

Looking for progenitors: a molecular approach to finding the origins of an invasive weed

J.M. Bond^{1,*}, E.M. Veenendaal², D.D. Hornby¹ & A.J. Gray¹

¹*Centre for Ecology and Hydrology Dorset, Winfrith Technology Centre, Dorset, UK;*

²*Harry Oppenheimer Okavango Centre, University of Botswana, Maun, Botswana;*

**Author for correspondence (e-mail: jomb@ceh.ac.uk; fax: +44-1305-213600)*

Key words: *Cenchrus biflorus*, genetic diversity, ISSR, Kalahari, origin of invasive populations

Abstract

One of the major problems in determining the origin of invasive species is that often their arrival is unnoticed. Only when population levels increase is their presence noticed but by this time determining the point of arrival is confounded by the extensive spread of the species. Here we use molecular markers (ISSRs) to determine the origin of an invasive weed in the Kalahari region, a species for which several origins could be possible. We show that molecular markers can be useful tools in determining the origin of invasive species

Abbreviations: ISSR – inter simple sequence repeat; RAPD – random amplified polymorphic DNA; AFLP – amplified fragment length polymorphisms; PCR – polymerase chain reaction; dNTPs – deoxynucleoside 5' triphosphates

Introduction

Human-mediated transport of plants and animals around the globe have increased dramatically; consequently the arrival of new species is increasingly commonplace. It is generally accepted that with the expanding global trade network the introduction of new species will tend to increase (Everett 2000). Some introduced species are barely noticed, others are tolerated because of a beneficial feature, and others out-compete native species become pests and are classed under the term 'invasive'. Unsurprisingly it is species in the last category that attract the most attention as invasions can affect ecological, economic, and social systems. Attempts to classify what characteristics make a successful invader are generally unreliable as exceptions frequently occur (reviewed in Lodge 1993). Consequently predicting invasive species is virtually impossible.

Species range changes have always been a natural component of ecosystem change, however, natural

range changes occur gradually and over many generations so that competing species co-evolve. The speed at which species boundaries are currently changing are not normal and typically artefacts of human-mediated movement aided by human manipulation of habitats. Although species ranges alter naturally, it is the unprecedented spread of a newly arrived species that leads to the classification of invasive. Once a species has become established, management becomes the more important issue.

Management of an invasive species generally looks at how the species can be controlled, but how the species first arrived and the prevention of new arrivals is another important aspect of management. Studying an invasive species in its natural range allows the factors that limit population spread, such as competition from other species, temperature extremes, water limitations, and predation pressures, to be established. Investigation of these parameters provides an insight into why the species is spreading unchecked in the new environment and how this may be prevented. When the

invasive species is causing large-scale economic damage a biological control agent may be considered. Studying the species in its original settings may identify a suitable candidate species for biological control. Often it is not clear how a species arrived in a new area, but establishing the country of origin may help identify the most likely route for import. Once the origin has been located, imports from this area can be screened to prevent further arrivals of the species. However, one generalisation regarding invasive species is that they often occupy a large native range (Lodge 1993), therefore, determining the origin of an invasive species is not always simple. Determining the origin of any invasive species is a vital component in instigating control methods.

In the absence of any substantial information to confirm or dispute the hypothesised routes of colonisation a genetic approach is necessary. Patterns of genetic similarity are often used to identify closely related populations or individuals (Blouin et al. 1996; Lynch 1988; Queller and Goodnight 1989). The theory behind the use of genetic markers is that populations with a shared ancestry (in this case the source population and the newly established invasive population) are genetically more similar. In the case of the arable weed *Capsella bursa-pastoris*, genetic data has been used to trace colonisation patterns and found high concordance with the idea that Spanish colonisers introduced the plant to California from Mexico (Neuffer 1996).

In order to detect genetic variation a marker is required. Allozymes are one of the easiest markers to screen, however, they are often associated with low variability and require a large amount of fresh tissue for analysis. Although microsatellite loci are the marker of choice for many investigations, identifying suitable microsatellite markers is a relatively long and costly process. In the case of invasive species the time factor is critical since spread can be rapid. However, a number of alternative genetic methods, that require no prior sequence knowledge and therefore can be immediately applied, exist. Random amplified polymorphic DNA (RAPD) is the longest established of these methods (Welsh and McClelland 1990; Williams et al. 1990). Amplified fragment length polymorphisms (AFLP) and inter simple sequence repeat (ISSR) are more recently developed and offer advantages over the RAPD method as they reveal a greater degree of polymorphism and are considered more reliable (Vos et al. 1995; Zietkiewicz et al. 1994). ISSRs are universal markers that utilise the fact that repetitive sequences (or SSRs) are a ubiquitous part of most genomes. The ISSR

method amplifies sections of DNA between adjacent repeats, with variation being revealed as changes in product length. Methodologically ISSR markers are simpler than AFLPs, involving only one round of amplification with polymerase chain reaction (PCR), therefore ISSRs are potentially less error prone.

