Characterization of Microsatellite Loci and Reliable Genotyping in a Polyploid Plant, *Mercurialis perennis* (Euphorbiaceae)

TANJA PFEIFFER, ANNA M. ROSCHANSKI, JOHN R. PANNELL, Grażyna Korbecka, and Martin Schnittler

From the Institute of Botany and Landscape Ecology, Ernst-Moritz-Arndt-University Greifswald, Grimmer Str. 88, 17487 Greifswald, Germany (Pfeiffer, Roschanski, and Schnittler); the Department of Plant Sciences, Oxford University, South Parks Road, Oxford OX1 3RB, UK (Pannell and Korbecka); and the Department of Plant Breeding and Biotechnology, Institute of Soil Science and Plant Cultivation, Czartoryskich 8, 24–100 Puławy, Poland (Korbecka). Anna M. Roschanski is now at the Faculty of Biology, Department of Conservation Biology, Philipps-University Marburg, Karl-von-Frisch-Straße 8, 35043 Marburg, Germany.

Address correspondence to Tanja Pfeiffer at the address above, or e-mail: tanja_pfeiffer@gmx.de.

Data deposited at Dryad: http://dx.doi.org/10.5061/dryad.8887

For many applications in population genetics, codominant simple sequence repeats (SSRs) may have substantial advantages over dominant anonymous markers such as amplified fragment length polymorphisms (AFLPs). In high polyploids, however, allele dosage of SSRs cannot easily be determined and alleles are not easily attributable to potentially diploidized loci. Here, we argue that SSRs may nonetheless be better than AFLPs for polyploid taxa if they are analyzed as effectively dominant markers because they are more reliable and more precise. We describe the transfer of SSRs developed for diploid *Mercurialis huetii* to the clonal dioecious *M. perennis*. Primers were tested on a set of 54 male and female plants from natural decaploid populations. Eight of 65 tested loci produced polymorphic fragments. Binary profiles from 4 different scoring routines were used to define multilocus lineages (MLLs). Allowing for fragment differences within 1 MLL, all analyses revealed the same 14 MLLs without conflicting with merigenet, sex, or plot assignment. For semiautomatic scoring, a combination of as few as 2 of the 4 most polymorphic loci resulted in unambiguous discrimination of clones. Our study demonstrates that microsatellite fingerprinting of polyploid plants is a cost efficient and reliable alternative to AFLPs, not least because fewer loci are required than for diploids.

**Key words:** genotyping errors, loci characterization, microsatellites, polyploidy, SSR markers

Over the last decade, genotyping with highly polymorphic DNA markers has become an increasingly important tool in plant ecology, with microsatellites (for review, see Selkoe and Toonen 2006) and amplified fragment length polymorphism (AFLP) fingerprinting (Mueller and Wolfenbarger 1999) being the current tools of choice. Both these techniques generate highly polymorphic profiles that are ideally suited for identifying distinct genetic individuals (genets), including clonal offspring. Whereas AFLPs utilize anonymous markers that can principally be used in all taxa and genomes, microsatellite analyses need specific primers for amplification of the regions with the respective simple sequence repeats (SSRs). The initial development of these primers is rather cost and time consuming, although genomic maps available for a growing number of taxa allow for a growing number of taxa allow for an easier scanning for SSR loci (e.g., Temnykh et al. 2001; Morgante et al. 2002; Tuskan et al. 2004). In cases where primers have been developed for related taxa, cross-amplification can be attempted, and there are now numerous studies reporting successful interspecific transferability of SSR primers (Peakall et al. 1998; Fraser et al. 2005; Gao et al. 2005; Ducarme et al. 2008; Eusemann et al. 2009). Generally, such attempts are the most successful for closely related species (76% transferability within the same genus and 90% within subgenera), although even then the amplifying loci are not always polymorphic (Rosetto 2001).

