RESEARCH ARTICLE

Estimation of *in situ* mating systems in wild sorghum (*Sorghum bicolor* (L.) Moench) in Ethiopia using SSR-based progeny array data: implications for the spread of crop genes into the wild

ASFAW ADUGNA[1,2*], PATTY M. SWEENEY[3] and ENDASHAW BEKELE[1]

[1]*Department of Biology, Addis Ababa University, P. O. Box 1976, Addis Ababa, Ethiopia*
[2]*Present address: Melkassa Agricultural Research Center, P. O. Box 1085, Adama, Ethiopia*
[3]*Department of Evolution, Ecology and Organismal Biology, Ohio State University, 318 W 12th Avenue, Columbus, OH 43210, USA*

Abstract

Because transgenic sorghum (*Sorghum bicolor* L.) is being developed for Africa, we investigated the potential for transgenes to spread to conspecific wild/weedy sorghum populations in Ethiopia, which is considered the centre of origin of cultivated sorghum. In the current study, the extent of outcrossing, and uniparental and biparental inbreeding were investigated in seven wild/weedy sorghum populations collected at elevations ranging from 631 to 1709 m. Based on allele frequency data of 1120 progenies and 140 maternal plants from five polymorphic microsatellite markers, outcrossing rates were estimated using standard procedures. The average multilocus outcrossing rate was 0.51, with a range of 0.31–0.65 among populations, and the family outcrossing rate was in the extreme range of 0 to 100%. The highest outcrossing ($t_m = 0.65$) was recorded in a weedy population that was intermixed with an improved crop variety in Abuare (Wello region). It was also observed that the inbreeding coefficient of the progenies (F_p) tend to be more than the inbreeding coefficient of both their maternal parents (F_m) and the level of inbreeding expected at equilibrium (F_{eq}), which is a characteristic of predominantly outbreeding species. Biparental inbreeding was evident in all populations and averaged 0.24 (range = 0.10–0.33). The high outcrossing rates of wild/weedy sorghum populations in Ethiopia indicate a high potential for crop genes (including transgenes) to spread within the wild pool. Therefore, effective risk management strategies may be needed if the introgression of transgenes or other crop genes from improved cultivars into wild or weedy populations is deemed to be undesirable.

[Adugna A., Sweeney P. M. and Bekele E. 2013 Estimation of *in situ* mating systems in wild sorghum (*Sorghum bicolor* (L.) Moench) in Ethiopia using SSR-based progeny array data: implications for the spread of crop genes into the wild. *J. Genet*. **92**, 3–10]

Introduction

Mating system is one of the most important aspects of the life history of a plant species that determine gene exchange (Hamrick and Godt 1990) and plays a major role in shaping plant population-genetic structure (Brown and Allard 1970; Ritland 1983; Enjalbert and David 2000). It is also a key parameter in plant evolutionary studies (Brown 1979). The dynamics of genetic diversity may be influenced by mating system, which in turn is influenced by different farming practices and environmental factors (Barnaud *et al.* 2008).

Information about mating systems and probable introgression of transgenes into wild relatives and its economic and ecological consequences are still rare despite increasing study (Schmidt and Bothma 2006). Worldwide, government regulatory agencies require information about the extent to which a transgenic crop will cross-pollinate with its wild or weedy relatives (Muraya *et al.* 2011a). Knowledge about gene flow is primarily needed to evaluate whether novel transgenes could unintentionally enter wild or weedy populations and confer fitness-enhancing traits that could exacerbate weed problems. Generally, there is interest in knowing whether transgenes are likely to spread from crop to crop, crop to weeds, and crop to wild relatives, especially where landraces and wild or weedy relatives represent valued germplasm for future crop breeding. Regardless of whether transgenic or nontransgenic cultivars are grown, information about the potential for crop-to-wild gene flow is needed to understand its effects on the genetic diversity of wild populations (Ellstrand 2003).

*For correspondence. E-mail: asfaw123@rediffmail.com.

Keywords. gene flow; inbreeding; outcrosssing; wild sorghum.