In order to test the utility of molecular markers in establishing the point of origin for invasive species we investigated the most plausible colonisation route for the invasive grass *Cenchrus biflorus* which is spreading throughout the Kalahari region of Botswana. *Cenchrus* L. is a genus of tropical and subtropical grasses, characterised by a highly evolved dispersal mechanism in the form of spines and bristles on the seeds. A few species (*C. ciliaris* and *C. setigerus*) are utilised as cattle fodder, however, the species we investigate here (*C. biflorus*) is only marginally palatable to cattle (Vij and Chaudhary 1981). In addition to its relatively unpalatable nature, *C. biflorus* is regarded as a pest because the seeds injure cattle and cause wounds through which infection can be introduced. *C. biflorus* is an annual species, germinating on disturbed land, found across India, north and east Africa, Arabia and Madagascar (DeLisle 1963). Studies of *C. biflorus* in other countries are limited because the plant is not regarded as a problem.

In northern Botswana, *C. biflorus* is found in natural grasslands but is usually restricted to the spoil heaps of rodent burrows (Veenendaal et al. 2000). However, in the central Kalahari the distribution of *C. biflorus* is somewhat different, with dense aggregations (which suppress the growth of native species) growing in the disturbed ground around boreholes, roads, and fields. Botswana has undergone considerable economic growth in cattle ranching, with conservative estimates suggesting a long-term average of 2–3 cattle for every person (Arntzen and Veenendaal 1986). Areas previously unsuitable for cattle, such as the central Kalahari, have been developed to support cattle with boreholes being drilled to supply water (see Perkins 1996; Thomas et al. 2000). Areas around these artificial sources of water are highly disturbed and support little vegetation, other than *C. biflorus*, because of the continual trampling by cattle. This change in land use had not only changed the landscape, but in the case of *C. biflorus* led to an altered distribution pattern. Since *C. biflorus* is generally less favourable than the native flora and capable of rapid spread it has been identified as an invasive species.

There are several theories regarding how *C. biflorus* came to the Kalahari region, two seem equally probable

thus an ideal situation to test the utility of genetic markers in determining the origin of an invasive species. The first theory is that populations in the north of Botswana, which have persisted since the early 1940s (unpublished survey data, Veenendaal pers. comm.), have spread southwards into the Kalahari. Many invasive species show an 'initial establishment' or 'lag' phase during which relatively little range expansion occurs (Williamson 1996, p. 99), and further examples in Shigesda and Kawasaki (1997). Therefore, *C. biflorus* may have spread from the northern populations and represent an example of the lag phase followed by expansion. However equally probable, because of the highly efficient dispersal mechanism, is the inadvertent introduction of seeds by human movements. As the central Kalahari has seen an increase in cattle ranching, the plant could have been introduced from another part of Africa by accidental human-mediated transport of *C. biflorus* seeds. Oral interviews with local farmers suggest that the grass may have spread into the Kalahari from the road connecting Botswana with eastern Namibia, but imported fodder from the Gauteng region of South Africa has also been blamed (van der Putten 2000). However, *C. biflorus* is not known in South Africa (R. Ellis, pers. comm.) therefore, this route of introduction is unlikely. It is also possible, although more unlikely, that the plants in the Kalahari are a novel hybrid species formed between *C. biflorus* and *C. incertus* (which has always been found in the region).

In order to determine which scenario is the more probable, a genetic approach using allozymes and ISSR markers has been utilised. In this case, close genetic relatedness between the northern and Kalahari populations would suggest that the northern populations had expanded to occupy the central Kalahari. In contrast, no genetic similarity would indicate the alternative scenario was more probable, with the search then widened to include plants from other regions of Africa. The third option is considered less probable but material from *C. incertus* will be included to refute this possibility.