Whereas both AFLPs and microsatellites can be used to study genetic diversity and structure, microsatellites have the distinct advantage for the study of diploid organisms because of their codominant expression, which allows an exact identification of genotypes. Unfortunately, many of the advantages of codominance are lost in the study of polyploids, even those with disomic inheritance, because the
allele dosage behind a specific peak can hardly be determined. As a result, SSR profiles of polyploid taxa are most commonly interpreted in terms of “allele phenotypes” (Esselink et al. 2004), and peak patterns are usually coded as dominant data (Gerber et al. 2000; Andreakis et al. 2009) largely comparable to anonymous markers like AFLPs or random amplified polymorphic DNA (RAPDs). For lower polyploids like tetraploids, it may sometimes be possible to determine allelic configurations (Landergott et al. 2006; Wirth et al. 2009), for example, by using specific techniques like microsatellite DNA allele counting-peak ratios (Esselink et al. 2004; Nybom et al. 2004) or other mathematical methods (Bruvo et al. 2004; Markwith et al. 2006), but this is usually not possible for higher ploidy levels. Nevertheless, polyploids may offer an important redeeming feature for population genetic analysis using microsatellites if their interpretation as dominant markers can be tolerated: Their SSR profiles are typically much more diverse, a fact that allows for fingerprinting with fewer markers, even if only allele phenotypes can be recognized.

Generally, the maximum number of possible genotypes (gt) in codominant microsatellite markers can be calculated through the formula that describes the number of k-combination with repetitions: $gt = \binom{n+k-1}{k}$; in our case, the number of elements (n alleles) equals the number of chosen elements (k chromosomes accommodating these alleles); hence, $gt = \frac{(2n-1)!}{n!(n-1)!}$, with n denoting both the ploidy level of the organism and the maximum number of different alleles at a locus in a single organism (excluding null alleles). Due to the difficulty (or impossibility) of distinguishing genotypes with the same alleles but in different allele dosages, only a lower number of distinct allele phenotypes (pt) can be readily discerned by their SSR profiles. When totally ignoring allele dosage (i.e., only noting presence or absence of alleles), a comparison of the number of all possible genotypes for different ploidy levels revealed a rule that can be described by the formula $pt = 2^{p} - 1$. Whereas for a diploid species, we thus obtain $gt = pt = 3$, for a tetraploid taxon, 35 possible genotypes will generate only 15 different allele phenotypes (for different phenotypic classes and inheritance patterns, see De Silva et al. 2005); for a decaploid, these numbers increase to 92 378 and 1023, respectively. Therefore, for highly polyploid species, allele phenotypes can generate a high diversity if an SSR locus is sufficiently polymorphic. Moreover, the amount of polymorphism is expected to be higher in polyploids because of the greater effective population size.

Another challenge presented by the SSR profiles of polyploid taxa concerns the difficulty of unambiguously identifying alleles (e.g., in the presence of stutter peaks). This issue renders it essential to rate the quality of the fingerprinting, that is, to assess the reproducibility of the profiles and to calculate error rates. Even in diploids, quality evaluation can be important but is often not reported (Hoffman and Amos 2005). For clonally growing plants, spatial position and sometimes also sexes or flowering morphs can be used as context information to recognize errors in identifying genetically distinct individuals, as demonstrated by Schnittler and Eusemann (2010) for the dioecious diploid Populus euphratica. Conflicts with this context information often indicate insufficient resolution of the marker system (compare Arnaud-Haond et al. 2005), that is, by wrongly merging similar but distinct genotypes into one multilocus genotype (MLG) (merging error, Schnittler and Eusemann 2010). High-resolution systems involving many loci reveal the opposite problem, that is, splitting a single genotype into 2 apparent MLGs. Unfortunately, such splitting errors are much more common because for a multilocus system, even a moderate per-locus error rate translates to a significant rate of per-sample errors, with most of them expressed as splitting errors. To compensate for this error, one or more deviating alleles may be allowed for MLG identity; we thus effectively deal with multilocus lineages (MLLs) as surrogates for genets. With the more frequent interference of stutter and allele peaks for SSR fingerprints from polyploid plants, the right choice of this threshold becomes especially important.

