are a classic example of adaptive radiation. Despite this evolutionary and ecological importance, phylogeographical studies on ants have relied largely on mitochondrial markers. In this study we design and test exon-primed intron-crossing (EPIC) markers, which can be widely used to uncover ant intraspecific variation. Candidate markers were obtained through screening the available ant genomes for unlinked conserved exonic regions interspersed with introns. A subset of 15 markers was tested in vitro and showed successful amplification in several phylogenetically distant ant species. These markers represent an important step forward in ant phylogeography and population genetics, allowing for more extensive characterization of variation in ant nuclear DNA without the need to develop species-specific markers.

KEYWORDS. Formicidae; molecular markers; phylogeny; population genetics.

Ants are a dominant component of terrestrial animal communities worldwide (Hölldobler & Wilson 1990). For instance, ants and termites account for nearly one third of the entire animal biomass of the Brazilian Amazon forest (Fittkau & Klinge 1973). Moreover, ant density in some temperate localities can exceed 140 workers per m² (Kajak et al. 1971). The diversity of life histories in ants is also unparalleled, including solitary foragers, seed harvesters, plant obligate mutualists, fungus-growers, and group predators (Hölldobler & Wilson 1990). As a consequence, measuring levels of ant population differentiation in a given biome should therefore provide a representative picture of what the remaining animal communities in a given region might have experienced during their evolutionary past. Finally, the availability of efficient sampling methods that could uncover dozens of species from a given location in a time-scale of hours (e. g. Agosti & Alonso 2000) suggests that ants are among the best available model systems to investigate patterns of animal diversity and distribution of species in terrestrial ecosystems.

Until recently, most of the studies in ant phylogeography were based exclusively on mitochondrial markers (e. g. Goropashnaya et al. 2007; Solomon et al. 2008; Resende et al. 2010; Leppänen et al. 2011; but see Pringle et al. 2012). Such reliance on linked and uniparentally inherited loci might be problematic, given that several mechanisms such as introgression, selective sweeps, and the stochastic nature of the coalescent process might cause a mitochondrial tree not to be representative of the underlying level of population differentiation of a given species (Avise 2000; Wiens et al. 2010). More robust inferences on population structure should therefore include additional unlinked loci, leading to the prevalent use of microsatellites (Schlöther et al. 2004). However, given that microsatellite development is a laborious process that can only be applied to at most a few closely related species, their applicability to ant phylogeography and population genetics has been limited (Goropashnaya et al. 2007; Ross et al. 2007). A promising alternative method to uncover nuclear genetic variation is the use of exon-primed intron-crossing (EPIC) markers (Lessa 1992; Slade et al. 1993; Palumbi 1996), which has been successfully used in a variety of vertebrate taxa (e. g. Li et al. 2010; Silva-Oliveira et al. 2012). In this approach, primers bind to conserved exonic regions flanking an intron, which tends to accumulate mutations at a higher rate than transcribed regions (Graur & Li 2000). The availability of nuclear genomes for phylogenetically distinct ant species, Harpegnathos saltator Jerdon, 1851 Camponotus floridanus (Buckley, 1866), Linepithema humile (Mayr, 1868), Pogonomyrmex barbatus (Smith, 1858), Atta cephalotes (Linnaeus, 1758) and Solenopsis invicta Buren, 1972 (Bonasio et al. 2010; Smith et al. 2011a; Smith et al. 2011b; Suen et al. 2011; Würm et al. 2011, respectively), provides an ideal situation in which candidate EPIC markers can be selected that could potentially allow for successful amplification in a variety of ant lineages. In this study we use a recently developed pipeline (Li et al. 2010) to screen for candidate EPIC markers based on the genomes of six ant...
species. Primers were synthesized for a subset of 15 EPIC markers, which were then tested in vitro. These markers are ready to be widely used, providing an important tool to increase the quality and accuracy in studies on ant phylogeography, as well as the inference of phylogenetic relationships among closely related ant species.

