Theor Appl Genet (1996) 93: 249–256

M. B. Hamilton · D. M. Rand

Relatedness measured by oligonucleotide probe DNA fingerprints and an estimate of the mating system of Sea Lavender (*Limonium carolinianum*)

Received: 6 November 1995 / Accepted: 19 January 1996

Abstract Using DNA fingerprint markers within species and populations of wild plants requires information on the relationship between fingerprint similarity and relatedness. We identified a hypervariable marker based on oligo-$(GATA)_4$-hybridization of *Dpn*II-cut genomic DNA from Sea Lavender (*Limonium carolinianum*). Banding patterns were somatically stable and highly variable among unrelated individuals. Band molecular-weight sizing errors (as a percent of band molecular weight) were estimated at 0.44%±0.003 within gels and 0.76%±0.964 between gels. Band sizing errors defined a 99% confidence bin of ±0.95% (1.90% total) of molecular weight. Band-sharing estimates were based on this bin size and on variance estimates that compensate for non-independent comparisons. Band-sharing among nine unrelated individuals ($\bar{\theta}$) was 0.198±0.011. Experimental pollinations designed to produce selfed, full- and half-sib progeny groups led to five selfed progeny groups and no outcrossed progeny (mean band-sharing, $\bar{S}$=0.468±0.074). A linear regression between band-sharing (S) and relatedness (r) assuming 17% inbreeding was r=0.006+0.914*S (R^2=0.973) and established the maximum amount of inbreeding. $\bar{S}$(0.392±0.022) estimated from wild pollinated seeds from four maternal families was intermediate to unrelated individuals and experimental selfed progeny, giving evidence for mixed mating in wild plants. More extensive plant pedigrees with known levels of inbreeding will be needed to measure variation in the relationship between S and r among populations and families.

Key words *Limonium carolinianum* · Plumbaginaceae · Multilocus fingerprint · Relatedness · Mating system

Communicated by p. M. A. Tigerstedt

M. B. Hamilton (✉)[1] · D. M. Rand
Department of Ecology and Evolutionary Biology,
Brown University, Box G-W, Providence, RI 02912, USA

Present address:
[1] National Zoological Park, Genetics Laboratory,
3001 Connecticut Avenue, NW, Washington, DC 20008, USA

Introduction

DNA fingerprinting has emerged as a powerful tool to examine identity and hereditary relationships among organisms. Rapidly evolving DNA markers created by variable numbers of tandem repeats (VNTR) have been especially useful in constructing pedigrees and in tracing lineages for those species that have low levels of variability in organelle and nuclear DNA markers or electrophoretic enzyme phenotype. The use of VNTR markers has facilitated the direct examination of genetic processes that operate over evolutionarily short-time spans (such as mate choice, kin selection, population subdivision) in a variety of mammals, birds and fishes (e.g., Packer et al. 1991; Haig 1993; Rave et al. 1994). Progress in identifying and using hypervariable genetic markers to test ecological and evolutionary hypotheses in plants has been somewhat slower. To-date, DNA fingerprinting in plants has been used most successfully to identify variation within and among cultivars and their wild relatives (e.g., Nybom 1991; Nybom and Hall 1991; Jung et al. 1993). Several studies with plants have found that multilocus fingerprints produced with repeat probes can distinguish evolutionary relationships among cultivars (Weising et al. 1994b) and identify ramets and genets of clonal plants (Van Heusden et al. 1991). There have been fewer applications of multilocus fingerprints to genetic studies of wild plant species and populations (reviewed in Weising et al. 1994a). More information on inheritance patterns, linkage among bands, and the relationship between fingerprint similarity and relatedness is needed before VNTR fingerprinting in wild plants can be widely applied to tests of evolutionary hypotheses within species and populations.

In addition to determining parentage, VNTR markers can be used to estimate the average relatedness within and among populations (however defined) if there is a non-ran-

dom relationship between banding pattern similarity and true relatedness. Identifying evolutionary phenomena, such as population subdivision, isolation by distance or adaptation with DNA profile similarity measures, requires quantitative tests of relatedness. In order to estimate relatedness accurately, the correlation between the similarity of multilocus banding patterns and overall relatedness must be significantly different from zero and comparisons of mean similarity measures will become more powerful tests of relatedness as the slope increases (Lynch 1988 1991). This correlation has been examined in a number of animal species and found to be significantly greater than zero, although not always large enough to provide a powerful measure of relatedness (Gilbert et al. 1991; Piper and Rabenold 1992; Haig et al. 1993; Butler et al. 1994). We know of no published studies that use multilocus fingerprints to examine the relationship between band-pattern similarity and relatedness within populations of a wild plant species.

Here we present data on a newly identified hypervariable microsatellite marker in the coastal plant Sea Lavender (*Limonium carolinianum*) which is being used for a broader study of population genetic structure. For this marker we estimated molecular-weight assignment errors for DNA-profile bands within and between gels. These errors were then used to construct intervals around point estimates of band molecular weight, or fixed-bins, for every band in order to obtain band-sharing estimates. Presumably unrelated individuals and arrays of progeny from controlled crosses were used to estimate the expected relationship between band-sharing and relatedness. Relatedness was also estimated from mean band-sharing scores of several wild collected maternal families of progeny to obtain an estimate of the mating system within one population.

Materials and methods

The species and breeding design

L. carolinianum, or Sea Lavender (hereafter *Limonium*), is a rosette-forming perennial herb that produces leathery spatulate leaves from a woody rootstock (Luteyn 1990). *Limonium* occurs in the intertidal zone of salt marshes and rocky beaches on the Atlantic coast of North America. Plants have one to more than 20 inflourescences that branch and bear many tiny (6 mm) purple to white flowers. Flowering occurs from late July through October and plants produce at most one seed per flower. Pollinator-exclusion experiments in the greenhouse (Baker 1953; Hamilton, unpublished data) indicate that the species is self-compatible, as are other *Limonium* species (Luteyn 1990).