Here, we demonstrate the use of SSR analyses for the fingerprinting of individuals of the clonal decaploid Mercurialis perennis L. (Euphorbiaceae), after their development from cross-amplification of markers originally developed for the related species M. annua and M. huettii (Korbecka et al. 2010). In particular, we 1) develop a specific scoring routine for polyploid SSR in M. perennis, 2) determine the level of polymorphism at individual loci and assess their resolution, and 3) validate the resulting allele phenotypes and MLLs by comparison with context data (sex, merigenet, and spatial position) to evaluate error rates and reliability of the data set. In previous attempts to distinguish genets of M. perennis on the basis of AFLP fingerprinting (Vandepitte et al. 2009, 2010), the genetic patterns obtained were not correlated with morphological data such as sex or merigenet relations. Our study thus provides an important case in point, where as few as 2 SSR loci appear to be superior to broadscale AFLP profiles in distinguishing different MLLs.

**Materials and Methods**

Mercurialis perennis L. (Euphorbiaceae) is a dioecious forest herb with stable sex expression (compare Jefferson 2008). The species grows clonally, with branched subterranean stoloniferous spacers connecting the aboveground shoots (modules). Decay of these spacers splits genetic individuals into independent merigenets (for terminology, see Pfeiffer 2005). Cytotypes with different ploidy levels are reported throughout the distribution range (Krähenbühl and Küpfer 1995). The population examined in this study is most probably decaploid; 2C values of 2 samples measured by flow cytometry are comparable to those for decaploids recorded by Vandepitte et al. (2009); compare also Krähenbühl and Küpfer 1995).

For the initial screening, 6 M. perennis samples from different German and 1 Polish population were used. The final analysis included 52 leaf samples from three $1 \times 1$ m²
plots collected in the Elisenhain forest (nature reserve “Eldena”) near Greifswald (Mecklenburg-Western Pomerania, 54°05’N, 13°27’E). As an out-group, 2 specimens from Pöltchow (Mecklenburg-Western Pomerania, ca. 85 km away) and Glashütte (Brandenburg, ca. 230 km away) were included. In the Elisenhain plots, all plants were carefully excavated to obtain information about connections between modules. As context information, we recorded the spatial position of the samples within 1 of 3 plots A, B, or D (distance between plots was at least 200 m); merigenet identity (which modules were connected morphologically); and sex (male, female, or nonflowering modules and merigenets). With 21 merigenets comprising up to 4 connected modules analyzed, most (50 of 54 samples) functioned as natural replicates, allowing us to estimate the error rate in discriminating between genets.

DNA was extracted from air-dried leaf material (30–40 mg). First, we used a standard Cetyltrimethylammonium bromide (CTAB) protocol (Doyle JJ and Doyle JL 1990 with small alterations), and for final analyses, we used the Invisorb Spin Food Kit II (Invitek). DNA extraction followed the manufacturer’s protocol, with a few exceptions: instead of lysis buffer P, 1 ml prewarmed (60 °C) 2× CTAB buffer was used; RNA digestion time was extended to 30–45 min at 37 °C; and final centrifugation in the second washing step lasted 5 min. Elution of DNA was carried out in 2 steps using 50 and 25 μl prewarmed elution buffer D. The DNA content of each sample was estimated in a spectrophotometer (BioMate3, Thermo Scientific).

