Application of plant DNA markers in forensic botany: Genetic comparison of Quercus evidence leaves to crime scene trees using microsatellites

Kathleen J. Craft\textsuperscript{a}, Jeffrey D. Owens\textsuperscript{b}, Mary V. Ashley\textsuperscript{a,\*}

\textsuperscript{a}Department of Biological Sciences, 845 W. Taylor St., M/C 066, University of Illinois at Chicago, Chicago, IL 60607, United States
\textsuperscript{b}Marion County Sheriff’s Office, Ocala, FL 34478, United States

Received 1 March 2005; received in revised form 3 March 2006; accepted 3 March 2006
Available online 2 May 2006

Abstract

As highly polymorphic DNA markers become increasingly available for a wide range of plant and animal species, there will be increasing opportunities for applications to forensic investigations. To date, however, relatively few studies have reported using DNA profiles of non-human species to place suspects at or near crime scenes. Here we describe an investigation of a double homicide of a female and her near-term fetus. Leaf material taken from a suspect’s vehicle was identified to be that of sand live oak, \textit{Quercus geminata}, the same tree species that occurred near a shallow grave where the victims were found. \textit{Quercus}-specific DNA microsatellites were used to genotype both dried and fresh material from trees located near the burial site and from the material taken from the suspect’s car. Samples from the local population of \textit{Q. geminata} were also collected and genotyped in order to demonstrate that genetic variation at four microsatellite loci was sufficient to assign leaves to an individual tree with high statistical certainty. The cumulative average probability of identity for these four loci was $2.06 \times 10^{-6}$. DNA was successfully obtained from the dried leaf material although PCR amplification was more difficult than amplification of DNA from fresh leaves. The DNA profiles of the dried leaves from the suspect’s car did not match those of the trees near the crime scene. Although this investigation did not provide evidence that could be used against the suspect, it does demonstrate the potential for plant microsatellite markers providing physical evidence that links plant materials to live plants at or near crime scenes.

* Corresponding author. Tel.: +1 312 413 9700; fax: +1 312 413 9462.
E-mail address: ashley@uic.edu (M.V. Ashley).

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Quercus; Oak; Microsatellites; Short tandem repeats (STR); Forensic botany

1. Introduction

While the field of forensic botany encompasses a variety of sub-disciplines ranging from plant anatomy to palynology (the study of pollen) [1], applications involving DNA profiling of plant material have to date been fairly limited. One application of DNA typing to forensics has involved drug enforcement, for example identifying \textit{Cannabis sativa} specimens [2–6]. Less common is the use of DNA profiling to produce physical evidence to link a suspect to a crime scene. The single well-known case involved seed pods of the Palo Verde tree (\textit{Cercidium} sp.) recovered from the pick-up truck of a suspect. Randomly Amplified Polymorphic DNA (RAPD) genetic profiling was used to link the seed pods to a single tree found at the murder site [7,8]. RAPD profiling was also applied to a lawsuit involving unauthorized commercialization of a patented variety of strawberry [9] and in a homicide case involving identifying the source of mosses found on a suspect [10].

Plant population biologists and ecologists are increasingly developing and employing microsatellite markers, also called short tandem repeats (STRs) or simple sequence repeats (SSRs), for reproductive and demographic studies of native plants. For example, microsatellites have been used to study pollination patterns [11–14], seed dispersal [15–17], gene flow [12,18–20] and hybridization [21,22]. While microsatellites have become the DNA marker of choice for forensic analysis of human DNA, to date they have received very limited application in forensic botany, restricted to applications in narcotics and drug enforcement [23–25]. However, the growing...
availability of microsatellite markers in a wide range of plant species provides potentially important tools for forensic applications involving a wider variety of criminal investigations. Because of the high levels of variability typical of microsatellite loci, composite genotypes across several loci can be used to assign trace plant material to living plants with high statistical probability.

