New Source of Genetic Polymorphisms in Lepidoptera?

Anna K. Hundsdoerfer a,b,* and Michael Wink a,*

a Institut für Pharmazie und Molekulare Biotechnologie, Abteilung Biologie, Im Neuenheimer Feld 364, D-69120 Heidelberg, Germany. Fax: +49(0)62 21 54 48 84.
E-mail: wink@uni-hd.de.

b Present address: Museum für Tierkunde, Königsbrücker Landstr. 159, D-01109 Dresden, Germany. Fax: +49(0)35 18926327. E-mail: anna.hundsdoerfer@snsd.smwk.sachsen.de

* Authors for correspondence and reprint requests

Z. Naturforsch. 60c, 618–624 (2005); received March 17, 2005

The variability level of the ISSR (inter-simple sequences repeat) primer (GACA) 4 was examined in the three Lepidoptera families Pyralidae, Sphingidae and Pieridae. Our study shows that the tetra-repeat (GACA)n is evidently present in sufficient numbers in these butterflies to provide informative DNA fingerprints. The variability is mostly rather high, but within a comparable range to other ISSR studies. Although less polymorphisms may be encountered in some butterfly families, this study indicates that high variability of this marker may be a common characteristic of Lepidoptera genomes. An appeal for a minimal level of standardization of ISSR-PCR data analysis is formulated to enable an exact comparison between the groups of organisms studied with this fingerprint technique.

Key words: GACA-ISSR-PCR, Standardization, Lepidoptera

Introduction

Inter-simple sequence repeat (ISSR) amplification is a relatively new technique that can rapidly differentiate closely related individuals (Gupta et al., 1994; Tsumura et al., 1996; Zietkiewicz et al., 1994). Unlike nucleotide sequences, ISSR markers describe DNA characteristics at several, mostly nuclear, chromosomal loci and thus avoid the use of gene trees as surrogates of species trees (Martin and Salamini, 2000). Revealing discontinuous markers, ISSR-PCR can provide a measure of genetic differences dispersed across the entire genome. The absence of a band is interpreted as primer divergence or the loss of a locus through either the deletion of the SSR site or a chromosomal rearrangement (Wolfe and Liston, 1998).

Until recently, the use of ISSR markers was restricted to cultivated plant species (e.g. Tsumura et al., 1996; Fang and Roose, 1997; Assefa et al., 2003), their pests (e.g. Kumar et al., 2001) or other animals of economic importance (e.g. Reddy et al., 1999; Nagaraju and Goldsmith, 2002). However, it is now increasingly being applied in population-level to interspecific studies of natural populations of plants and animals (e.g. Culley and Wolfe, 2001; Kauserud and Schumacher, 2003; King and Ferris, 2000; Sudupak, 2004; Guicking et al., 2002a, b; Luque et al., 2002; Nagy et al., 2002; Treutlein et al., 2003; Hundsdoerfer et al., 2005). ISSR markers have proven an efficient method for detecting hybridisation in natural populations of plants (Wolfe et al., 1998) and animals (Wink et al., 2001). The use of the primer (GACA)n has been widened to include an application for determining the gender of birds (Wink et al., 1998).

ISSR marker analysis involves PCR amplification of DNA using a single primer composed of a microsatellite sequence such as (GACA)n. The primer can be anchored at either the 3’ or 5’ end by 2–4 arbitrary, often degenerate, nucleotides. The amplification products are then separated by PAGE (polyacrylamide gel electrophoresis) or on an agarose gel. ISSR-PCR can reveal a large number of fragments, and thus many potentially polymorphic loci, in one PCR with good reproducibility. These characteristics are great advantages compared to RAPD amplification. In addition, ISSR-PCR markers have been reported to show the greatest variability in Bombyx mori (Saturniidae, Lepidoptera) strains (Reddy et al., 1999) when compared to RAPD-PCR and microsatellites. But the major advantage of ISSR-PCR over microsatellite analyses is the cost-efficiency, because no initial investment in primer design is necessary. However, a disadvantage is that loci are...
usually interpreted as dominant markers, so no genotypic allele information is acquired, as in microsatellite analyses.