We used primer pairs designed by Korbecka et al. (2010) based on their M. huetii SSR library. As M. perennis is only distantly related to M. huetii (Krähenbühl et al. 2002; Obbard et al. 2006), we chose to test 65 out of 73 available primer pairs that had amplified successfully in M. huetii plus at least some lineages of M. annua in the cross-amplification tests carried out by Korbecka et al. (2010). We assumed that these primers were most likely to amplify M. perennis as well. We screened for amplification in M. perennis using M. huetii DNA as positive control. The PCRs contained 10–20 ng DNA template, 0.2 μM of each forward (F) and reverse (R) primers, 1× reaction buffer supplied with polymerase (AppliChem) with 2.5 mM MgCl₂, 0.2 mM dNTP mix (Piqlab), 0.25 μg bovine serum albumin (New England Biolabs), 0.45 U Tag-Polymerase (AppliChem), and ddH₂O in 10 μl final volume. The loci were amplified in a thermocycler (Eppendorf Mastercycler) using the following cycle profile: 3 min at 94 °C, 30 cycles of 30 s at 94 °C, 30 s at 55 °C, and 45 s at 72 °C and an additional 20 min at 72 °C and cooling to 4 °C. PCR products were screened for bands after electrophoresis in 2% agarose gels stained with ethidium bromide.

PCRs of loci that produced no PCR bands or that differed in size by more than ~100 nt from the ones in M. huetii were repeated at least once. Successfully amplified loci with products of roughly the expected lengths were tested further. For these tests, fluorescently labeled forward primers were used with the same reaction mix (see above) and the following step-down cycle protocol: initial 3 min at 94 °C, followed by 20 cycles of 30 s at 94 °C, 30 s at locus-specific annealing temperature (Tₘ; Table 2), and 45 s at 72 °C and another 10 cycles of 30 s at 89 °C, 30 s at Tₘ, 5 °C, and 45 s at 72 °C, followed by 20 min at 72 °C before cooling down to 4 °C. PCR products were checked in agarose gels and diluted depending on band quality (2.5- to 10-fold). One microliter of the dilution was mixed with 8.85 μl HiDi formamide and 0.15 μl GeneScan 500 ROX size standard (Applied Biosystems), denatured for 5 min at 95 °C, and cooled down before screening the samples in an ABI Prism 310 Genetic Analyzer for length polymorphisms. For the final analysis, 8 polymorphic loci were tested on 54 samples of M. perennis. Products from singleplex PCRs were pooled for simultaneous runs of 4 loci per sample.

To analyze SSR loci in a highly polyploid taxon, a suitable analysis routine had to be developed, including a check for reliability in distinguishing genets. Sample profiles were scored blindly in random order using GeneMapper v3.7 (Applied Biosystems). Three different peak-scoring routines were applied, a mainly automatic scoring (further referred to as AUTO), a semiautomatic (SEMI), and a manual routine (MANUAL), aiming at discerning different peak types. For all routines, the same bin sets were applied, and independent binary profiles were generated for each locus, coding presence (1) and absence of peaks (0). The matrices for 8 individual loci were combined for further analyses (multilocus approach).

For AUTO, the scoring was carried out using default settings but allowing a maximum of 10 alleles per locus (corresponding to the maximum number of alleles in a decaploid). In rare cases, we manually corrected for minimal peak shifts and deleted pull-up peaks from other loci. For the SEMI scoring, we manually added to AUTO profiles 1) post-peaks directly following scored peaks and

---

**Table 1** Scores and qualities of different peak types distinguished in the MANUAL scoring routine of the SSR profiles (compare with Figure 1)