Here we report the results of an investigation that utilized dried leaf material found in the trunk of a suspect’s car. The case was a double homicide, where the bodies of a pregnant woman and her near-term fetus were found buried in a shallow grave in a small woodlot in central Florida. Three large trees, identified as *Quercus geminata* (sand live oak), were found within a few meters of the gravesite, including one tree with a canopy that spread over the gravesite. Two leaves from the suspect’s car were identified as *Q. geminata* and a third as either *Q. geminata* or the closely related species *Q. virginiana* (Virginia live oak). Because the case against the suspect was comprised of strong circumstantial evidence but little physical evidence, and because the suspect denied ever being at the site, we decided to attempt to genetically compare the leaf material from the suspect’s car to the living trees at the burial site, in an effort to provide physical evidence to place the car at the crime scene.

Our study had four objectives. First, we would determine whether microsatellite loci developed from other species of oak would successfully produce genetic profiles in *Q. geminata*. Second, we would determine if the levels of variability for these markers were sufficiently high to link leaves to an individual tree with high statistical confidence. Third, we would determine whether we could extract DNA from older, dried leaf material and if so, if the DNA would yield acceptable genotypes. This step was necessary, as we would be working with dried leaves that had been stored in an evidence envelope for approximately 5 years. Our final objective was to extract DNA from the evidence leaves and compare those to the genotypes of the three trees located near the burial site. A genetic match would provide evidence to place the suspect’s car at the scene.

## 2. Materials and methods

A sample collection of mature leaves from 24 individual trees identified as *Q. geminata* was made within 2–10 km of the crime scene in Ocala, Florida. Leaves were collected at 13 sites, with 1–5 trees sampled per site. GPS coordinates were taken for each scene in Ocala, Florida. Leaves were collected at 13 sites, with 23 QpZAG9 and QpZAG110, were developed for *Q. petraea* [28] and two, MSQ13 and MSQ4, were developed for *Q. macrocarpa* [29].

Polymerase chain reaction (PCR) was performed using 0.2–0.4 µg genomic DNA, 100 µM dNTP (Perkin-Elmer), 0.15–0.6 µM primer, 2.0–3.0 mM/L MgCl₂, 1.0 µg/µL bovine serum albumin, PCR buffer (Promega), and 0.04 µL Taq polymerase (Promega). One primer of each pair was end-labeled with a fluorescent tag (Cy-3, Fam, or Tet) (IDT DNA Technologies). Polymerase chain reactions consisted of a 94 °C preheat followed by 38 cycles of 94 °C denaturation for 30 s, 54 °C annealing for 30 s, 72 °C extension for 30 s, and one 72 °C extension for 5 min. PCR products were then mixed with a fluorescent size standard, ROX 350 (Genescan) and run on an MJ BaseStation automated sequencer (MJ Research, Waltham, Massachusetts, USA). The PCR products (‘alleles’) consist of the microsatellite repeat, the

### Table 1

Description of polymorphism for markers used in investigation

<table>
<thead>
<tr>
<th>Microsatellite locus</th>
<th>Source species</th>
<th><em>Q. geminata</em> sample size</th>
<th>Number of alleles</th>
<th><em>Hₐ</em> observed heterozygosity</th>
<th><em>Hₑ</em> gene diversity</th>
<th>PIC</th>
<th>CE PIave</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSQ4</td>
<td><em>Q. macrocarpa</em> sample size</td>
<td>20</td>
<td>12</td>
<td>0.65</td>
<td>0.92</td>
<td>0.879</td>
<td>1.01 × 10⁻²</td>
</tr>
<tr>
<td>MSQ13</td>
<td><em>Q. macrocarpa</em></td>
<td>23</td>
<td>12</td>
<td>0.96</td>
<td>0.86</td>
<td>0.828</td>
<td>3.51 × 10⁻⁴</td>
</tr>
<tr>
<td>QpZAG9</td>
<td><em>Q. petraea</em></td>
<td>24</td>
<td>10</td>
<td>0.75</td>
<td>0.87</td>
<td>0.831</td>
<td>1.65 × 10⁻⁵</td>
</tr>
<tr>
<td>QpZAG110</td>
<td><em>Q. petraea</em></td>
<td>24</td>
<td>8</td>
<td>0.59</td>
<td>0.66</td>
<td>0.611</td>
<td>2.06 × 10⁻⁶</td>
</tr>
</tbody>
</table>
two primer-binding regions, and flanking regions on each side of the repeat. Differences in the PCR product size will reflect changes in the number of microsatellite repeat units. To assure gel-to-gel consistency, both a size standard and several individuals whose genotypes were already scored were run on each gel. All genotypes were scored using Cartographer software (MJ Research).