Microsatellites had been assumed to be rather rare in Lepidoptera because only few could be identified and these were difficult to find (Meglecz and Solignac, 1998; Traut et al., 1992; Palo et al., 1995). However, an increasing number of studies contradict this hypothesis by reporting either successful use of microsatellite or ISSR-PCR polymorphisms (Nagaraju and Goldsmith, 2002; Bogdanowicz et al., 1997; Keyghobadi et al., 1999; Reddy et al., 1999; Harper et al., 2000; Kumar et al., 2001; Luque et al., 2002; Hundsdoerfer et al., 2005). In these publications the simple repeat motifs described and used most often are (CA)n, (AC)n, (GA)n, (AG)n, (TG)n, (GT)n, (TC)n, (CT)n, and more rarely (TGA)n and (GATA)n. More compound motifs have also been reported in butterflies, however, only few show polymorphisms. Luque et al. (2002) have reported no polymorphisms for ISSR-PCR primers (CTGT)4, (TCC)5, (GATA)4, (GTA)4, and only few for (GGAT)4 and (GACA)4 in noctuid moths. In contrast, the ISSR-PCR primer (GACA)4 has been shown to exhibit a surprisingly high level of variation in the Hyles euphorbiae-complex (Sphingidae; Hundsdoerfer et al. 2005). In this study we examined the utility of the ISSR-PCR primer (GACA)4 for the examination of intraspecific polymorphisms of Lepidoptera in more detail and were able to show great fingerprint variability in the three butterfly families Pyralidae, Sphingidae (both “Heterocera”) and Pieridae (Rhopalocera).

Materials and Methods

Specimens

Three species from different families of Lepidoptera were sampled with six individuals each. All specimens were collected in Germany. Caterpillars of Acentria ephemerella Denis & Schiffermüller 1775 (Pyralidae) were collected from different patches of its foodplant (the pondweed Potamogeton perfoliatus Linnaeus 1753; Potamogetonaceae), in Lake Constance (Konstanz) in 2002. Hyles euphorbiae Linnaeus 1758 (Sphingidae) caterpillars were collected from different foodplants of Euphorbia cyparissias Linnaeus 1753 (Euphorbiaceae) near Viernheim (region of Mannheim) in 2001. Pieris rapae Linnaeus 1758 (Pieridae) adults were caught near Schwetzingen (region of Mannheim) in 2004.

DNA extraction

For total DNA extraction, legs or a piece of the anterior abdomen were placed into Buffer B [25 mM EDTA, 75 mM NaCl, 10 mM Tris [tris(hydroxymethyl)aminomethane], pH 7.5, 1% SDS] in an Eppendorf-cup. The material was then either cut into small pieces with scissors or crushed with a small pestle and incubated with Proteinase-K (1 mg/sample) at 50 °C for at least 10 h. The remaining proteins were precipitated with at least ½ volume of saturated NaCl-solution and removed by centrifugation. The supernatant was cleaned with a phenol/chloroform reagent. After removal of the remaining phenol with chloroform, 0.8 volume ice-cold isopropanol was added and the DNA was pelleted by centrifugation at 4 °C. Prior to dissolving the DNA-pellet in 20–800 µl TE-Buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5), it was washed twice with ice-cold 70% ethanol. If the DNA solution still contained PCR-inhibitors after this procedure it was cleaned over NucleoSpin® Tissue columns (Macherey-Nagel), adapting the first step by mixing 30 µl DNA solution with 180 µl of the Lysis buffer T1 and then following the standard protocol.