We screened the complete nuclear genome of six ant species representing a phylogenetically diverse sample of the family: *H. saltator* (Ponerinae), *L. humile* (Dolichoderinae), *C. floridanus* (Formicinae), *S. invicta* (Myrmicinae), *P. barbatus* (Myrmicinae), and *A. cephalotes* (Myrmicinae) to identify conserved exons of single-copy genes. First, we retrieved coding sequences from *H. saltator* using its genome annotation information. We then compared the selected coding sequences of *H. saltator* with the genome of other ant species to identify single-copy and conserved exon among all six species. We screened for EPIC makers that are flanked by two single-copy and conserved exons. After this initial screening, potential markers were further selected based on the level of conservation of the flanking exons (to maximize the chance of cross-species successful amplification) and the size of the intron (450–650 pb, to allow for the amplification of the entire intron in a single reaction), from which 15 candidate markers were chosen for primer design and used for further analyses (Table I).

All primers were tested in vitro with the following ant species: *Gnamptogenys striatula* Mayr, 1884 (Ectatomminae), *Hylomyrma reitteri* (Myrmicinae), *Brachymyrmex sp.* (Formicinae), *Pheidole incisa* Mayr, 1870 (Myrmicinae) and *Linepithema sp.* (Dolichoderinae). We chose those species because they form a phylogenetically diverse sample and they belong to different species from those whose genomes were used in the initial screening. The phylogenetic relationships among all species involved in the present study are shown in Fig. 1. Total genomic DNA was extracted from entire ants using the PureLink™ Genomic DNA kit (Invitrogen™, USA), and 1 µL each (3.2 pmol) primer. The ultra-pure water and template were obtained in both directions. The sequencing reaction protocol was performed in 10 µL: 0.5 µL ABI Prism® BigDye™ v3.1 (Applied Biosystems Inc., Foster City, CA), 1.0 µL 5x buffer and 1 µL each (3.2 pmol) primer. The ultra-pure water and template to give 40–50 ng of DNA in each reaction was composed in the remainder of the mixture. The cycle sequencing reaction protocol contained an initial denaturation step of 96°C for 1 min, followed by 35 cycles of 10 s at 96°C denaturation, 60°C for 1 min and 70°C for 2 min, followed by final extension at 72°C for 6 min. Reactions were done in 25-µL reactions with 2 units of AmpliTaq DNA polymerase, 1x PCR buffer, 1.5mM of MgCl₂, 0.5 mM of dNTPs and 0.5 µM of each primer. PCR products were electrophoresed on 1.5% agarose gels with E-Gel® 1 Kb Plus DNA Ladder (to evaluate the size of the fragments), stained with ethidium bromide and visualized under UV light. Samples of the PCR products were purified using PEG 8000. All sequenced samples were obtained in both directions. The sequencing reaction protocol was performed in 10 µL: 0.5 µL ABI Prism® BigDye™ v3.1 (Applied Biosystems Inc., Foster City, CA), 1.0 µL 5x buffer and 1 µL each (3.2 pmol) primer. The ultra-pure water and template to give 40–50 ng of DNA in each reaction was composed in the remainder of the mixture. The cycle sequencing reaction protocol contained an initial denaturation step of 96°C for 1 min, followed by 35 cycles of 10 s at 96°C denaturation,

![Fig. 1. Phylogenetic relationships among all species involved in the present study, based on Brady et al. (2006) and unpublished results on *Hylomyrma reitteri*.](image-url)

Table I. EPIC primers developed in the present study and the corresponding position in the genome of *Harpegnathos saltator.*