Analyses of genetic variation including number of alleles, observed heterozygosity, and gene diversity were calculated with the aid of the FSTAT program [30]. Polymorphic Information Content (PIC) was calculated using CERVUS software [31]. The average probability of identity (PIave) was calculated using the program API-CALC [32]. This is the probability of observing two copies of any profile in the population. Parameter values for API-CALC were set to assume that trees were unrelated. Q. geminata is diploid; therefore polyploidy was not a concern in this study.

We obtained leaf samples previously collected from the burial site by forensic botanist David Hall. These samples were taken in April of 2003 and included both dried and fresh leaves from the three Q. geminata trees (2352, 2353 and 2354) located at the burial site. We used these leaves to assess our ability to extract amplifiable DNA from older leaves. In this case, the dried, brown leaves were collected from the ground beneath the canopy of each tree, while the fresh, green leaves were taken directly from the trees. DNA was extracted and microsatellite loci were amplified using the methods described above.

On February 6, 2004, two evidence envelopes containing leaves collected from the trunk of the suspect’s car by Ocala detectives in 1999 were transported to our laboratory at the University of Illinois at Chicago. Approximately half of each of three dried oak leaves, two previously identified as Q. geminata and one as Q. virginiana by David Hall, was removed using a sterile razor blade. These were samples NFGQ A, NFGQ C and NFGQ D, taken from the trunk of the suspect’s vehicle (Table 2). The remaining leaf material was returned. These samples were extracted and genotyped as described above, with extra caution taken to avoid any possibility of cross-contamination of samples. Once we obtained genotypes, we compared these to the three trees at the burial site. Tree 2352 was of particular interest as it was closest to the burial site and had branches overhanging the grave.

3. Results

All four primer pairs developed for other species of oaks amplified alleles in the population sample of Q. geminata from the Ocala area with only minor modifications of PCR conditions. Table 1 shows descriptive statistics of variability. All four loci were highly variable, having 8–12 alleles per locus, gene diversity ranging from 0.66 to 0.92, and PIC ranging from 0.611 to 0.879. Our findings of high variation are similar to those previously reported for Q. geminata [26]. All 24 trees sampled and genotyped had unique genetic profiles with these four markers. The cumulative expected PIave using these four loci was $2.06 \times 10^{-6}$. Because of the relatively small sample size, these values should be considered to be estimates of the variability and exclusion probability. We did not undertake more extensive sampling because mismatches between the evidence leaves and the trees at the burial site were identified.

DNA extractions from both fresh and dried leaves were successful (Fig. 1). However, the DNA from the fresh leaf material amplified much more consistently than that obtained from dry leaf material, suggesting that DNA from dried leaf material was degraded or contained more PCR inhibitors than did DNA from fresh material. Nevertheless, when amplifications were successful, the genotypes from fresh and dry leaf material from the same tree consistently matched. That is, there was no indication that analysis of degraded material leads to inaccurate or altered genetic profiles. Fig. 2 shows examples of