In July of 1993 eight haphazardly selected individuals (each >10 m apart) with developing inflorescences were dug from a salt marsh population of *Limonium* at Rumstick Cove, Barrington, Rhode Island. These plants were potted in peat-based media (Fafard Growing Mix No. 2), placed in a greenhouse, watered as necessary and fertilized weekly. When the plants began to flower in early August they were kept in groups of two or three plants in growth chambers to exclude potential pollinators and reduce the chance of accidental pollen exchange. Self and cross pollinations among plants with open flowers and dehiscent anthers were made daily until flowering ceased. Pollen was transferred with bee-sticks, which were used only once and then discarded. Pollen was observed to adhere to the bee-sticks and stigmatic surfaces under a dissecting microscope. Unfortunately, *Limonium* flowers are too small to bud-emasculate in large numbers. Although every effort was made to transfer pollen to stigmas without dislodging self pollen, it is possible that some or all of the presumed crosses were actually selfed progeny.

When seeds were mature, all labeled fruits were individually removed from inflorescences and sorted by cross. Seeds were planted in 12 × 24 plug trays, germinated on a mist bench with bottom heat, and seedlings were then transplanted into cell packs (TFI Plastics 806). From 255 self pollinations and 205 biparental crosses, 58 progeny (two selfed and three outcrossed families) survived until DNA extraction. Additionally, wild-pollinated seed families were collected from permanently marked plants in a rocky beach population of *Limonium* at Brown University's Haffenreffer Grant property in Bristol, Rhode Island. Fifteen Haffenreffer plants were randomly selected, 36 seeds per family were sampled and grown as above. When plants were about 8-months old and had grown several large leaves, 1–3 g of leaf tissue was harvested for DNA extraction.

DNA profiles

Fresh leaf-tissue samples were powdered under liquid nitrogen in a mortar and pestle and total genomic plant DNA was extracted using the CTAB method of Doyle and Doyle (1987) including the addition of extra NaCl/CTAB solution and final precipitation with polyethelene glycol and NaCl as described by Mak and Ho (1993). Final pellets from DNA extraction were re-suspended in 100 µl TE [10 mM Tris-HCl (pH 8.0), 1 mM EDTA]. These methods were critical to the quality of the DNA profiles obtained, with poorer quality DNA and uninterpretable profiles resulting when using several other extraction methods.

Twelve microliters of plant DNA were digested with ten units of restriction endonuclease in the supplied buffer according to the manufacturer's protocol for 5 h to overnight. The digested DNA was separated on 0.7% agarose, 1×TAE gels run in "Buffer Puffer" gel boxes (Owl Scientific) at 1.1 V/cm for approximately 31 h. Included on each 24-lane gel were five lanes of a molecular-weight marker composed of a 1-kb ladder and the gel-slice purified (Heery et al. 1990) 23 130-bp band of *Hin*dIII digested lambda DNA (both standards from Gibco BRL). No plant profile was more than three lanes distant from a molecular-weight marker lane. Gels were de-purinated for 5 min in 0.242 N HCl, given two 20-min washes in 1.5 M NaCl/0.5 M NaOH, washed once for 40 min in 1.5 M NaCl/1 M Tris-HCl (pH 7.5), and capillary blotted to nylon membranes (0.45 mm Magnagraph, Micron Separations Inc.) using 10×SSC as a transfer buffer (Sambrook et al. 1989). After blotting, membranes were baked at 65°C for 2 h and UV crosslinked.

Membranes were pre-hybridized in 25 ml of hybridization solution (7% SDS, 1 mM EDTA, 0.263 M Na_2HPO_4, and 1% fraction-V bovine serum albumin; Westneat 1988) for at least 5 h. Seventy five picomoles of an oligonucleotide probe and 100 ng of 1-kb ladder DNAs [treated with alkaline phosphatase (Boeringer Mannheim) to remove 5′ phosphate groups] were end-labeled with polynucleotide kinase (New England Biolabs) in separate reactions with supplied buffer and 5 µl of $\gamma dAT^{32}P$ (NEN Dupont). Labeling reactions were incubated at 37°C for 30 min and then diluted without removal of unincorporated nucleotides in 10 ml of hybridization solution. Hybridizations proceeded overnight at room temperature plus five degrees (about 26–30°C) in a rotating canister hybridization oven. Membranes were washed in 2×SSC/0.1% SDS for 15 minutes once at room temperature, twice at 37°C, once at 42°C and rinsed in room temperature 2×SSC. The 42°C final wash temperature is in the lower range of melting temperatures for a 16-bp oligonucleotide in 2×SSC solution and is therefore a moderate-to-high-stringency wash (section 6.4, Ausubel et al. 1993). Membranes were sealed in Cling Wrap and exposed to X-ray film (Fuji RX) for 24 h to 21 days with two intensifying screens. Most membranes required two exposure times, one for strongly hybridizing lanes and a longer duration for less intense lanes.