ISSR-PCR analysis

The fragments between the microsatellites consisting of the tetra-repeat (GACA)a were amplified using the non-anchored primer (GACA)4 (i.e. GACAGACAGACAGACA; Epplen et al., 1992). Each PCR was performed with about 500 ng template DNA in a 25 µl volume [10 pmol of the primer and 0.625 nmol of each dNTP, except dATP: 0.28 nmol cold dATP plus 0.1 µl radioactive α-33P-dATP solution (370 MBq/ml, Amersham Biosciences), 0.1 units of Taq-Polymerase (SIGMA) and water, buffered with 10 mM Tris-HCl, 50 mM KCl, 0.5% Triton X-100, 1.5 mM MgCl2] and covered by two drops of mineral oil. Thermo cycling was performed with a Trio Thermo block TB1 (Biometra, Göttingen). Following the initial 5 min denaturation at 94 °C, the program consisted of 28 cycles of 50 s at 94 °C, 30 s at 55 °C, 120 s at 72 °C and 25 min at 72 °C for final elongation. The DNA fragments were separated by PAGE (polyacrylamide gel electrophoresis) in a vertical apparatus (Base Acer Sequencer, Stratagene) for 4 h at 65 W.
The denaturating gels [6 m urea, 100 ml Long Ranger Solution, Biozym (PA), 100 ml TBE-Buffer (10x: 1 m Tris, 0.83 m boric acid, 10 mm EDTA, pH 8.6)] had a size of 45 × 30 cm and a thickness of 0.25 mm. After drying, the gel was exposed to an X-ray film (Hyperfilm-MP, Amersham) for at least 12 h and developed (Kodak). The film was then scanned with a resolution of 300–600 dpi. The bands were analysed visually on the film itself, but marked on an A3-sized print of the film.

The bands of the ISSR-PCR fingerprints were interpreted as representing independent characters (Assefa et al., 2003) and were visually scored into a data matrix as either absent (“0”) or present (“1”). Qualitative differences in band intensity were not considered (Assefa et al., 2003). We assumed that markers from different loci did not co-migrate to the same position on the gel (Culley and Wolfe, 2001). Generally, ISSR bands are scored as dominant markers (e.g. Zietkiewicz et al., 1994), so this procedure was adopted in this analysis as well. This implies that slight differences in the retention index of apparently homologous bands or band-patterns were not scored as alternate allelic conditions or as separate characters.

The bands with high molecular weight (not illustrated) were excluded due to non-

Results

The gels of the ISSR-PCR fingerprints are illustrated in Fig. 1. Characters were coded species-specifically, since the three species did not share common bands. The bands with high molecular weight (not illustrated) were excluded due to non-

Fig. 1. Illustration of the gels of the amplified ISSR fragments (six samples each) of the three species of the families Pyralidae, Sphingidae and Pieridae (all Lepidoptera) as visualized on a PA-gel by the use of radioactively labelled adenine; a) Acentria ephemrella (Pyralidae); b) Hyles euphorbiae (Sphingidae); c) Pieris rapae (Pieridae).
Table I. Comparison of the variability of the ISSR data of the three species examined.

<table>
<thead>
<tr>
<th></th>
<th>Acentria ephemerella</th>
<th>Hyles euphorbiae</th>
<th>Pieris rapae</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total no. of bands</td>
<td>24</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>Fixed, constant</td>
<td>6 (25.0%)</td>
<td>3 (20.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>variable bands</td>
<td>18 (75.0%)</td>
<td>12 (80.0%)</td>
<td>11 (100.0%)</td>
</tr>
<tr>
<td>Shared, parsimony</td>
<td>11 (45.8%)</td>
<td>8 (53.3%)</td>
<td>11 (72.7%)</td>
</tr>
<tr>
<td>informative bands</td>
<td>7 (29.2%)</td>
<td>4 (26.7%)</td>
<td>3 (27.3%)</td>
</tr>
</tbody>
</table>

reproducibility. The data set did not contain missing data.

In total, the ISSR-PCR fingerprints of the six samples of *Acentria*, *Hyles* and *Pieris* each led to 24, 15 and 11 bands, respectively, that were scored for data analyses. The variability of the ISSR-PCR fingerprint data is presented in Table I. The highest number of both variable and shared (parsimony informative) sites of the ISSR-PCR fingerprints are detected in *Pieris*, the smallest number in *Acentria*.