<table>
<thead>
<tr>
<th>Microsatellite genotypes&lt;sup&gt;a&lt;/sup&gt;</th>
<th>QpZAG9</th>
<th>QpZAG110</th>
<th>MSQ13</th>
<th>MSQ4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trees at site</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2352</td>
<td>264, 264</td>
<td>214, 226</td>
<td>246, 250</td>
<td>205, 205</td>
</tr>
<tr>
<td>2353</td>
<td>264, 264</td>
<td>216, 216</td>
<td>244, 254</td>
<td>218, 232</td>
</tr>
<tr>
<td>2354</td>
<td>264, 268</td>
<td>216, 222</td>
<td>244, 258</td>
<td>203, 203</td>
</tr>
<tr>
<td><strong>Leaves, suspect’s car</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFGQ A</td>
<td>264, 268/270</td>
<td>208, 210</td>
<td>n.a.</td>
<td>203&lt;sup&gt;b&lt;/sup&gt;, 203/226</td>
</tr>
<tr>
<td>NFGQ C</td>
<td>264, 264</td>
<td>208/210, 210&lt;sup&gt;b&lt;/sup&gt;</td>
<td>n.a.</td>
<td>210&lt;sup&gt;b&lt;/sup&gt;, 210/214</td>
</tr>
<tr>
<td>NFGQ D</td>
<td>264, 264</td>
<td>208/214/210</td>
<td>246, 248</td>
<td>214, 226</td>
</tr>
</tbody>
</table>

<sup>a</sup> Numbers represent PCR product sizes in nucleotide base pairs. Each PCR product includes the microsatellite alleles, primer-binding regions, and flanking regions. Each individual has two alleles at each locus, separated by a comma. A forward slash indicates ambiguous scoring of allele size.

<sup>b</sup> Denotes ambiguous genotypes possibly due to allelic dropout.

Fig. 2. Comparison of PCR amplification of fresh (top panel) and dry (lower panel) leaves from tree 2353 at loci QpZAG9 and QpZAG110.
electropherograms obtained from DNA extracted from fresh and dry material collected from tree 2353, one of the three trees found at the burial site, at loci QpZAG110 and QpZAG9.

The genotypes of the three trees at the burial site and the three evidence leaves (NFGQ A, NFGQ C and NFGQ D) are given in Table 2. There is some missing (locus MSQ13 for NFGQ A and NFGQ B) and equivocal data for the evidence leaves. Despite several attempts at amplification, we could not obtain complete and unambiguous genotypes at all four loci. One likely problem with DNA extracted from the older leaves was allelic dropout. Allele dropout occurs when one allele in a heterozygote produces a much stronger signal than the other allele, resulting in false homozygote scoring [33,34]. Possible instances of allelic dropout are denoted in Table 2. The one other ambiguous genotype in Table 2 was due to small gel-to-gel shifts in scoring allele size. We found matching genotypes between evidence leaves and trees at the site for individual loci (e.g., 264 homozygote genotype was found for four of the six trees) but there were no matches between evidence leaves and trees at the site across all four loci. Fig. 3 shows a comparison of electropherograms for one evidence leaf and the three trees at the burial site for locus QpZAG110. All four leaves have different genotypes at this locus. The evidence leaves appear to have come from three different trees (Table 2), but not from the trees at the burial site.

4. Discussion

DNA typing of native plant species has only occasionally been used for solving criminal and civil cases. Given our growing knowledge of plant genomes and genetic variability, there exists an increasing opportunity to develop this resource for forensic applications. In our investigation, we were unable to produce genetic evidence that placed a suspect at a crime scene. However, this was not due to technical limitations of our method and our results demonstrate that this approach could and should be used in future forensic investigations. In our study, the genetic variability found at four microsatellite loci was sufficient to produce unique DNA profiles for trees, and the probability of two individuals in the population having the same...
profile was approximately one in 500,000 \((2.06 \times 10^{-5})\). Additional loci would dramatically decrease this number, but since the evidence leaves did not produce a matching profile, we did not increase the number of loci.