<table>
<thead>
<tr>
<th>Score</th>
<th>Peak characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Main, high peak, and adjacent peaks with at least 90% (forward) or 85% (back) of the height of the respective main peak</td>
</tr>
<tr>
<td>9</td>
<td>Distinct main peak, significantly smaller than those scored with 10 in same profile</td>
</tr>
<tr>
<td>8</td>
<td>Post-peaks: high peak after main peak, attaining &gt;50–85% of main peak height</td>
</tr>
<tr>
<td>7</td>
<td>Pre-peak: distinct, high peak before main peak (&gt;66–90% of its height); fused shoulder pre-peaks were only scored as 3</td>
</tr>
<tr>
<td>6</td>
<td>Rather small but distinct peak</td>
</tr>
<tr>
<td>5</td>
<td>Higher peak between adjacent main peaks cannot be assigned unambiguously to either main peak and/or identified as stutter versus smaller distinct peak (rarely used)</td>
</tr>
<tr>
<td>4</td>
<td>Smaller post-peak (10–50% of main peak height or ~50% of peak with score 8)</td>
</tr>
<tr>
<td>3</td>
<td>Smaller pre-peak (&gt;33–66% of main peak height)</td>
</tr>
<tr>
<td>2</td>
<td>Putative stutter peak with height &gt;50 rfu</td>
</tr>
<tr>
<td>1</td>
<td>Putative stutter peak with height ≤50 rfu</td>
</tr>
</tbody>
</table>
reaching at least 10% of their height, 2) peaks (≥50 rfu) not identified in the AUTO mode, and 3) peaks of very similar height to adjacent scored ones. Small peaks between neighboring large peaks were only scored if they were clearly recognized as post-peaks of the first peak. In the MANUAL setting, each peak in the electropherograms was scored manually and given peak quality scores ranging from 1 to 10 (Table 1). These scores were applied in descending order with declining putative peak quality, that is, with higher scores for presumable post-peaks (because stutter bands usually occur before main peaks) and lower scores for obvious stutter (see Table 1, Figure 1). For data analyses, these peak scores were translated into different binary matrices, applying a peak quality threshold to exclude all peaks with lower scores. (Example: Using a peak quality threshold of 8, all peaks with scores of 8–10 were coded as present in a scoring routine called MANUAL 8.)

To differentiate allele phenotypes (in the following, we refer to these as SSR phenotypes because the allelic nature of peaks was not verified) and to check for conflict with context information, we modified the Excel routine described in Schnitller and Eusemann (2010). Basic functions of this routine are comparable with those of the program GenClone (Arnaud-Haond and Belkhir 2007). However, we applied a variable allowance of 0, 1, 2, … deviating peaks for the identity of SSR phenotypes, thus effectively defining MLLs (Arnaud-Haond et al. 2007). For values greater than zero, we defined all SSR phenotypes that differed from at least one other SSR phenotype by a number of peaks below the respective threshold as 1 MLL (clone), even if 2 members of this MLL differed by a number of peaks above the threshold. This routine enabled us to test variable allowances of tolerated fragment differences and to compare with frequency histograms of pairwise genetic distances between samples (as presented in Douhovnikoff and Dodd 2003). The binary matrices for AUTO, SEMI, and for selected peak score thresholds of MANUAL were further analyzed in PAUP 4.0b (Swofford 2002), calculating neighbor joining and unweighted pair group method with arithmetic mean dendrograms and implying bootstrap searches with each 100 000 replicates.

To assess the resolution of our system with different scoring routines, we calculated locus accumulation curves by selecting all combinations $C^{l}_{n}$ from the $l = 8$ available loci with $1 ≤ n ≤ 8$ as in the respective routine in GenClone but additionally monitored conflicts with context information. If the obtained curves leveled off and did not result in conflicts, the resolution of the marker system seemed to be sufficient for genet discrimination using the respective data set (Schnittlter and Eusemann 2010). Conflicts with context information include 1) assigning samples of 1 merigenet to different MLLs, 2) recording the same MLL in different plots, or 3) having shoots of opposite sex in 1 MLL. The first problem represents a splitting error and the latter 2 are merging errors in the sense of Schnitller and Eusemann (2010).

Using DNA sequencing, microsatellite regions in *M. perennis* were examined to compare them with *M. huetii* SSR library clones. For each of the 8 loci used in the analysis, 2 PCR products (1 male and 1 female sample) with low peak numbers were used as a template. Sequencing reactions further contained the BigDye 1.1 terminator kit, sequencing buffer (both Applied Biosystems), and unlabeled primer.
After cycle sequencing for 2 min at 96 °C, 24 cycles with 10 s at 96 °C, 5 s at T_a, 4 min at 60 °C, and cooling to 4 °C, sequencing products were purified in 0.2-ml strips using the Dynabeads Sequencing Clean-Up kit (Invitrogen) according to the manufacturer’s protocol but with an additional washing step with 85% ethanol. The samples were run on a Genetic Analyzer 3130xl (Applied Biosystems); the resulting sequences were checked manually for the presence of microsatellite motifs. For most loci, only one primer was used in the sequencing reaction; in cases of ambiguity, both forward and reverse sequences were checked.