Brown and Allard (1970) first applied isozyme polymorphisms as markers for the estimation of mating system parameters in maize and later Ritland and Jain (1981) developed a computer program to analyse them. Since then the method of computing mating systems has been employed in many plant species following the advent of more polymorphic genetic markers such as microsatellites. Previous studies in crop sorghum showed outcrossing rates in the range of less than 10% to 73% (Ellstrand and Foster 1983; D'je *et al.* 2004; Barnaud *et al.* 2008) and to nearly 100% in individual sudan grass plants (Pedersen *et al.* 1998). Moreover, Muraya *et al.* (2011a) found multilocus outcrossing rates of 8.9 to 70% in wild sorghum in Kenya. However, such information is not available in Ethiopia, the centre of origin of sorghum.

Conservation of genes in the wild has become a concern, especially in the past two decades, owing to development and official release of transgenic crop plants (e.g. Gepts and Papa 2003). The concerns are for the most part associated with the trepidation that transgenic plants containing herbicide tolerance gene(s), for example, may escape to the environment and create increased weediness or even 'super weeds', which can be difficult to control. Owing to the increasing deployment of genetically modified (GM) crops in developing countries (James 2011), scientific risk assessment of transgenic crops and its impact on conventionally bred crops and crop wild relatives is needed to establish adequate biosafety regulations. Research is under way to develop GM sorghum (*S. bicolor* ssp. *bicolor* (L.)) cultivars with enhanced nutritional factors (Zhao 2007) for Africa, where sorghum originated (Mann *et al.* 1983; Doggett 1988) and is a staple food crop. Gene flow from modern cultivars to local landraces and wild and weedy relatives is expected to occur (Ejeta and Grenier 2005; Tesso *et al.* 2008). Thus, use of transgenic sorghum and accompanying increases in adoption of modern cultivars raise questions about genetic erosion and possible unwanted effects of transgenes that introgress into wild populations. Moreover, knowledge of their level of outcrossing under current field conditions is highly desirable for designing effective *in situ* conservation of wild and domesticated plants (Ritland and Jain 1981) as outcrossing can be modified by phenology, field size and planting practices (Barnaud *et al.* 2008).

Wild relatives (the races under *S. bicolor* ssp. *arundinaceum*) of cultivated sorghum (*S. bicolor* ssp. *bicolor*) often occur in close proximity to sorghum crop fields in Africa (Ejeta and Grenier 2005; Tesso *et al.* 2008). While gene flow from the wild species to the crop might lead to reduced yield and loss of the genetic purity of the cultivated varieties (Felber *et al.* 2007), gene flow in the reverse direction and introgression of crop alleles into wild germplasm could lead to the contamination of wild sorghum populations, which in turn could lead to increased weediness in the sorghum and/or erosion of genetic diversity of wild sorghum populations. Crop–weed hybridization is known to have been involved in evolution of numerous weeds (Barrett 1983; Small 1984), and weedy sorghum subspecies such as shattercane (*S. bicolor* ssp. *drmmondii*) already exist (Ejeta and Grenier 2005; Defelice 2006). The introgression of crop sorghum genes into its close species, Johnsongrass (*S. halepense* (L.) Pers.) has been implicated in the increased weediness of Johnsongrass (De Wet and Harlan 1975; Holm *et al.* 1977; Pedersen *et al.* 1998). It has been suggested that all *S. bicolor* ssp. *bicolor* races can freely hybridize with wild *S. bicolor* ssp. *verticilliflorum* to produce shattercanes (*S. bicolor* ssp. *drummondii*) (Dahlberg 2000; Defelice 2006). Identification and quantification of the risks associated with gene flow, and details of how to measure rare hybrid events and their ecological fitness costs require knowledge of the mechanisms of pollen dispersal and hybridization events. Therefore this study was carried out to investigate the extent of *in situ* outcrossing rate and inbreeding of wild/weedy sorghum populations to predict the consequences of such outcrossing and to suggest further germplasm conservation measures.