Limonium DNAs were surveyed for restriction fragment length variation using simple repetitive sequence probes in an effort to identify a "DNA fingerprint" marker capable of uniquely distinguishing among individuals in wild populations. DNA from four individual

plants was digested with each of 14 restriction endonucleases (*Alu*I, *Bam*HI, *Bfa*I, *Dde*I, *Dpn*II, *Eco*RI, *Hin*dIII, *Hha*I, *Hin*fI, *Msp*I, *Pa-ll*, *Pst*I, *Rsa*I, and *Sma*I from New England Biolabs and Promega) and hybridized to each of three oligonucleotide probes, $(GATA)_4$, $(GTG)_5$, and $(GGAT)_4$ (Weising et al. 1991a,b). DNAs hybridized with $(GTG)_5$ and $(GGAT)_4$ produced few, if any, resolvable bands and in most cases showed very little hybridization. DNAs cut with *Alu*I, *Bam*H I, *Dde*I, *Dpn*II, *Hha*I, *Hin*dIII, *Pal*I and *Rsa*I showed moderately to highly variable banding patterns when probed with $(GATA)_4$. The *Dpn*II-$(GATA)_4$ combination was chosen for further use due to the large number and wide size range of resolvable fragments it produced, as well as the high signal-to-background ratio of the resulting DNA profiles.

Band scoring and relatedness estimates

All autoradiographs were scored for the molecular weight of bands that composed individual plant DNA profiles using a transmission scanning densitometer (BioRad Model GS-670) and Molecular Analyst software (version 1.1.1, BioRad). The 23 130 bp and the 1018-bp marker bands were used to define the upper and lower limits of each DNA profile. A ruler and sharp tip marker were used to draw lines across the experimental lanes between these "spike" bands for each set of two marker lanes. These lines defined the top and bottom of each experimental lane. The largest 12 bands (12 216 to 1636 bp) of the "1-kb ladder" molecular-weight size standard were used to estimate the relationship between fragment mobility and log molecular weight which was then used to score fragments of unknown size. Autoradiograph films were scanned at 200 µm resolution with edge enhancement set at "sharp" and the filter set at "grayscale". Each profile was "smoothed" once over 5 pixels, the minimum setting, to reduce signal noise. "Peak Criteria" options depended on the intensity and background levels of individual profiles but minimum peak width was never less than 1 mm or more than 2 mm. Every attempt was made to set "Peak Criteria" heights and widths so only absorbance peaks clearly visible as bands on film were assigned as profile bands. Peak assignment by the densitometer software was interpreted literally, even in cases of wide absorbance peaks that scored as one band but were most likely two bands with little separation due to background. Infrequently, bands were clearly visible on film but were not detected as peaks, or else film artifacts (e.g., spots) were identified as peaks. In these cases peaks were added to or removed from DNA profiles.

The similarity of DNA profiles between pairs of individuals was determined using band-sharing coefficients

$$S_{xy} = \frac{2n_{xy}}{n_x + n_y} \tag{1}$$

where x and y are individual plants that share n_{xy} bands and n_x and n_y are the total number of bands observed in each individual (Lynch 1988 1991). DNA molecular weights estimated from gel electrophoresis are subject to the usual point-estimation errors of any measurement. Running an identical DNA fragment on an agarose gel a number of times will produce a distribution of estimated molecular weights with a mean value that is very close to the true weight of the fragment but with a non-zero variance. Further, measurement error is introduced by comparing DNA profiles from different gels, a necessity when comparing DNA profiles of a large number of individuals. In order to account for this within- and between-gel molecular-weight estimation error, all DNA fragments were treated as molecular-weight ranges or "variable bins" (Budowle et al. 1991; Weir 1992). Including this step when estimating S_{xy} requires estimating the upper and lower range of a band for individual x given some error value and then asking if any bands in the profile of individual y fall inside this bin. This method should over-estimate S_{xy} because VNTR repeat units that potentially differentiate "alleles" are very small (4 bp in this case) compared to even small fragments within DNA profiles (Budowle et al. 1991). The size of bins used in band-sharing calculations was derived from direct estimates of band-sizing error within gels (one profile scored with two molecular-weight size standards on one gel) and between gels (the profile of one individual scored on two separate gels).

Band-sharing estimates within groups of a given relatedness were based on all unique pairwise comparisons among individual DNA profiles. Because a given profile is employed in multiple comparisons, the data points used to compute mean band-sharing, $\bar{S}$, are not independent. In order to estimate the entire sampling variance of $\bar{S}$ we must measure the variance due to independent profile comparisons and add to it the co-variance due to non-independent profile comparisons (Lynch 1991). The sampling variance of $\bar{S}$ for all pairwise band-sharing comparisons of DNA profiles is then

$$\mathrm{var}(\bar{S}) = \frac{n\,\mathrm{var}(S_{xy}) + 2n'\,\mathrm{cov}(S_{xy}, S_{xz})}{n^2} \tag{2}$$

where n is the total number of band-sharing comparisons used to estimate $\bar{S}$ and n' is the number of those band-sharing measures that share an individual. The variance term is due to independent comparisons and the co-variance term due to comparisons that share members (see Lynch 1991 for full formulas). If the co-variance estimator in equation 2 provided negative values, the absolute value of the covariance was used instead. This correction inflates the overall variance estimate and provides very conservative estimates. Confidence intervals for $\bar{S}$ estimates were constructed using the variance of equation 2 and the number of pairwise comparisons [N(N–1)/2]. All fixed-bin band-sharing and variance calculations were performed with original C programs written by, and available from, the corresponding author.