<table>
<thead>
<tr>
<th>Locus</th>
<th>The position of exon 1 in <em>Harpegnathos saltator</em> genome</th>
<th>The position of exon 2 in <em>Harpegnathos saltator</em> genome</th>
<th>Average size of the intron across six species used for designing primers</th>
<th>Identity of the exon sequence across six species used for designing primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>ant1076</td>
<td>scaffold1076:27056:270920</td>
<td>scaffold1076:269369:269215</td>
<td>616.0</td>
<td>94.0</td>
</tr>
<tr>
<td>ant1087</td>
<td>scaffold1087:158069:157789</td>
<td>scaffold1087:157182:157026</td>
<td>938.0</td>
<td>96.0</td>
</tr>
<tr>
<td>ant1225</td>
<td>scaffold1225:12532:12329</td>
<td>scaffold1225:11932:11705</td>
<td>432.0</td>
<td>94.5</td>
</tr>
<tr>
<td>ant1281</td>
<td>scaffold1281:238506:238274</td>
<td>scaffold1281:237845:237709</td>
<td>406.0</td>
<td>94.5</td>
</tr>
<tr>
<td>ant1401</td>
<td>scaffold1401:28851:28726</td>
<td>scaffold1401:27273:27163</td>
<td>678.0</td>
<td>94.5</td>
</tr>
<tr>
<td>ant1503</td>
<td>scaffold1503:164317:164731</td>
<td>scaffold1503:166007:166262</td>
<td>932.0</td>
<td>94.0</td>
</tr>
<tr>
<td>ant346</td>
<td>scaffold346:68688:68870</td>
<td>scaffold346:69142:69264</td>
<td>291.0</td>
<td>96.5</td>
</tr>
<tr>
<td>ant384</td>
<td>scaffold384:472316:473086</td>
<td>scaffold384:472110:471826</td>
<td>833.0</td>
<td>96.5</td>
</tr>
<tr>
<td>ant389</td>
<td>scaffold389:104575:104358</td>
<td>scaffold389:103782:103644</td>
<td>94.5</td>
<td>508.0</td>
</tr>
<tr>
<td>ant505</td>
<td>scaffold505:742686:742795</td>
<td>scaffold505:743138:743431</td>
<td>95.5</td>
<td>408.0</td>
</tr>
<tr>
<td>ant839</td>
<td>scaffold839:296212:296350</td>
<td>scaffold839:297615:297720</td>
<td>97.0</td>
<td>1042.0</td>
</tr>
<tr>
<td>ant965</td>
<td>scaffold965:696156:696310</td>
<td>scaffold965:696571:696741</td>
<td>94.0</td>
<td>259.0</td>
</tr>
</tbody>
</table>
15 s at annealing 50°C and 4 min at 60°C. The final DNA precipitation was performed with isopropanol and run on an ABI 3500 sequencer to assess the level of intraspecific sequence variability. Chromatograms were edited in the Staden Package (Staden et al. 2000) and unambiguously aligned by hand using BioEdit program version 7.1.3 (Hall 1999), all sequences being available in GenBank (accession numbers KF738831 to KF738842).

The initial screening of potential fragments smaller than 3 kbp indicated 16,089 candidate loci (mean = 232.8 bp, SD = 381.4). Although not explored in the present study, shorter fragments (<200 bp) could be particularly useful in the context of next-generation sequencing (Puritz et al. 2012). We were able to successfully amplify all loci selected for the in vitro tests (Table II). Amplified product size varied between loci and species, from 400 to 1,000 bp. The size discrepancy for the same amplification product across species ranged from 50 to 500 bp. (mean = 196.6, Table II).

In order to test if the developed EPIC markers would provide sufficient intraspecific variability, we sequenced a subset of four samples from the amplified products of G. striatula obtained at locations between 600 km and 2,900 km apart (Brazilian states of Santa Catarina, São Paulo, Sergipe and Pernambuco). Different loci provided varying levels of divergence, from 0.22–0.56% (ant.1F/R, ant.202F/R) to 1.1–1.4% (ant.263F/R, ant.965F/R, and ant.1281F/R), with a maximum of 8.9% (ant.346F/R). None of the amplified sequences showed heterozygous indel positions. However, that is indeed a possibility with EPIC markers. Given the broad variety of product sizes among species, preliminary tests on a variety of loci might be necessary to select those with sufficiently small lengths such that these issues are minimized. Given that we specifically chose phylogenetically distant species, it is not surprising that sequences from different species cannot be aligned. However, the levels of intraspecific variability detected in the present study suggest that at least some loci might be valuable to assess the phylogenetic relationships of closely-related species.

The EPIC markers developed in the present study represent an important addition to the ant phylogeography toolbox, allowing for unprecedented new sources of information from nuclear markers. However, given that ants are a remarkably diverse and ancient lineage, it is possible that some markers might not be equally useful in all cases. Future studies should therefore include an initial assessment of fragment size and sequence variability for the species in question before proceeding to more extensive sampling. However, the fact that we obtained successful amplification of all loci for a phylogenetically diverse sample of ant lineages strongly suggests that these markers should be applicable to ants in general.