Materials and methods

Plant material

Seeds were collected in 2008 from October through November from different races of the annual wild *S. bicolor* ssp. *verticilliflorum* and the weedy *S. bicolor* ssp. *drummondii* in different regions of Ethiopia (figure 1) where the presence of wild sorghums occurring sympatrically and allopatrically with the cultivated sorghum have been reported by Ayana *et al.* (2001) and Tesso *et al.* (2008). Samples were collected from seven populations distributed in the major sorghum-growing regions of Ethiopia: Ghibe valley (N8°11′E37°33′, alt. 1425 m), Bisidimo (N9°12′E42°13′, alt. 1386 m), Mandura (N11°05′E36°25′, alt. 1404 m), Humera-Donkey (N14°14′E36°41′, alt. 631 m), Hagereselam (N14°05′E36°58′, alt. 820 m), Abuare (N12°05′E39°39′, alt. 1426 m), and Kobo (N12°08′E39°37′, alt. 1500 m). From

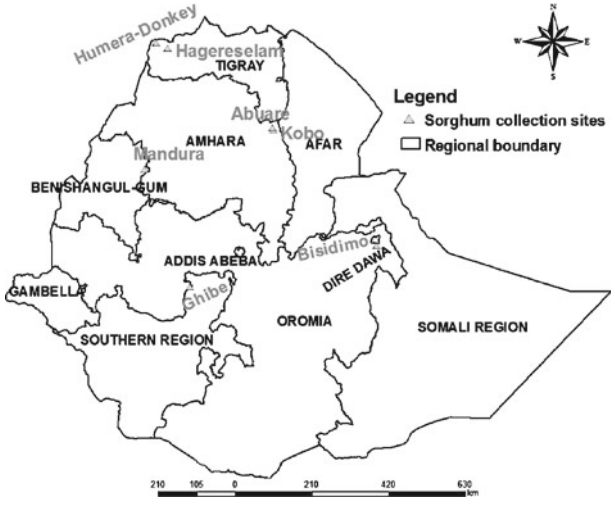

Figure 1. Map of Ethiopia showing the regions from where wild sorghum was collected.

each population, 20 widely spaced individuals were sampled. Seeds from each of the 20 maternal plants (panicles) were sown in individual pots in the greenhouse at Melkassa Agricultural Research Center in Ethiopia in January 2009. From each maternal plant eight progeny plants (hereafter referred to as family) were sampled for DNA extraction. Hence a total of 1120 progeny plants (eight seedlings × 20 families × seven populations (collection sites) were included in the experiment. The 140 maternal plants were also included. Sampling from the geographically different regions was aimed at assessing environmental effects on outcrossing rate and/or the possible inclusion of more races prevailing in different regions.

DNA extraction, polymerase chain reaction (PCR) and fragment size determination

DNA was extracted in two steps: (i) *in situ* collection of leaf squashes from maternal plants in the field and from 3–8-week-old seedlings in the greenhouse in Ethiopia using FTA PlantSaver card and Whatman chromatography paper; and (ii) sample purification following the manufacturer's procedure as optimized for sorghum by Adugna *et al.* (2011) at the Stanley J. Aronnof Laboratory, Department of Evolution, Ecology and Organismal Biology, Ohio State University, Columbus, USA.

PCR was carried out using five sorghum simple sequence repeat (SSR) markers (Brown *et al.* 1996) in one multiplex. The markers were selected based on their high polymorphism in earlier studies and after screening their compatibility for multiplexing. PCR followed the Qiagen® multi-master mix kit protocol (Germantownm, Maryland, USA) for SSR multiplex, and for ABI Fluorescence detection of PCR products, forward primers were labelled with fluorochromes of different colours fluorochromes (FAM, HEX and NED) (table 1). PCR was carried out in 12 μL total volume of reaction mix containing 200 nM of each primer pair in a multiplex, 1 μL of template DNA, 2.6 μL of sterile double distilled water (ddH_2O), 6 μL of Qiagen Multiplex PCR 2× Master mix. PCR reactions were run in a Master cycler (Eppendorf, Hamburg, Germany) with an initial denaturation step of 15 min at 95°C, followed by 35 cycles of 30 s at 94°C, 90 s at 58°C, 60 s at 72°C, 30 min at 60°C, and held at 4°C following Qiagen protocol for microsatellite multiplexes.

For fragment-size determinations, 2 μL of the PCR product was diluted with 14 μL of ddH_2O and then 2 μL of the diluted PCR product was added to 14 μL of 36 : 1 Hi-Di-Formamide : GenScan^{TM}/350 Rox^{TM} size standard in a 96-well microtitre plate and denatured at 95°C for 5 min and cooled on ice for at least 5 min. Allele size scoring of the PCR fragments was performed by ABI 3100 Genetic Analyzer (DNA sequencer) and sizes were read using the associated GeneMapper 3.7 software (Applied Biosystems, Foster City, USA). To verify the repeatability of allele scoring, we included samples of a well-studied sorghum inbred line, BT×623, by obtaining seeds from the Department of Agronomy, Purdue University, West Lafayette, USA. To exclude the possible effects of imprecise DNA fragment sizes due to stuttering, large-allele dropout or null alleles on genotyping, the software Allelobin (Prasanth *et al.* 2006) was used to classify observed SSR allele sizes into representative discrete allele sizes using a variation of the least-squares minimization algorithm of Idury and Cardon (1997).