Materials and methods

Allozyme analysis

Variation was screened in 11 *C. biflorus* populations with 10 enzyme systems: phospho-glucose isomerase (PGI); peroxidases (PER); glucose-6-phosphodehydrogenase (G6PHD); 6-phospho-glucose

dehydrogenase (6 PGDH); isocitrate dehydrogenase (IDH); glucose dehydrogenase (GDH); esterase (EST); shikimic dehydrogenase (SKDH); glucose oxaloacetic transaminase (GOT); superoxide dismutase (SOD). These 10 allozymes failed to detect any variation in 11 *C. biflorus* populations collected across Botswana, therefore, an allozyme study was inappropriate in this case.

DNA analysis, plant material, DNA extraction, and PCR

C. biflorus seeds from various regions of Botswana (Figure 1) and two populations of *C. incertus* were collected. Samples from the north of the Botswana represent established populations, Kalahari populations are more recent and considered invasive. Populations are readily delimited as seeds germinate predominantly on areas of bare ground. Seeds were posted to the UK and germinated in an unheated glasshouse. In order to avoid sampling closely related individuals, seeds from different flowering spikes were germinated separately, and one seedling from each spike selected for analysis. The number of samples screened from each population is shown in Table 1. A leaf was removed, freeze dried and stored at -20°C until required for DNA extraction, which was carried out using a Nucleon Phytopure DNA extraction kit (Amersham).

PCR reactions were carried out in 10 μl reaction volumes containing 1.5 mM MgCl_2 , 0.2 mM deoxynucleoside 5' triphosphates (dNTPs), 0.2 μM primer, 1% formamide and $\sim 0.01 \mu\text{g}$ DNA template. The thermal cycling conditions consisted of 30 cycles of 1 min at 94°C , 2 min at 55°C and 40 s at 72°C , followed by a final extension of 5 min at 72°C . Every PCR reaction was duplicated to allow the reliability of the ISSR markers to be assessed. Banding patterns were visualised by running the samples on 6% acrylamide gels and silver staining using Promaga's silver staining reagents. We screened for variation in eight ISSR primers selected from University of British Columbia's Primer set #9, (811, 812, 834, 841, 848, 856, 857, 888, and 981). Of the primers selected, 888 and 891 produced the greatest number of clearly resolvable bands and were used to screen all samples.

Statistical analyses

The Shannon diversity index was used to assess the genetic diversity because this index is sensitive to small