The average population distances based on Dice (1945; also Nei and Li, 1979) and Jaccard (1901, 1908) are presented in Table II. The standardized scale of these distance measures define a value of 1 as representing maximal genetic distance, and reflecting the situation of the two samples analysed sharing no bands, likewise, a value of 0 represents no genetic distance and reflects that all analysed bands are shared. Therefore, a higher value reflects lower fingerprint similarity. The variability of *Acentria* and *Hyles* fingerprints is similar. The variability of those of *Pieris* is again the highest of the three species.

**Table II.** Dice (1945; also Nei and Li, 1979) and Jaccard (1901, 1908) distances of the ISSR-PCR data within the three species studied. Presented are the minimal (min), maximal (max) and mean distance values, with the standard deviation (stdev).

<table>
<thead>
<tr>
<th></th>
<th>Dice distance</th>
<th>Jaccard distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>min</td>
<td>max</td>
</tr>
<tr>
<td><em>Acentria ephemerella</em></td>
<td>0.231</td>
<td>0.481</td>
</tr>
<tr>
<td><em>Hyles euphorbiae</em></td>
<td>0.091</td>
<td>0.600</td>
</tr>
<tr>
<td><em>Pieris rapae</em></td>
<td>0.333</td>
<td>1.000</td>
</tr>
</tbody>
</table>

**Discussion**

Of the three species studied, *Acentria ephemerella* (Pyralidae), *Hyles euphorbiae* (Sphingidae) and *Pieris rapae* (Pieridae), the first species showed the least variability. This may simply be a characteristic of the occurrence of the \((GACA)_n\) motif in this species genome. However, an interesting characteristic of this species’ semi-aquatic lifestyle is that most of its females are immobile. This leads to reduced genetic exchange within this species compared to those with mobile females. The spurge hawkmoths (*Hyles euphorbiae*), on the other hand, are strong-flying (males and females), enabling genetic exchange over large distances. This behaviour may contribute to the higher genomic variability observed in the occurrence of the \((GACA)_n\) motif. The Cabbage White (or Small White, *Pieris rapae*) is of smaller body size and does not fly especially well. However, the species feeds on different plant species of the family Brassicaceae, which include extensively cultivated crops. When harvested, these crops are transported over large distances and eggs or caterpillars of the Cabbage White may be transported with them passively. The differences of variability observed between the species examined may thus either reflect various distinct genomic occurrences of the \((GACA)_n\) motif or may be caused by different levels of genetic exchange due to the species’ lifestyles. On the basis of the present data it is not possible to assess which one of these two possibilities holds, since we do not yet fully understand the genetics of the ISSR markers.

The motif \((GACA)_n\) appears to occur numerous Lepidoptera genomes. This also holds for other organisms, especially birds (e.g. Wink et al., 2002 and unpublished data). Since \((GACA)_n\) is not one of the classical microsatellite motifs, this fact is somewhat surprising. ISSR-PCR studies with this primer combined with a microsatellite analy-
sis based on primers flanking the tetra-repeat \((\text{GACA})_n\) promise interesting insight into the genetics of the ISSR markers.

The comparison between the genetic ISSR distances measured in this study and those of other organisms revealed some difficulties: Representing a relatively new technique, the ISSR-PCR fingerprint data analysis is not standardized yet. Many computer programs (freeware, shareware and costly ones) are used for the calculation of several distance measures, based on different coefficients. Representing a widely used coefficient and excluding shared absences (which we regard as an important aspect), we decided to calculate Jaccard (1901, 1908) and Dice (1945; also Nei and Li, 1979) distances in this study, enabling a comparison to several other publications. Culley and Wolfe (2001) studied the population genetic structure of \textit{Viola pubescens} Aiton (Violaceae) and also found a surprisingly high ISSR variability compared to other species of the genus. This was reflected by 100% polymorphic bands, which is also the case in \textit{Pieris rapae} of our study (Table I). Although a different software was used for the calculations of the Nei and Li (1979) distances [Culley and Wolfe (2001) used \textsc{javasiml}, while we used \textsc{RAPDistance}], the mathematical algorithm is expected to be the same, resulting in comparable results. With a maximum of 1.000 the Nei and Li (1979) genetic distances within the population of \textit{Pieris rapae} are even higher than those of \textit{Viola pubescens} which exhibit a maximum of 0.672.