While we were able to obtain genetic profiles from evidence leaves that had been stored in paper envelopes for more than 5 years, the DNA was not of high quality and was difficult to amplify. Evidence collection teams should consider the application of DNA technologies to the organic material that they collect, and consider how such materials should be collected and stored. Commonly used approaches for storing plant material for DNA analysis include the use of silica or other drying agents or freezing. If possible, material from different species should be identified and stored separately to reduce cross-contamination. Certain types of tissues (e.g., young leaves or buds) typically yield better quality DNA than other tissues (e.g., dried leaves or bark), and this should be taken into consideration when evidence is collected.

The case mentioned in the introduction involving seed pods from a Palo Verde tree found in a suspect’s pickup truck has some noteworthy differences from our case [7,8]. In the seed pod case, additional physical evidence, notably a beeper, linked the suspect to the crime scene. The finding of seed pods of a distinctive desert plant cast more suspicion than the finding of live oak leaves in a Florida vehicle; live oaks are common species throughout the southeastern U.S. Indeed, live oak leaves were also collected from the victim’s car. The method used for the Palo Verde study was randomly amplified polymorphic DNA, or RAPDs, rather than the microsatellite approach we employed here. Both techniques have certain advantages and limitations for plant forensic applications. RAPD analysis can be done quickly and inexpensively and can be used on species without previous marker development or genome characterization. However, the reproducibility and reliability of RAPD markers have been questioned [35–37]. The loci characterized by the generic primers used in RAPD analysis are not well defined and data analysis must accommodate the dominant inheritance of RAPD markers.

Database construction of RAPD profiles is probably not possible, rather blind tests and genetic “line-ups” are probably the most suitable approaches for RAPD evidence.

Microsatellites require a significant development stage involving the construction and screening of a genomic library and testing of primers designed from sequences flanking repeats. Once developed, markers can generally be used on closely related species, as in this case among members of the ‘white oak’ section of Quercus. Our successful use of primers developed for other species of oak is consistent with other reports of successful transfer of primers among oak species [21,22,38–40]. Microsatellite alleles can be reliably sized and genotypes can be stored in a database. Since microsatellite analysis of plants is analogous to STR markers widely used in forensic applications in human populations, statistical procedures and estimations of match probabilities developed for human forensics (and the debates surrounding them) can be easily transferred to plant microsatellite data e.g., [41–46].

Although the additional time needed for microsatellite loci development will limit their application to certain types of plants, there is a growing list of plant taxa for which microsatellites are available. Trees may be more promising than smaller, more abundant plant species for forensic botany applications because trees can, in many cases, be thoroughly sampled around a crime scene. The somatic stability of microsatellites among leaves and branches of individual trees and among clonally propagated individuals has been established e.g., [12,47,48]. Since microsatellite primers have been published for both the “white” [28,29] and “red” oak [49] sections of the genus Quercus, oak leaves of any species could be genotyped. Microsatellites have also been developed for many other tree species, including larch [50], sycamore [51], eucalyptus [52], white pine [53] and Douglass fir [54]. Given that this list will continue to grow, evidence collection teams should be trained in botanical trace evidence and its potential to provide linkages between crime scenes and suspects. Forensic botanists should routinely be called to crime scenes to aid in identification of unusual plant species or species amenable to DNA profiling.

5. Conclusions

We have demonstrated that microsatellite markers developed for ecological studies can be applied to forensic botany, and have the potential to link suspects to crime scenes. DNA profiles were obtained from evidence leaves (Q. geminata) stored for 5 years, and were compared to living trees near a crime scene. A population survey indicated the average cumulative probability of identity using these four loci was one in 500,000. In this case, the DNA profiles of sand live oak leaves taken from a suspect’s car did not match those of trees near a gravesite, and could not provide physical evidence in this particular case. However, our results demonstrate that botanical genetic profiling, in particular microsatellite genotyping, has strong potential in solving criminal and civil cases, and growing knowledge of plant genomes will provide new opportunities in forensic botany.

Acknowledgments

The authors kindly thank Dr. David Hall and Ginger Clark for their assistance on species identification and sample collection. This work was completed in partial fulfillment of the doctoral degree (to KJC) from the graduate college at the University of Illinois at Chicago.

References