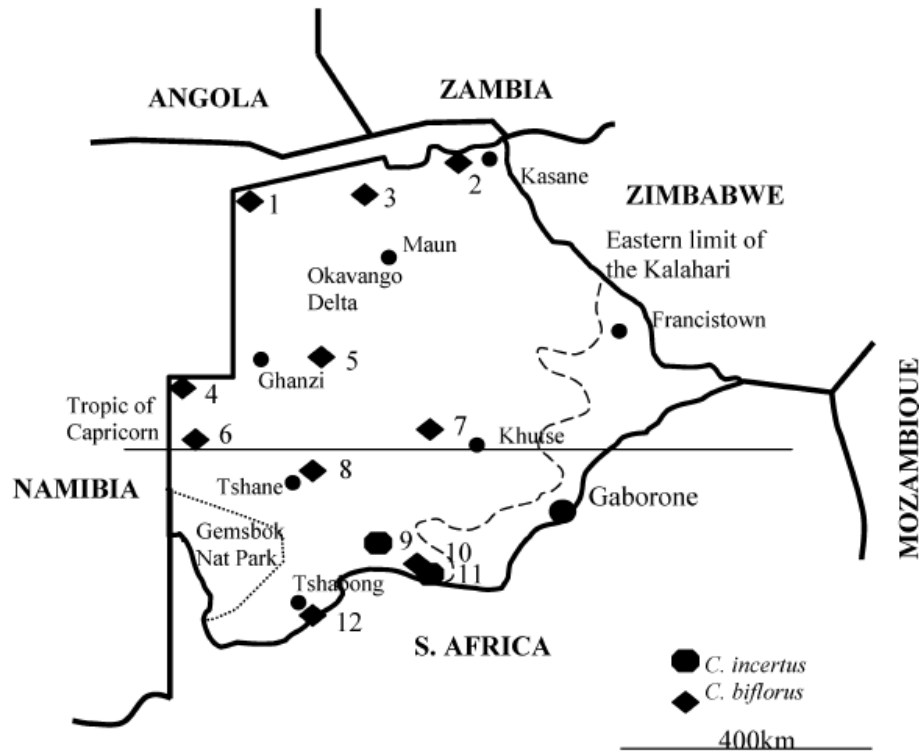


Figure 1. Sampling sites within Botswana. Populations are represented by a three letter code: 1, SHO Shongomoshu; 2, SAK Satau kgotla; STR road to Satau and PAR Parakarungu kgotla; 3, BET Betsha; 4, MAK Makunda; 5, FLF Fani Lewis farm; 6, KKB Karakubis; 7, XAD Xade; 8, KAB Kang Brigades; AIC Kaninyane artificial insemination camp and KRA Kang regional agricultural offices; 9, ITS Tswaitswai borehole (*C. incertus* population); 10, WER Werda; 11, ITE Werda (*C. incertus* population); 12, TSA Tsabong. Populations collected at site 2 represent the populations established in the 1940s. The Kalahari sites are more recently established. Samples from Burkina Faso are not shown on this map because of the physical separation involved.

changes in band abundance (Mouillot and Lepêtre 1999). Diversity was calculated within and between each population as follows (Lewontin 1972):

$$H_j = -\sum p_i (\log_2(p_i))$$

where p_i is the frequency of the band in the population. The species diversity (H_{sp}) was calculated in the same way, but using the frequency of a band in the whole sample. The average diversity over all populations (H_{pop}) was calculated for each primer as $(1/n)\sum H_j$, where n is the number of populations. From these values the proportion of diversity within each population is calculated as H_{pop}/H_{sp} and the proportion of diversity between populations is $(H_{sp} - H_{pop})/H_{sp}$.

The banding patterns were also used to calculate genetic distances between populations. The absence of a band is not necessarily an indication of similarity, since a band may be absent in two samples for

different reasons. Therefore, it is important to use a distance measure that only takes into consideration the sharing of bands, such as Nei's genetic distance: $D = 1 - (2N_{xy}/N_x + N_y)$ (Nei and Li 1979); where N_{xy} is the number of fragments shared between samples x and y ; N_x is number of fragments in sample x and N_y the number of fragments in sample y . Nei's genetic distance was calculated between all pairs of individuals within the TREECON program (Van de Peer 1994).

Genetic distances were displayed as a geometric picture using non-metric multi-dimensional scaling (MDS) techniques using the PRIMER package (Carr 1996). The MDS technique takes a matrix (similarity or dissimilarity) and plots all the samples in the configuration that best satisfies all the constraints of the matrix. Samples with the greatest similarity will be plotted closest together; with the axis of MDS plot

Table 1. Diversity indices between and within populations of *C. biflorus*.

| Population | <i>n</i> | 888 | 891 | Average |
|---|----------|-------|-------|---------|
| Burkina Faso – H_j | 5 | 8.93 | 5.55 | 7.24 |
| Mozambique – H_j | 15 | 17.81 | 17.41 | 17.61 |
| Botswana, southern border (site TSA) – H_j | 5 | 1.31 | 1.43 | 1.37 |
| Botswana, southern border (site WER) – H_j | 9 | 5.69 | 0.00 | 2.85 |
| Botswana, western border, central (site MAK) – H_j | 10 | 2.35 | 0.39 | 1.37 |
| Botswana, western border, central (site KKB) – H_j | 14 | 1.18 | 1.05 | 1.12 |
| Botswana, western border, north (site SHO) – H_j | 13 | 2.05 | 0.11 | 1.08 |
| Botswana, central Kalahari (site FLF) – H_j | 8 | 5.11 | 2.72 | 3.92 |
| Botswana, central Kalahari (site KAB) – H_j | 12 | 3.50 | 0.50 | 2.00 |
| Botswana, central Kalahari (site XAD) – H_j | 13 | 2.05 | 3.00 | 2.53 |
| Botswana, central Kalahari (site AIC) – H_j | 13 | 1.57 | 1.30 | 1.44 |
| Botswana, central Kalahari (site KRA) – H_j | 11 | 1.12 | 1.41 | 1.27 |
| Botswana, northern border (site SAK) – H_j | 14 | 3.71 | 5.05 | 4.38 |
| Botswana, northern border (site BET) – H_j | 12 | 8.61 | 5.25 | 6.93 |
| Botswana, northern border (site STR) – H_j | 11 | 0.84 | 2.24 | 1.54 |
| Botswana, northern border (site PAR) – H_j | 9 | 0.87 | 2.22 | 1.55 |
| <i>C. incertus</i> , southern Botswana – H_j | 25 | 0.00 | 0.00 | 0.00 |
| H_{sp} (excluding <i>C. incertus</i>) | | 15.41 | 11.12 | 13.27 |
| H_{pop} | | 4.14 | 3.10 | 3.64 |
| % diversity within populations (H_{pop}/H_{sp}) | | 0.27 | 0.28 | 0.27 |
| % diversity between populations ($(H_{sp} - H_{pop})/H_{sp}$) | | 0.73 | 0.72 | 0.73 |

representing arbitrary units rather than actual genetic distance.