Results and Discussion

Cross-amplification of SSR markers from *M. huetii* in *M. perennis* was successful in 12 of the 65 loci tested. Eight loci were polymorphic and suitable for distinguishing genets: Mh01, Mh35-2, Mh57, Mh61, Mh72, Mh80, Mh81, and Mh83. Of these, only Mh35-2 has been characterized before (Korbecka et al. 2010). These 8 loci were used to assign 54 samples to MLLs. The lengths of PCR products of the individual loci in *M. perennis* correspond fairly well to their respective lengths in *M. huetii* (Table 2).

The sequencing revealed identity of simple repeat motifs between *M. huetii* and *M. perennis* in Mh72, Mh80, Mh81, and Mh83 (Table 2). In Mh57, the 2 sequenced *M. perennis* samples showed a slightly altered motif [(CA)_n (GA)_n repeat] but no consequences for the dinucleotide repeat structure. Even the complex motif of Mh35-2 in *M. huetii* and *M. annua* (Korbecka et al. 2010) was present in *M. perennis* with slight alterations (Table 2), highlighting the interspecific transferability of the SSR loci. For these loci, observed small (±1 nt) differences in fragment lengths in some samples most probably indicate single nucleotide indels in the different SSR copies, resulting in fragment lengths departing from typical repeat distances but not affecting the motif itself. Korbecka et al. (2010) also detected such small differences in the allele lengths for loci originating from the same SSR library and attributed this length variation to mononucleotide runs detected in the sequence next to the microsatellite region. Because most sequenced loci in *M. perennis* also showed short mononucleotide repeats (especially poly-A and T), this might explain the observed length differences.

For only 2 loci, Mh01 and Mh61, the repeat motifs remain ambiguous because of difficulties in sequencing the microsatellite region. For these loci, observed small (±1 nt) differences in fragment lengths in some samples most probably indicate single nucleotide indels in the different SSR copies, resulting in fragment lengths departing from typical repeat distances but not affecting the motif itself. Korbecka et al. (2010) also detected such small differences in the allele lengths for loci originating from the same SSR library and attributed this length variation to mononucleotide runs detected in the sequence next to the microsatellite region. Because most sequenced loci in *M. perennis* also showed short mononucleotide repeats (especially poly-A and T), this might explain the observed length differences.

Out of the 4 scoring routines compared, that is, AUTO, SEMI, MANUAL 8, and MANUAL 4, AUTO resulted in the lowest mean fragment numbers scored for all loci (24.5 ± 2.9), whereas the highest numbers were obtained for SEMI and MANUAL 4 (36.5 ± 3.6 and 38.7 ± 4.0, respectively). The least polymorphic were loci Mh01, Mh72, and Mh81, showing lowest minima as well as lowest mean numbers of fragments per sample. The loci Mh35-2, Mh57, and Mh80 performed best (Table 2, see also Supplementary Material, Supplement 1).

Peak scoring and the generated binary profiles proved to be rather constant for modules of the same merigenet (natural replicates), different runs, and independent PCR
products (data not shown) but also for different scoring routines. Genet discrimination is highly reliable: The same 14 MLLs are identified with all 3 scoring routines and as clusters with (nearly) maximum bootstrap support (≥94–100%) in the dendrograms (compare Figure 2). The 2 samples from outlier populations Glashütte and Pöltchow are readily differentiated as distinct, singular genets. The 52 modules (2 singular samples and 21 merigenets with each 2–4 analyzed modules) from the Elisenhain belong to 12 clonal MLLs, each comprising 2–11 samples from 1 to 4 independent merigenets (Figure 2).