Table II. EPIC primers developed in the present study and the corresponding estimated fragment sizes based on gel electrophoresis.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Length</th>
<th>Sequence (5’ to 3’)</th>
<th>Brachymyrmex sp.</th>
<th>Pheidole incisa</th>
<th>Hylomyrma reitteri</th>
<th>Gnamptogenys striatula</th>
</tr>
</thead>
<tbody>
<tr>
<td>ant.1076F</td>
<td>24</td>
<td>AAAATCTNAGTGGAAACGCGTCA</td>
<td>600</td>
<td>600</td>
<td>550</td>
<td>550</td>
</tr>
<tr>
<td>ant.1076R</td>
<td>23</td>
<td>CCGATGCTTATYTCACAGTACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.1087F</td>
<td>21</td>
<td>ACCAGGGAGGGCTCGGAAGCTGGA</td>
<td>700</td>
<td>650</td>
<td>800</td>
<td>700</td>
</tr>
<tr>
<td>ant.1087R</td>
<td>27</td>
<td>GCAATGGTGAAGCTGACAGACCTTCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.1225F</td>
<td>26</td>
<td>TAAACCTATTGGAAGGACGCCAGGAG</td>
<td>500</td>
<td>600</td>
<td>650</td>
<td>1000</td>
</tr>
<tr>
<td>ant.1225R</td>
<td>27</td>
<td>GACTGATGCTTAAGCTGAGACGCTGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.1281F</td>
<td>23</td>
<td>GACGCAGGTTGYAACGAAAATACAC</td>
<td>500</td>
<td>500</td>
<td>400</td>
<td>550</td>
</tr>
<tr>
<td>ant.1281R</td>
<td>24</td>
<td>GCCRCAATATCCAGCTTCACAGCAG</td>
<td>700</td>
<td>550</td>
<td>800</td>
<td>700</td>
</tr>
<tr>
<td>ant.1401F</td>
<td>22</td>
<td>GYAGGAGGGAGCGCTCTCTAAATCT</td>
<td>800</td>
<td>700</td>
<td>800</td>
<td>800</td>
</tr>
<tr>
<td>ant.1401R</td>
<td>26</td>
<td>AAGCTTACTCTGAAAGACTCCACATC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.1503F</td>
<td>21</td>
<td>GRTTGCCCTCCAGGAGGATACA</td>
<td>800</td>
<td>700</td>
<td>800</td>
<td>800</td>
</tr>
<tr>
<td>ant.1503R</td>
<td>23</td>
<td>AAGGTAGTCAGGCAAGACACAC</td>
<td>800</td>
<td>700</td>
<td>800</td>
<td>800</td>
</tr>
<tr>
<td>ant.1F</td>
<td>27</td>
<td>CTTGCGTGCTTAYGAAGAACAGTYGTTAC</td>
<td>600</td>
<td>450</td>
<td>400</td>
<td>650</td>
</tr>
<tr>
<td>ant.1R</td>
<td>21</td>
<td>AAGCAGCTGCSAGCCGGTTCCAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.202F</td>
<td>26</td>
<td>CCYACTAATCTGTTATATCAGAAGCAG</td>
<td>650</td>
<td>500</td>
<td>500</td>
<td>600</td>
</tr>
<tr>
<td>ant.202R</td>
<td>22</td>
<td>GACACAAGTGGAGGACGCCCTTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.263F</td>
<td>27</td>
<td>GACTGTCGAGAATACACTCTCTCCAC</td>
<td>450</td>
<td>550</td>
<td>550</td>
<td>600</td>
</tr>
<tr>
<td>ant.263R</td>
<td>24</td>
<td>GTTTGGTGTGGAGGCAATTTGGAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.346F</td>
<td>26</td>
<td>GTTGGACACATCGTGGGATCT</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>500</td>
</tr>
<tr>
<td>ant.346R</td>
<td>26</td>
<td>GAATTCTTTGTTGGAATCTCGGGCTCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.839F</td>
<td>25</td>
<td>CAATGGCGATTTACAACGAATTTCT</td>
<td>850</td>
<td>850</td>
<td>850</td>
<td>850</td>
</tr>
<tr>
<td>ant.839R</td>
<td>22</td>
<td>TAYCCGGAGACACCTCACTTATCAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ant.965F</td>
<td>24</td>
<td>AGTTACAGGTTCACAGCGTGCATTAC</td>
<td>450</td>
<td>450</td>
<td>850</td>
<td>400</td>
</tr>
<tr>
<td>ant.965R</td>
<td>25</td>
<td>GAGAAGGYYGAYTTAAGACTGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Revista Brasileira de Entomologia 57(4): 427–430, December 2013
ACKNOWLEDGMENTS

We thank Rogério R. Silva for providing specimens used in the in vitro tests. We also thank Paula F. B. Bassi for assistance during laboratory work. This study was partly funded by a grant from CNPq/MCT (571334/2008-3) to MRP.

REFERENCES