Non-metric MDS places n points representing the n samples in a two-dimensional ordination plot with population i at co-ordinates X_{1i} on ordination axis 1 and at X_{2i} on axis 2. The distance A_{ij} between populations i and j in the MDS ordination plot is therefore: $A_{ij} = \sqrt{[(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2]}$. The MDS optimisation procedure places the n points in a two-dimensional plot so as to maximise the agreement between the ranks of actual genetic distances (D) and the ranks of MDS plot distances (A_{ij}). The lack of agreement is measured by a statistic called the STRESS (short for Standardised Residual Sum of Squares) defined as:

$$\text{STRESS} = \sqrt{\frac{\sum_{i < j} (D_{ij} - A_{ij})^2}{\sum_{i < j} D_{ij}^2}}$$

MDS and STRESS are well described in Krzanowski (Krzanowski 1987). STRESS values <0.1 correspond to an excellent representation of the data with little chance of a misleading diagram. STRESS values <0.2 still provide a potentially useful picture, although for the higher values the detail of the plot cannot be considered truly accurate. Plots with a STRESS value >0.3 should be treated with scepticism especially when less than 50 points have been plotted. The advantage of non-metric MDS over metric MDS (which tries to maximise

ordination distances to the actual distances, D_{ij} , rather than their rank) is that ordination plots from the former are less dominated by odd outlier populations. Hence, non-metric plots often contain a better representation of inter-population distances within two-dimensional plots.

Results

The ISSR markers revealed relatively high levels of genetic variation; in total 152 bands were scored (74 from primer 891 and 78 from primer 888). In *C. biflorus* levels of polymorphism were high, with 90% of bands being polymorphic at primer 891 and 78% in primer 888. Inconsistencies in PCR were not revealed between the duplicate PCR reactions. Despite this reliability, faint bands were not scored as they are more likely to be poorly amplified or inconsistently stained. Some bands were only found in one population, these are termed private alleles, and are unique to that population. Of Botswana samples, none of the Kalahari populations contained private alleles, whereas the northern population of SAK contained two unique bands and BET one unique band. Diversity indices as calculated for each primer and population are shown in Table 1. In the *C. incertus* plants screened, there was no variation in banding pattern, either within or between sites. In comparison, the *C. biflorus* populations were

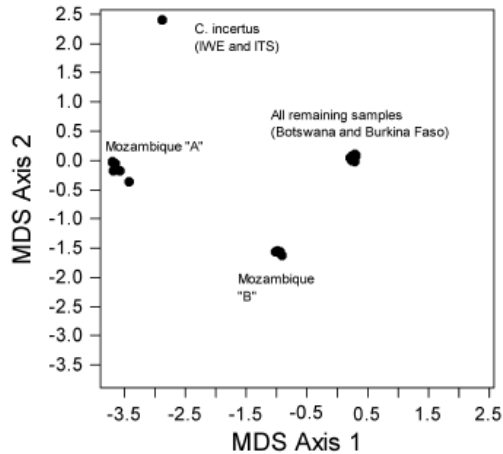


Figure 2. MDS plot showing all samples.

exceptionally variable. The more variable the banding pattern within a population the higher the diversity value; the samples obtained from Burkina Faso, Mozambique, and northern Botswana sites BET and SAK were highly variable. In contrast, the remaining populations from Botswana showed lower levels of genetic diversity.

Figure 2 shows all the *Cenchrus* spp. samples plotted on a MDS plot. The STRESS value for Figure 2 is exceptionally low (0.03) for such a large number of samples ($n = 199$). Therefore, the distance plotted between each population pair reflects the actual genetic distance without extreme distortion. For greater resolution, the main cluster (Burkina Faso/Botswana) was plotted on a second MDS plot in Figure 3. Figure 3 shows that within the Burkina Faso/Botswana cluster, each population clusters together but the spread of data points for each population varies. Although the stress value for this plot is higher than for the previous plot (STRESS = 0.12) the data points still provide a good indication of the genetic distances between all individuals.