Using yet another software for the calculation of Nei and Li (1979) distances (MVSP-pc), the values in a very recent publication on ISSR data (Sudupak 2004; on natural populations of chickpea species: \textit{Cicer}, Fabaceae) unfortunately do not appear to be comparable. Sudupak (2004) used the following mathematical formula to calculate the distances: the genetic distance between samples \(a\) and \(b\) \((GD_{ab})\) is defined as \(GD_{ab} = -\ln (S_{ab})\), where \(S_{ab}\) is the Nei and Li (1979) “similarity” or “distance coefficient”. These values were calculated with the software MVSP-pc and transformed to the “distance estimates” by the above mentioned formula. In his nomenclature, our distance estimates (calculated by \textsc{RAPDistance}) are only the “distance coefficients” and a transformation with the above formula would be necessary to enable comparison. Since a presentation of all possible procedures of ISSR data analyses would extend the scope of this paper, we refrain from comparing the ISSR distances quantitatively to further publications. However, a qualitative comparison can be drawn to Sudupak’s (2004) study: some intraspecific distance averages are reported to be 0 in Sudupak (2004), indicating very low intraspecific (population-level) variability (this is also formulated in his text). The intrapopulation distances were rather high in our dataset: no pairwise comparisons resulted in a genetic distance of 0 (indicating identical band patterns).

The Jaccard distances (1901, 1908) are often only used to construct distance matrices for cluster analyses rather than be reported as values. In Galván et al. (2003), however, the Jaccard distances of ISSR patterns within the common bean (\textit{Phaseolus vulgaris} L.) are presented in the text: they ranged from 0.23 to 1.00 (calculated with NTSYS-pc), with a mean of 0.80. These values appear comparable to ours. In the case of \textit{Pieris rapae} in the study at hand, the maximum and the mean values are the same as in the common bean (Galván et al., 2003), although the minimum value of \textit{Pieris} distances is higher. Thus, the relatively high level of variation we observed in our ISSR study of Lepidoptera proves to be within the same range as that of other groups of organisms.

**Conclusion**

ISSR-PCRs with the primer \((\text{GACA})_4\) appear to result in fingerprints with a high level of polymorphisms in Pyralidae, Sphingidae and Pieridae (Lepidoptera). The high variability found in the \textit{Hyles euphorbiae}-complex (Sphingidae; Hundsdoerfer et al., 2005) does not appear to be restricted to this group of moths. However, differences can be observed between taxonomic groups. This may simply reflect differing genomic occurrence of the \((\text{GACA})_n\) motif or may be caused by different levels of genetic exchange due to the species’ specific lifestyles. Further research into the genetics of ISSR markers in Lepidoptera will contribute to the understanding of pattern differences between species. This might elucidate whether the lifestyle of the species, along with the resulting differences in genetic exchange can influence differences of variability of ISSR-PCR patterns. In addition, we would like to formulate an appeal for a minimal level of standardization of ISSR-PCR data analysis to enable an accurate comparison between the groups of organisms studied. It could greatly enhance the understand-
The occurrence of the SSR repeat (CA)_n and its variations in Lepidoptera has been often reported (e.g., Palo et al., 1995; Bogdanowicz et al., 1997; Reddy et al., 1999; Kumar et al., 2001; Luque et al., 2002). In contrast, the SSR repeat (GACA)_n has only been mentioned once (Luque et al., 2002), and appeared not to show sufficient polymorphisms in Noctuidae. We report here various variability levels of this primer in three families of Lepidoptera. Our study shows that the tetra-repeat (GACA)_n is evidently present in sufficient numbers to provide informative fingerprints in the three species of Lepidoptera studied. Although less polymorphisms may be encountered in some butterfly families, our investigations do indicate that high variability of this marker may be a common characteristic of Lepidoptera genomes.

Acknowledgements

We thank Michael Korn and Dietmar Straile (Konstanz) for the Acentria samples and Daniela Guicking for the introduction to the ISSR-PCR lab methods. Special thanks also go to Harry Hasford and Margarita Tippmann for language corrections that greatly improved the manuscript.