In all scoring routines, these 14 MLLs were obtained without conflicts with context information over a broad range of tolerated fragment differences (Figure 3). SEMI, MANUAL 7, and MANUAL 8 seem to be the best scoring routines because, in order to avoid splitting errors, only 4 fragment differences within 1 MLL have to be allowed. A higher allowance is required for AUTO and other MANUAL routines (Figure 3). These values are additionally verified by the frequency histograms of pairwise comparisons between samples, which showed bimodal distributions with pronounced minima (data not shown).

Figure 2. Neighbor joining (NJ) dendrograms for SEMI (left) and MANUAL 8 routines (right; different scales). Numbers at branches are bootstrap values (>50% (NJ search with 100 000 replicates in PAUP). Context data are coded in the sample tag (with letters A, B, and D referring to plots) and color of the boxes (sex): The 12 MLLs from Elisenhain consist of 5 males (gray) and 7 females (blank); the numbers of merigenets assigned to these MLLs are listed in brackets. Female 3 (marked by asterisk, different unsupported position in both trees) includes a nonflowering merigenet and a single female sample allowing for sex determination in the former.
If fewer fragment differences are allowed within 1 MLL, splitting errors occur, artificially augmenting the number of genets. They can be reliably detected as they assign modules from 1 merigenet to different MLLs (merigenet conflict). In contrast, very high allowances for fragment differences reduce the resolution of the system and introduce merging errors that wrongfully reduce the number of MLLs. Merging errors can only be seen by conflicts with plot and/or sex assignment and thus can go undetected in cases where the 2 merged genets 1) belong to the same plot (probability \( P > 0.33 \) for 3 plots) or 2) belong to the same sex (\( P > 0.5 \)). With unequal numbers of samples between plots, biased sex ratios, or a significant proportion of nonflowering plants (of unknown sex), these probabilities increase further (Schnittler and Eusemann 2010). To minimize these problems, we included in our study mostly flowering modules from 3 different plots, which should be of different origin due to the limited colonization capacity of \( M. \) perennis (cf. Jefferson 2008).

For genotyping in diploid taxa, most often 8–12 loci are required for reasonable precision (Arnaud-Haond et al. 2005; Eusemann et al. 2009; see also Koskinen et al. 2004). Usually a system with fewer loci lacks resolution, indicated by merging errors (Arnaud-Haond et al. 2005; Schnittler and Eusemann 2010). However, with more loci, the error rate in distinguishing genets increases in a nearly linear fashion due to multiplication of genotyping errors at single loci (Hoffman and Amos 2005). Due to the larger number of alleles per locus, the discriminative power of a locus is potentially higher for polyploids than for diploids, even when only allele phenotypes can be differentiated (allowing up to 1023 possible phenotypes in a decaploid compared with only 3 genotypes in a diploid organism). Successful fingerprinting of polyploids with a high level of precision will thus typically require substantially fewer loci than equivalent diploids. Our study of decaploid \( M. \) perennis bears this out.

To assess the resolution of the marker system adopted here, locus accumulation curves were generated. Figure 4A,B shows the results for the AUTO and SEMI scoring routines, allowing 7 and 4 fragment differences within 1 MLL, respectively. For the latter scoring mode, any combination of 5 loci, irrespective of individual degree of polymorphism, results in unambiguous assignment to MLLs. The same was the case for all but one of the combinations of 4 loci and even for 6 combinations of only 2 loci (including \( Mh35-2, Mh57, Mh80\), and \( Mh83\); see upper dotted line in Figure 4B). In the AUTO scoring mode, the values were not as good: only the 3 most polymorphic loci together (\( Mh35-2, Mh57, Mh80\)) allowed for reliable discrimination of clones; in most cases, the maximum number of discerned MLLs was reached only by combinations of all 8 loci. Figure 4C,D shows the corresponding curves for mean conflicts per sample, which decreases slower for AUTO than for SEMI routine.
Given the high reliability and discriminative power obtained with the SEMI routine, we strongly recommend this scoring method; it provides high performance with little additional effort. In contrast to the MANUAL scoring, SEMI does not require further laborious differentiation of peak types. Compared with the fully automatic method (AUTO), the addition of post-peaks in particular adds much information. However, including a higher number of loci and a somewhat higher allowance of fragment differences for MLL identity, a fully automated scoring routine yields the same results. Two (SEMI) or 3 (AUTO) of the most polymorphic markers are sufficient for unambiguous recognition of clones and clonal replicates in our study (Figure 4A,B). In the combined analysis of all 8 loci, however, the optimum constellation (14 MLLs, no conflicts) is more robust in SEMI (and MANUAL) than in AUTO, and fewer fragment differences have to be tolerated for MLL identity (Figure 3).