Discussion

Any attempt to examine genetic similarity between natural populations must take care with the sampling strategy. If closely related individuals are inadvertently screened the sample will be biased and the level of genetic variation revealed may be lower than in the population as a whole. For this reason we were careful not

to sample adult plants because they could, by chance, represent the seed from one parent plant. Instead seeds from individual plants were sampled and germinated in the greenhouse, with care taken not to mix the seeds from different plants. One seed from each parent plant was analysed, the disparity in sample sizes resulting from none of the seeds from the parent plant germinating.

The diversity revealed in each population varied tremendously, the populations from Mozambique were the most diverse whilst no variation was shown in the 25 *C. incertus* plants screened from two distant populations. However, the high diversity in the Mozambique sample was an artefact of ascribing the Mozambique samples into one group when they are genetically as differentiated as *C. incertus* and *C. biflorus*. Species diversity (H_{sp}) was high, but this value reflects the large difference between the Mozambique and Botswana samples. In contrast the population diversity (H_{pop}) was small, but this reflects the fact that the majority of populations from Botswana contained little diversity. The portioning of diversity within and between populations revealed that very little diversity (27%) was found within a population, despite our efforts to maximise the variation revealed (by avoiding sampling families). Very little diversity will be seen within populations when the species is inbreeding (Bussell 1999) therefore, it appears that *C. biflorus* is predominantly, but not restricted to, inbreeding.

The MDS plot (Figure 2) shows that *C. incertus* is clearly differentiated from *C. biflorus*. A hybrid would occupy a position intermediate to the two species. Therefore, from the genetic data we can immediately discount the possibility that the Kalahari plants are a novel hybrid. Genetic differentiation of the same magnitude is seen between the Mozambique plants and the remaining *C. biflorus* samples. Within the Mozambique samples, two distinctly different banding patterns were revealed both of which are highly distinct from the Botswana samples and from each other. The genetic distance between the two Mozambique groups (arbitrarily termed 'A' and 'B') was as extreme as that between *C. incertus* and *C. biflorus*. In terms of a possible origin of *C. biflorus* in the central Kalahari, it is clear that Mozambique plants, or seeds, do not represent the source population.

In order to establish if the northern populations could be the progenitors of the Kalahari populations the MDS plots were redrawn removing the Mozambique and *C. incertus* samples (Figure 3). The three northern populations (SAK, STR, and PAR), which have

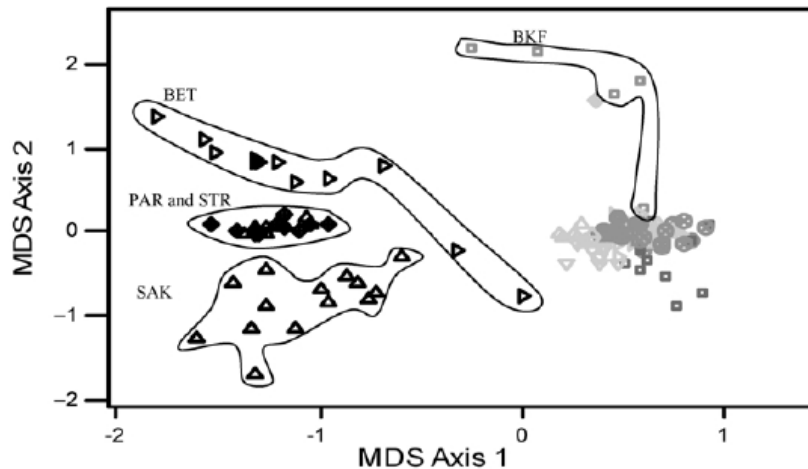


Figure 3. MDS plot showing Botswana and Burkina Faso samples. Each point represents an individual plant with tones and shapes distinguishing each population. Black diamonds or triangles represent the northern populations of BET, PAR, STR, and SAK. The tight clustering of the Kalahari samples makes distinguishing populations virtually impossible. However, the lightest grey represents the central Kalahari samples of FLF, KAB, XAD AIC, and KRA. The mid-grey circles indicate the western Kalahari samples (MAK, KKB, and SHO). Whilst the darker grey squares represent the southern Kalahari populations (TSA and WER). The samples from Burkina Faso are shown as mid-grey squares and labelled by the code BKF.

perpetuated since the 1940s, are clearly genetically differentiated from the remaining Botswana samples. The other northern population (BET) is distinct from the Kalahari samples and similar to the established northern populations. Therefore, we suggest that population BET originated from the established northern populations, unfortunately we have no data to indicate if this population is also long standing. The SAK and BET populations form discrete clusters, each cluster contains a wide spread of data points and consequently relatively high diversity (H_j) values. In direct contrast the PAR and STR populations are genetically indistinguishable and contain comparatively little variation.