In order to utilize a VNTR marker to test evolutionary hypotheses that involve genetic relatedness, it is necessary to first calibrate band-sharing with known genetic relatedness. Band-sharing of DNA profiles is expected to increase linearly with the relatedness of individuals according to

$$E(S) = \bar{\theta} + r(1 - \bar{\theta}) \tag{3}$$

where r is the degree of relatedness and $\bar{\theta}$ is the average band-sharing among non-relatives (Lynch 1988 1991). Estimating $\bar{\theta}$ within a population can be accomplished by determining $\bar{S}$ for a group of non-relatives (r=0). Further, it is possible to directly estimate the slope of the relationship between band-sharing and relatedness (and the amount of variance in $\bar{S}$ for a given relatedness class) by determining $\bar{S}$ for individuals of known relatedness. Using the relationship between S and r, relatedness can be estimated for a group of individuals with known $\bar{S}$.

Results

Multiple profiles for the same individual obtained using DNA extracted independently from separate leaf samples always gave identical banding patterns (data not shown). The molecular-weight sizing error of bands within and between gels was estimated directly using *Limonium* DNA samples. Table 1 shows that the overall mean of deviations from individual band-size means was larger between gels than within gels. The upper bound of the 99% confidence interval around the between-gel error estimate provides the size of error bins. Using a bin of ± 0.95% (1.90% total) of molecular weight gives 99% confidence that bands which overlap within a bin have the same molecular weight and their size differences are due to measurement error. This bin size was used for all subsequent band-sharing estimates.

Fingerprint profiles from two parental plants, a group of selfed progeny, and two presumed outcrossed progeny are shown in Fig. 1. The six selfed progeny (solid dia-

Table 1 Estimates of band molecular-weight sizing errors within and between gels. Within-gel errors are based on two scorings for each profile, using the closest molecular-weight standard to the left and right of the profile. Between-gel errors are based on profiles scored from separate gel runs of the same DNA sample. For all profiles, bands larger than, or equal to, 1900 base pairs were scored. Mean percent deviation is the overall mean of individual band-size differences from individual band-size means expressed as a proportion of mean band molecular weight. Band-sharing was calculated for each pair of profiles from the same DNA sample and the mean is given ($\bar{S}$). Profiles scored between gels sometimes differed by the presence/absence of one or several bands, giving band-sharing coefficients of less than one for duplicate runs of the same DNA

Type	Pairs of profiles	Total bands	Mean % deviation ± standard deviation	99% Confidence interval (% MWt)	$\bar{S}$ for pairs of profiles
Within gels	10	189	0.44 ± 0.003	0.383 – 0.497	1.000
Between gels	10	177	0.76 ± 0.964	0.575 – 0.947	0.957

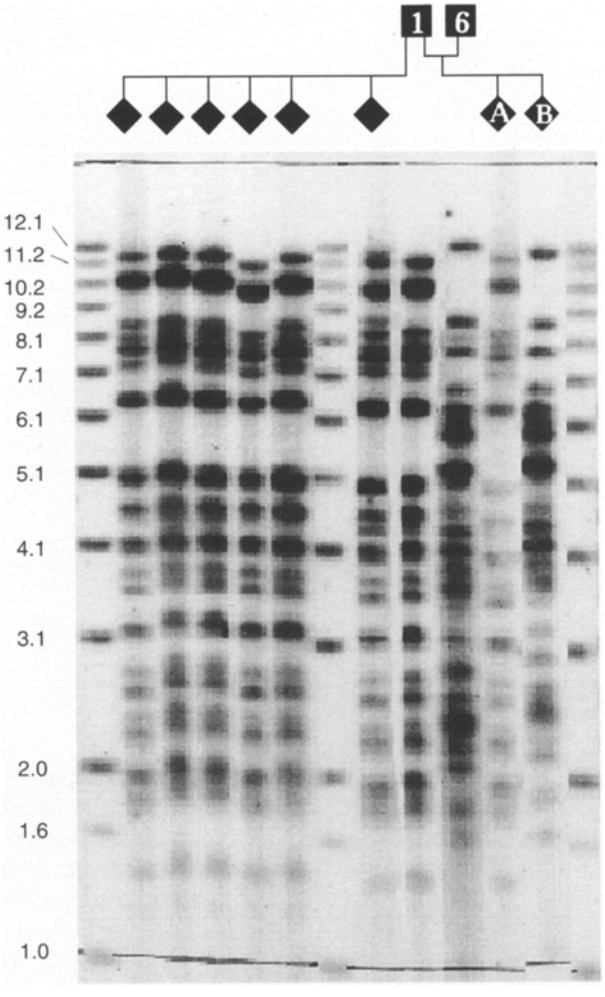

Fig. 1 DNA profiles complied from several autoradiograms and exposures that show band patterns for selfed and presumed outcrossed progeny. The *squares* represent parental plants and the *diamonds* F_1 progeny plants. The *solid diamonds* are selfed progeny of parent 1. *Diamond A* is the progeny from a presumed cross between plant 1 (dam) and plant 6 (sire) and *diamond B* is the reciprocal cross. DNA profiles of selfed progeny are very similar to the parental plant. Profiles of each of the crosses closely resemble the maternal parent and strongly suggest that the crosses were actually self pollinations. *Unlabeled lanes* are molecular-weight size markers whose size is given in kb on the left. *Lines* on the autoradiogram defined the top and bottom of profiles

Table 2 Mean band-sharing coefficients ± one standard deviation for progeny groups derived from experimental crosses. For outcrosses, the maternal individual is given first. Selfed progeny are expected to be at least half related (r=0.5) on average when parent plants are not inbred. DNA profiles strongly suggest full-sibs are actually selfed progeny

Type	Relatedness class	N	$\bar{S}$ ± SD
Plant 1	Selfs	11	0.412 ± 0.029
Plant 6	Selfs	8	0.471 ± 0.046
Plant 1× Plant 2	Full-sibs	11	0.381 ± 0.010
Plant 1× Plant 4	Full-sibs	8	0.564 ± 0.008
Plant 1× Plant 6	Full-sibs	3	0.510 ± 0.086[a]
All families			0.468 ± 0.074