Of course, the codominant state of microsatellite loci loses its utility in higher polyploids, as the dosage of individual alleles can usually not be determined (see in the Introduction). Furthermore, the discrimination of low-copy alleles is often hampered by stuttering. As long as stutter peaks are reproducible, they do not diminish the feasibility of SSR for the purposes of distinguishing genets. Due to these features, SSRs for polyploid organisms are a partially anonymous marker system that may be substantially better when interpreted analogously to AFLP or RAPD profiles than with normal SSR analyses. Technically, however, they enjoy the advantage of the classical SSR method, especially the use of specific primers in a single PCR reaction. Due to short DNA fragments amplified by specific primers, DNA degradation and contamination (often a problem in field-collected material) are much less critical than in AFLP fingerprinting, and the resulting rate of experimental errors in assigning genet identity should be lower. Moreover, AFLPs are prone to contamination by DNA from other organisms such as phytoparasitic fungi, which cannot be separated from the target. In polyploid organisms, the potential advantage of AFLP over SSR markers (i.e., their ability to generate many polymorphic markers with only a few primer combinations) is compensated by the higher discriminative power of SSRs.

SSR fingerprinting in *M. perennis* offers new opportunities for small-scale studies of the biology of the species. Even in the small data set analyzed in this study, up to 4 physically independent merigenets were identified as members of the same clone (MLL), highlighting the relevance of clonal reproduction in the reproductive and habitat colonization strategy of *M. perennis*. Additionally, the analyses may help to assign nonflowering modules to their sex if flowering samples of the clone are included (compare Figure 2). Using AFLP fingerprinting, Vandepitte et al. (2009, 2010) have provided first genotypic data on sex ratios and spatial patterns in *M. perennis*. However, with established polymorphic markers with sufficient resolution, the SSR technique is a real alternative to AFLP, even for polyploid organisms.

**Figure 4.** Locus accumulation curve for the AUTO (A) and SEMI scoring mode (B), applying thresholds of 7 and 4 different fragments for MLL identity, respectively. Shown are (A, B) mean numbers of differentiated MLLs with standard deviations (solid lines) and combinations of loci performing best and poorest (upper and lower dotted lines, respectively) and the resulting mean sex (C) and plot (D) conflicts per sample for AUTO (triangles) and SEMI (circles) routine, with standard deviations in solid and dotted lines, respectively.
Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding

Käthe-Kluth scholarship from the Ernst-Moritz-Arndt-Universität Greifswald to T.P.; Natural Environment Research Council (UK) to J.R.P.; and FP7 EC project “PROFICIENCY” (contract no. 245751) to G.K.

Acknowledgments

We are much indebted to Tristan Lemke (Rostock) for providing the specimen from Polchow; to Pascal Eusemann, Anja Klahr, and Anne-Marie Fiore-Donno for useful hints, discussions, and technical assistance in the laboratory; to Martin Haase (Institute of Zoology, Greifswald) for assistance with sequencing; and the Staatliches Amt für Umwelt und Natur Uckermark for permission to collect in the nature reserve “Eldera”.

References


**Received September 24, 2010; Revised February 10, 2011; Accepted March 10, 2011**

**Corresponding Editor: James L. Hamrick**