The northern populations are clearly genetically differentiated from the Kalahari samples, with the differences not just based on changes in band frequency. The northern samples (BET, PAR, STR, and SAK) contained six bands that were absent in the Kalahari samples whilst the Kalahari samples contained eight bands that were not present in the northern populations. Gene flow from *C. biflorus* populations in the south of Angola, Zambia, and Zimbabwe may provide an input of genetic material which causes the northern populations to be genetically diverse and differentiated from the Kalahari samples. The genetic screening has shown that it is unlikely that the northern populations

have spread to the Kalahari region. Although many invading species show a characteristic lag phase prior to range expansion, it does not appear that the northern population has expanded after a lag period.

The levels of diversity within the Kalahari populations can be used to reveal information about their origin. Diversity can be introduced into a population in one of two ways, either through gene flow (i.e. interbreeding) with a genetically differentiated population, or through mutation. Given that genetic markers mutate slowly, and that the Kalahari populations have only been established for a maximum of 20 years, it is unlikely that significant genetic variation has been caused by mutation. All of the Kalahari *C. biflorus* populations screened were genetically variable, with more than one genotype within each population. Therefore, we conclude that the Kalahari populations are not derived from a single introduction event. We suggest that these newly formed populations were established through several introductions of genetically similar material, either from one population, or from a geographic region containing genetically similar populations. Although the northern population BET shows some similarity with the Kalahari populations, it is unlikely, given the extent of variation within BET, that populations derived by multiple introductions from BET would show so little variability. If the northern

populations were the source of progenitors for the central Kalahari populations, the MDS plot would indicate a greater degree of genetic similarity.

Samples from Burkina Faso are highly differentiated from each other. Had more samples been screened an even greater spread of genetic diversity may have been revealed. Despite the large spread of points, one of the Burkina Faso samples is remarkably similar to those from the main Botswana cluster. Additionally, an outlying central Kalahari sample clusters closely with the Burkina Faso samples. This genetic similarity is surprising considering the low number of samples collected from Burkina Faso, and the geographical separation.

It is clear that genetic interchange with the northern populations is limited; otherwise a greater degree of genetic similarity would be expected. The seeds of *C. biflorus* have specialised spines that attach to virtually any surface. Seeds are spread by attaching to animals and birds, but spread is increased through human-mediated movement, such as attaching to tyres, or within fodder for cattle. In a survey of farmers within the Kalahari region, all considered cattle as the main element in the spread of the grass. However, half of the respondents also identified humans as being instrumental in transporting seeds, usually in connection with cattle or vehicles (Van der Putten 2000). In support of these observations many *C. biflorus* populations are clustered around centres where cattle are gathered (e.g. agricultural stations, show grounds, and artificial insemination camps). We would expect populations at sites where cattle gather to contain higher levels of genetic variation, simply because seeds from many different regions are brought in by the cattle. However, the genetic diversity within the populations at the artificial insemination camp (AIC) and agricultural offices (KRA) do not contain raised levels of genetic variation. From this observation we conclude that our sampling regime has detected a large proportion of the genetic diversity within the Kalahari populations.

Despite the huge potential for spread, it would appear that human and animal movement does not spread the northern populations either towards the south or west. Geographically, the western population SHO is closer to the northern population of BET than the other central Kalahari populations. Despite this relatively close geographic association, population SHO remains genetically distinct from the northern group of samples. The north east of Botswana is culturally

separated from the rest of the country, in addition the northern cattle grazing areas are isolated by large game fences and poor road infrastructure. These cultural factors may explain why movement of *C. biflorus* across the Okavango Delta southward is comparatively rare.

We suggest that the Kalahari plants have been introduced from outside Botswana. One local name for the plant (Makunda) suggests that the plant first appeared around Makunda, a village close to the Namibian border. However, samples from Namibia were unavailable for screening. Of the samples screened here, remarkable similarity is found between some Burkina Faso and Kalahari samples. However, because of the geographic separation, Burkina Faso is unlikely to be the source of the newly established populations (unless there is substantial movements of fodder or cattle between these two areas). Since diversity is low in the introduced populations, but not invariable, a number of introductions are probably responsible for the establishment of *C. biflorus*. The progenitor plants, or seeds, are thought to have come from a genetically restricted source, which may reflect a limited geographic origin. A comprehensive genetic survey may pinpoint the possible progenitor population. We predict that human-mediated movement has introduced the seeds of *C. biflorus*, which has led to the formation of new populations in the Central Kalahari.