[a] Parametric standard deviation; at least four data points are required for non-independence-adjusted variance estimates

monds) show a high degree of resemblance to their parent (square 1). The presumed outcrossed progeny (diamonds A and B) were derived from reciprocal crosses of parent 1 and parent 6; parent 1 was the dam of progeny A and parent 6 was the dam of progeny B. These two progeny in Fig. 1 are indicative of all presumed outcrossed progeny in that they appear highly similar to their maternal parent but have no apparent resemblance to their presumed paternal parent. Because of this high degree of resemblance to the maternal parent, as seen in selfed progeny arrays, and the fact that crosses were made without emasculation, it appears that most, or all, presumed outcross progeny are actually selfed progeny. Mean band-sharing coefficients for the controlled pollination progeny are shown in Table 2. The mean band-sharing coefficient from all families is plotted in Fig. 2 assuming the parents of these progeny are 17% inbred (hence expected r=0.585 for selfs, Falconer 1989, p. 93).

The expected relationship between band-sharing, S, and relatedness, r, for *Limonium* is shown in Fig. 2. Band-sharing among unrelated individuals, $\bar{\theta}$ of equation 3, was estimated at 0.198±0.011 from nine plants (five experimental cross parents and four other plants) sampled from several areas within the Rumstick population. Assuming that there is no inbreeding in this population, these individuals

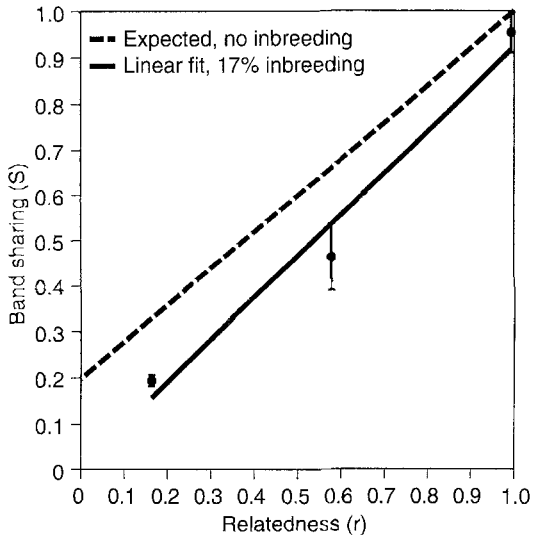

Fig. 2 A plot of band-sharing coefficient (S) by relatedness (r). The plot shows the expected relationship between S and r (see equation 3 in the text) estimated from nine presumably unrelated plants ($\bar{\theta}$=0.198±0.011) assuming these plants are completely outbred (*dashed line*). The *solid line* is a linear regression fit to the band-sharing values of unrelated individuals, the overall mean of experimental crosses (see Table 2), and the maximum band-sharing allowable given measurement error (see Table 1). The relatedness values for the regression assume that the population is 17% inbred on average. The regression equation is r=0.006+0.914*S, R^2=0.973

are unrelated and $\bar{\theta}$ estimates S at r=0. However, since *Limonium* is self compatible, wild progeny show evidence of selfing (see below), and the Rumstick population does not contain more than several thousand individuals, it is likely that both uni- and bi-parental inbreeding occur. While no independent data to estimate the level of inbreeding are available, it is still possible to use the mean band-sharing estimates from the unrelated and selfed progeny to obtain a linear approximation to the relationship between S and r. One approach is to determine if there an analytical solution to the simultaneous equations for the known quantity of S that will solve for r [using equation 3 and the equation for inbreeding of selfs as relatedness increases, F_t=0.5(1 + F_{t-1}); see Falconer 1989]. There is no value of r that will solve these simultaneous equations, suggesting that the relationship between r and S may not be linear. Another approach is to estimate the relationship between r and S with linear regression based on the band-sharing values of the unrelated plants, the selfed progeny and the maximum value of band-sharing when relatedness is unity while assuming relatedness values. Using a value of r=0.17 produces a regression line, r=0.006+0.914S, that gives a y-intercept (relatedness) very close to zero when band-sharing is zero. This value of r is an approximate maximum amount of inbreeding because larger values give regression lines with negative y-intercepts, an impossibility since individuals cannot share less than none of their bands. This line and the three points it is based on are shown in Fig. 2.

The DNA profiles of one Haffenreffer dam (591) and the nine wild-collected progeny that gave scorable profiles are shown in Fig. 3. Variable bands in these progeny occur mostly in the larger molecular-weight ranges. From these profiles it is apparent that some progeny show banding patterns highly similar to their dams (progeny 3, 8, 6 and 1), as do the selfed progeny from hand pollinations (Fig. 2). There are also a number of progeny that show bands not present in the dam's profile (progeny 10) or else lack bands that are present in the dam (progeny 4, 9, 2 and 5). Although selfed progeny do exhibit a small number of band differences from their dam, this latter group of wild progeny appear less similar to their dam than was characteristic of selfed progeny. Table 3 gives the observed $\bar{S}$ and estimated relatedness for four maternal families of wild collected seeds. The overall mean of $\bar{S}$ estimated within each wild-collected seed family was 0.392. Expected relatedness, estimated from the regression equation fit between r and S for the wild-pollinated progeny, ranged between 0.323 and 0.363.

Table 3 Mean band-sharing coefficients ± one standard deviation for progeny groups derived from greenhouse-grown, wild collected seeds from permanently marked plants in the Haffenreffer population. These four maternal families had the largest number of living progeny at the time of DNA extraction. Relatedness (r) was inferred using a linear regression equation estimated from unrelated individuals and selfed individuals. The range for inferred relatedness was based on the 95% confidence interval for $\bar{S}$

Maternal plant	N	$\bar{S}$ ± SD	Inferred r
526	9	0.364 ± 0.021	0.311 – 0.324
591	9	0.395 ± 0.010	0.343 – 0.350
595	10	0.391 ± 0.009	0.340 – 0.345
600	6	0.417 ± 0.069	0.343 – 0.400
All families	4	0.392 ± 0.022	0.323 – 0.363

Discussion

The $(GATA)_4$ minisatellite marker found in *Limonium* is somatically stable and exhibits a high degree of variability among individuals. The direct estimate of band molecular-weight scoring errors among gels, to determine the size of fixed-bins for band-sharing analyses, is important for future work with this VNTR marker in *Limonium*. We found that the between-gel component of band-sizing error was larger and had a greater variance than the within-gel component. The size of fixed-bins estimated here is smaller than that used by some labs for forensic DNA typing with single-locus probes (Budowle et al. 1991) and is similar to the smaller fixed-bin sizes estimated by other DNA-typing labs (see Risch and Devlin 1992). This finding re-inforces the need for direct estimates of band molecular-weight sizing error for newly identified VNTR markers (see Galbraith et al. 1991). Error estimates for the application of existing markers are also highly dependent on the methodology employed by different labs. For ex-

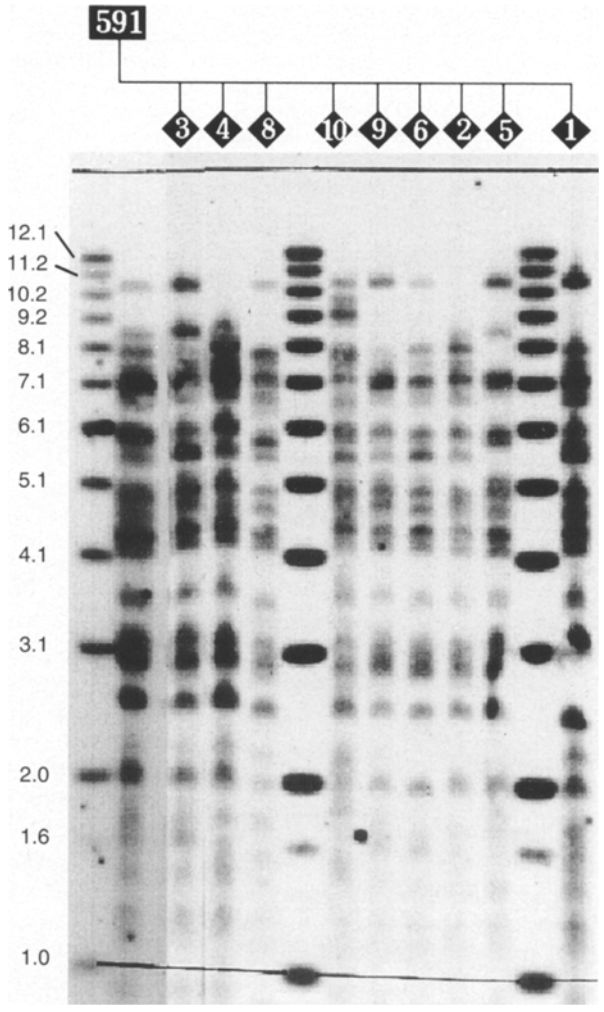

Fig. 3 DNA profiles of a *Limonium* individual (591) from the Haffenreffer population and a sample of its wild-pollinated seed progeny (diamonds). *Progeny 3, 8, 6 and 1* show banding patterns highly similar or identical to their dam while *progeny 10* shows bands not present in the dam's profile and *progeny 4, 9, 2 and 5* lack bands that are present in the dam's profile. Although selfed progeny profiles differ from their dam's profile, *progeny 10, 4, 9, 2 and 5* appear less similar to their dam than was characteristic of experimental selfed progeny and are most likely due to outcrossing. Unlabeled lanes are molecular-weight size markers whose size is given in kb on the left. The image was compiled from several exposures

ample, we found that running gels more slowly with buffer re-circulation greatly reduced variation in sample mobility across lanes of gels ("smiling") and thereby reduced the within-gel band-scoring error.

The experimental pollinations were designed to provide selfed, full- and half-sib progeny groups for an empirical validation of the expected relationship between S and r. Unfortunately, the inability to emasculate plants led to five apparently selfed progeny groups and no outcrossed progeny. Although this could suggest that *Limonium* is apomictic, as are a few species in the genus (Baker 1953), progeny do not show identical DNA profiles as expected for parthenogenesis but rather a lower $\bar{S}$ consistent with sexual selfing.

The linear regression fit between S and r for the three band-sharing data points available gives a y-intercept at approximately zero and implies that, on average, unrelated individuals share no bands. Background similarity among unrelated individuals for multilocus markers in other species ranges between 10% and 70% (see Lynch 1991). Since it is unlikely that the background similarity of *Limonium* $(GATA)_4$-profiles for unrelated plants is actually zero, the empirical relationship between S and $r$estimated with regression shows that there is not more than 17% inbreeding. This regression line is a rough estimator of the relationship between S and r because it is based on only three points. The slope of the line is also an approximate maximum because the point for highest relatedness (r=1.0) is based on re-runs of individuals and not highly inbred lines, which would be likely to exhibit more variation and a lower mean band sharing due to sampling and family effects. The amount of band-sharing among selfed experimental-cross progeny arrays is greater than that of unrelated individuals even if the linear regression fit for S and r is approximate. The number of unrelated individuals used to estimate $\bar{\theta}$ is small and may not be entirely indicative of true $\bar{\theta}$ within the Rumstick population. It is likely that these unrelated individuals are actually partly inbred on average due to the presence of a mixed mating system and a relatively small population size. It is also possible that the parents used for the experimental crosses could differ somewhat from the "unrelated" individuals in their average degree of inbreeding despite the assumption that both groups are equally inbred.