C. biflorus satisfies many of the generalisations made about invasive species such as r-selected traits (e.g. quick maturation, low generation time, large seed output), high dispersal capabilities, inbreeding species, lack of predators, and coloniser of disturbed ground. From these general rules it is anticipated that *C. biflorus* will continue to spread as cattle ranching in the Kalahari increases. The molecular approach we adopted has been successful in distinguishing which of two contrasting theories of introduction was more probable. The genetic data has indicated that the Kalahari populations originated from outside Botswana, screening more samples may identify the exact location. We suggest that the future spread of *C. biflorus* in Botswana will closely match human and livestock movements (as already implied by the lack of gene flow southwards from the northern populations).

References

- Arntzen J and Veenendaal E (1986) A Profile of Environment and Development in Botswana. NIR/IES, Gaborone/Amsterdam, 275 pp

- Blouin MS, Parsons M, Lacaille V and Lotz S (1996) Use of microsatellite loci to classify individuals by relatedness. *Molecular Ecology* 5: 393–401
- Bussell J (1999) The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). *Molecular Ecology* 8: 775–789
- Carr M (1996) PRIMER User Manual (Plymouth Routines in Multivariate Ecological Research), Plymouth Marine Laboratory, UK
- DeLisle D (1963) Taxonomy and distribution of the genus *Cenchrus*. *Iowa State College Journal of Science* 37: 259–351
- Everett R (2000) Patterns and pathways of biological invasions. *Trends in Ecology and Evolution* 15: 177–178
- Krzanowski W (1987) Principles of Multivariate Analysis, a User's Perspective. Clarendon Press, Oxford
- Lewontin R (1972) The apportionment of human diversity. *Evolutionary Biology* 6: 381–398
- Lodge D (1993) Biological invasions: lessons for ecology. *Trends in Ecology and Evolution* 8: 133–137
- Lynch M (1988) Estimation of relatedness by DNA fingerprinting. *Molecular Biology and Evolution* 5: 584–599
- Mouillot D and Leprêtre A (1999) A comparison of species diversity estimators. *Research in Population and Ecology* 41: 203–215
- Nei M and Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Science USA* 76: 5269–5273
- Neuffer B (1996) RAPD analyses in colonial and ancestral populations of *Capsella bursa-pastoris* (L.) Med. (*Brassicaceae*). *Biochemical Systematics and Ecology* 24: 393–403
- Perkins J (1996) Botswana: fencing out the equity issue. Cattlepost and cattle ranching in the Kalahari Desert. *Journal of Arid Environments* 33: 503–517
- Queller DC and Goodnight KF (1989) Estimating relatedness using genetic-markers. *Evolution* 43: 258–275
- Shigesda N and Kawasaki K (1997) *Biological Invasions: Theory and Practice*. Oxford University Press, Oxford
- Thomas D, Sporton D and Perkins J (2000) The environmental impact of livestock ranches in the Kalahari, Botswana: natural resource use, ecological change and human response in a dynamic dryland system. *Land Degradation and Development* 11: 327–341
- Van de Peer Y (1994) TREECON for windows: a software package for the construction and drawing of evolutionary trees for the microsoft windows environment. *Computer Applications in the Biosciences* 10: 569–570
- Van der Putten WE (2000) The impact of invasive grass species on the structure, function and sustainable use of coastal and inland sand dune ecosystems in southern Africa (INVASS), p 77. Netherlands Institute of Ecology, Heteren, The Netherlands
- Veenendaal E, Sekhute B, Ripley B and Burns A (2000) *Cenchrus biflorus*; alien or sahelien? Investigations into the invasive properties of an alien annual grass in the Kalahari. In: Masundire Chanda R (ed) *Towards Sustainable Management in the Kalahari Region – Some Essential Background and Critical Issues*, Directorate of Research and Development
- Vij S and Chaudhary J (1981) Cytological investigations in three species of *Cenchrus* L. (Gramineae). *Cytologia* 46: 661–669
- Vos P, Hogers R, Bleeker M, Jijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 18: 6531–6535
- Welsh J and McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18: 7213–7218
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA and Tingey SV (1990) DNA Polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531–6535
- Williamson M (1996) *Biological Invasions*. Chapman & Hall, London
- Zietkiewicz E, Rafalski A and Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR) anchored polymerase chain reaction amplification. *Genomics* 20: 176–183