Assuming that "unrelated" plants are not inbred, the 20% background similarity in DNA profiles is in the lower range of reported estimates for animals. It seems likely that these individuals are partly inbred (no more than about 17%) and that band similarity among truly unrelated individuals would be even less. Since the expected power of relatedness estimates based on band-sharing scores will increase as similarity among unrelated individuals decreases, band-sharing among groups for this $(GATA)_4$-marker in *Limonium* should be a reasonably powerful test of relatedness. Band-sharing and relatedness estimates become more precise as the sample size used to estimate $\bar{S}$ increases. Employment of this multilocus marker to determine relatedness will necessitate the use of multiple individuals (perhaps ten or more) to estimate $\bar{S}$ within groups in order to be able to distinguish small differences in $\bar{S}$ and r among groups. Even with these sample sizes there is a considerable range of mean band-sharing values within progeny groups that should have approximately equal average relatedness, as demonstrated by progeny groups from experimental crosses (Table 2). This VNTR marker may have considerable variation among families for the relationship between r and S. This possibility needs further investigation as it would considerably increase the variance in inferred relatedness for a VNTR marker used across a wide range of genotypes.

The use of individuals with a uniform and known amount of past inbreeding as parents for progeny arrays (such as fully inbred lines and F_1s from crosses among such

lines) would provide more accurate estimates of the empirical relationship between S and r based on regression. Pre-existing pedigrees can also be used to determine the empirical relationship between relatedness and band-sharing and this method has demonstrated a significant fit in chickens (Kuhnlein et al. 1990), cattle (Mannen et al. 1993) and geese (Rave et al. 1994). Although the use of extensive pedigrees is desirable and may be possible with other plants, especially annuals or domesticated species, it was not realized in *Limonium* (despite the experimental design) due to difficulty in producing outcrossed individuals and the long growth time before first flowering (at least 18 months under greenhouse conditions).

The DNA profiles obtained for wild-bred progeny from the Haffenreffer population show a degree of band-sharing between unrelated individuals and experimental selfed progeny. The banding patterns of these wild progeny give evidence for both self- and outcross-fertilization occurring in natural populations, although banding patterns consistent with selfing are more common, suggesting that *Limonium* has a mixed mating system. These wild-bred progeny are from a different population than the hand-pollinated progeny. $\overline{S}$ from a large sample of plants from each population shows that the Haffenreffer population has a slightly higher $\overline{S}$ than the Rumstick population (0.262 and 0.232, respectively; Hamilton, unpublished). Progeny from the Haffenreffer population should therefore be somewhat more inbred than progeny from the Rumstick population. This means that there may be less current inbreeding in the mating system than the inferred relatedness estimate suggests because part of this relatedness is due to population inbreeding. That $\overline{S}$ of progeny within families containing selfed and outcrossed families is somewhat less than $\overline{S}$ among selfed-sibs is further evidence that $\overline{S}$ does actually co-vary with the degree of relatedness.

These results show that VNTR genetic markers in plants are capable of resolving relatedness differences within wild species with several potential complications. There may be considerable variation in band-sharing among families that have similar mean relatedness, which would complicate efforts to relate band-sharing and relatedness at the level of populations. Determining the relationship between r and S may also be complicated by differing levels of inbreeding among the population units where measurement takes place, in the extreme necessitating a calibration of r versus S in each population or family. Use of VNTR markers, such as the one identified here for detecting differences within species, will require background work that establishes the relationship between band-sharing and identity by descent before the marker can be used to infer relationships in wild populations or progeny groups. Although this is true of any genetic marker, it is particularly important given our present limited understanding of VNTR inheritance and the mechanisms that generate variability in such repeat systems (see Wolff et al. 1991; Weising et al. 1994a). Independent determinations of population inbreeding, or the use of pedigrees with variable and known levels of inbreeding, will make it much easier to obtain an accurate calibration of the relationship between r and S. Effort devoted to calibrating band-sharing with relatedness for a multilocus marker will allow investigators to reap many benefits when the markers are applied to wild populations and will considerably strengthen the inferential power of a VNTR marker. This hypervariable marker in *Limonium* provides an important tool to test hypotheses regarding the operation of evolutionary forces that depend on tests of relatedness among groups, although such measurements will be accompanied by a degree of measurement error inherent to the genetic marker.

Acknowledgments We thank A. Coleman, S. Gaines, K. Holsinger, M. Lynch, E. Perry, C. Purrington, B. Schaal and J. Schmitt for technical assistance, discussion and helpful reviews of this work. F. Jackson and A. Sanford provided invaluable assistance in the greenhouse. M. B. H. thanks Dr. and Mrs. Zeglier for bees that became bee-sticks as well as a stimulating introduction to apitherapy. This research was supported by a Rhode Island Wild Plant Society Educational grant to M. B. H., NSF grants to M. B. H., D. M. R and J. Schmitt (Dissertation Improvement DEB 9212849) and to D. M. R. (DEB 9120293), and was submitted in partial fulfillment of the requirements for a doctoral dissertation at Brown University.

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1993) Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York

Baker HG (1953) Dimorphism and monomorphism in the Plumbaginaceae. II. Pollen and stigmata in the genus *Limonium*. Ann Bot 67:433–445

Budowle B, Giusti AM, Waye JS, Baechtel FS, Fourney RM, Adams DE, Presley LA, Deadman HA, Monson KL (1991) Fixed-bin analysis for statistical evaluation of continuous distributions of allelic data from VNTR loci, for use in forensic comparisons. Am J Hum Genet 48:841–855

Butler MA, Templeton AR, Read B (1994) DNA fingerprinting in Speke's gazelle: a test for genetic distinctness, and the correlation between relatedness and similarity. Mol Ecol 3:355–361

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Falconer DS (1989) Introduction to quantitative genetics. 3rd edn. Longman, London

Gilbert DA, Packer C, Pusey AC, Stephens JC, O'Brien SJ (1991) Analytical DNA fingerprinting in lions: parentage, genetic diversity, and relatedness. J Hered 82:378–386

Galbraith DA, Boag PT, Gibbs HL, White BD (1991) Sizing bands on autoradiograms: a study of precision for scoring DNA fingerprints. Electrophoresis 12:210–220

Haig SM, Belthoff JR, Allen DH (1993) Examination of population structure in red-cockaded woodpeckers using DNA profiles. Evolution 47:185–194

Heery DM, Gannon F, Powell R (1990) A simple method for subcloning DNA fragments from gel slices. Trends Genet 6:173

Jung C, Pillen K, Frese L, Fähr S, Melchinger AE (1993) Phylogenetic relationships between cultivated and wild species of the genus *Beta* revealed by DNA "fingerprinting." Theor Appl Genet 86:449–457

Kuhnlein U, Zadworny D, Dawe Y, Fairfull RW, Gavora JS (1990) Assessment of inbreeding by DNA fingerprinting: development of a calibration curve using defined strains of chickens. Genetics 125:161–165

Luteyn JL (1990) The Plumbaginaceae in the flora of the southeastern United States. Sida 14:169–178

Lynch M (1988) Estimation of relatedness by DNA fingerprinting. Mol Biol Evol 5:584–599

Lynch M (1991) Analysis of population genetic structure by DNA fingerprinting. In: Burke T, Dolf G, Jefferies AJ, Wolff R (eds)

DNA fingerprinting: approaches and applications. Birkhauser Verlag, Basel, Switzerland, pp 113–126
Mak Y-M, Ho K-K (1993) Use of polyethelene glycol for purification of DNA from leaf tissue of woody plants. BioTechniques 14: 735–736
Mannen H, Tsuji S, Mukai F, Goto N, Ohtagaki S (1993) Genetic similarity using DNA fingerprinting in cattle to determine relationship coefficient. J Hered 84:166–169
Nybom H (1991) Applications of DNA fingerprinting in plant breeding. In: Burke T, Dolf G, Jefferies AJ, Wolff R (eds) DNA fingerprinting: approaches and applications. Birkhauser Verlag, Basel, Switzerland, pp 294–311
Nybom H, Hall HK (1991) Minisatellite DNA 'fingerprints' can distinguish *Rubis* cultivars and estimate their degree of relatedness. Euphytica 53:107–114
Packer C, Gilbert DA, Pusey AE, O'Brien SJ (1991) A molecular genetic analysis of kinship and cooperation in African lions. Nature 351:562–565
Piper WH, Rabenold PP (1992) Use of fragment-sharing estimates from DNA fingerprinting to determine relatedness in a tropical wren. Mol Ecol 1:69–78
Rave EH, Fleischer RC, Duvall F, Black JM (1994) Genetic analyses throught DNA fingerprinting of captive populations of Hawaiian geese. Conservation Biol 8:744–751
Risch NJ, Devlin B (1992) On the probability of matching DNA fingerprints. Science 255:717–720
Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
Van Heusden AW, Van der Voort JR, Bachmann K (1991) Oligo-(GATA) fingerprints identify clones in asexual dandelions *Taraxacum* (Asteraceae). Fingerprint News 3:13–15
Weir BS (1992) Independence of VNTR alleles defined as fixed bins. Genetics 130:873–887
Weising K, Beyermann B, Ramser J, Kahl G (1991a) Plant DNA fingerprinting with radioactive and digoxigenated oligonucleotide probes complementary to simple repetitive DNA sequences. Electrophoresis 12: 159–169
Weising K, J Ramser, D Kaemmer, G Kahl, and JT Epplen (1991b) Oligonucleotide fingerprinting in plants and fungi. In: Burke T, Dolf G, Jefferies AJ, Wolff R (eds) DNA fingerprinting: approaches and applications. Birkhauser Verlag, Basel, Switzerland, pp 312–329
Weising K, Nybom H, Wolff K, Meyer W (1994a) DNA fingerprinting in plants and fungi. CRC Press, Boca Raton, Florida
Weising K, Ramser J, Kaemmer D, Kahl G (1994b) Multilocus DNA fingerprinting and genetic relatedness in plants: a case study with banana and tomato. In: Schierwater B, Streit B, Wagner GP, DeSalle R (eds) Molecular ecology and evolution: approaches and applications. Birkhauser Verlag, Basel, Switzerland, pp 45–59
Westneat DF, Noon WA, Reeve HK, Aquadro CF (1988) Improved hybridization conditions for DNA 'fingerprints' probed with M13. Nucleic Acids Res 16:4161
Wolff R, Nakamura Y, Odelberg S, Shiang R, White R (1991) Generation of variability at VNTR loci in human DNA. In: Burke T, Dolf G, Jefferies AJ, Wolff R (eds) DNA fingerprinting: approaches and applications. Birkhauser Verlag, Basel, Switzerland, pp 20–38