Data analysis

Estimation of in situ outcrossing rate: Mating system parameters were estimated both at the population level and at family level based on the mixed-mating model of Brown and Allard (1970) using the most popular procedure of Ritland and Jain (1981) (revised in 1990 and 1996) and the multilocus mating system program (MLTR for Windows revised v3.3 (Ritland 2008) accessible at http://genetics.forestry.ubc.ca/ritland/programs.html. For estimation of outcrossing rate at population level and family level, the families within populations and the individuals within families, respectively, were used as the units of resampling. To solve the likelihood equation for the maximum-likelihood estimates, the Newton–Raphson (NR) numerical method was used as recommended in the manual. Standard errors of the estimates were based on 1000 bootstraps. The recommended initial values that allow iterations to start for each parameter

Table 1. Sequences and repeat motifs of the sorghum SSR markers used in the study.

Marker	Flanking sequences (5′–3′)	Repeat motif
Sb5-206	F: HEX-8ATTCATCATCCTCATCCTCGTAGAA	
	R: AAAAACCAACCCGACCCACTC	$(AC)_{13}/(AG)_{20}$
Sb1-1	F: FAM-6TCCTGTTTGACAAGCGCTTATA	
	R: AAACATCATACGAGCTCATCAATG	$(AG)_{16}$
Sb6-34	F: HEX-8AACAGCAGTAATGCCACAC	
	R: TGACTTGGTAGAGAACTTGTCTTC	$((AC)/(CG))_{15}$
Sb5-256	F: FAM-6AATTTGCTTTTTGGTCCGTTT	
	R: TAGGAAAGACAGTACTAGAGGTCA	$(AG)_8$
Sb4-72	F: NED-TGCCACCACTCTGGAAAAGGCTA	
	R: CTGAGGACTGCCCCAAATGTAGG	$(AG)_{16}$

were used as they were checked by default. Assumptions of the model are described in Ritland (2002). This model specifies that both selfing and outcrossing (mixed mating) occur in a given population (Shaw and Allard 1982; Ritland 2002). Mixed-mating model is based on three important assumptions (Clegg 1980): (i) self-fertilization occurs at a rate of s and random mating (outcrossing) at a rate of $t = 1 - s$, (ii) the gene frequencies among pollen are distributed identically over all maternal plants; and (iii) maternal genotype does not determine the rate of outcrossing. The mating system parameters that were estimated from the progeny array data and the maternal genotype data of each population and the combined analysis include (i) the multilocus population outcrossing rate (t_m); (ii) the single-locus population outcrossing rate (t_s); (iii) the single-locus inbreeding coefficient of maternal parents (F_m). The significance of the multilocus outcrossing rates of the different populations was tested using Wilcoxon paired-match test using GenStat software v7.2 (VSN International 2007). Pearson's correlation of the multilocus outcrossing rate of the progenies with plant height, population density, and gene diversity of the maternal plants was also calculated.

Inbreeding

Three types of inbreeding coefficient—inbreeding coefficient of the maternal plants (F_m), inbreeding coefficient of the progenies (F_p), and equilibrium inbreeding coefficient expected from observed outcrossing (F_{eq})—were computed. The single-locus inbreeding coefficient of maternal parents (F_m) was computed with MLTR program. Coefficient of inbreeding of the progenies (F_p) was computed using FSTAT (Goudet 2002), by randomly taking one progeny individual from each family. The inbreeding coefficient expected at equilibrium (F_{eq}) was computed using the method of Fyfe and Bailey (1951) as $F_{eq} = \frac{(1-t_m)}{(1+t_m)}$, where t_m is the multilocus population-level outcrossing rate.

Estimation of biparental inbreeding

The presence of biparental inbreeding, mating between related individuals, was tested in several ways as follows: (i) the difference between t_m and t_s: the presence of biparental inbreeding is declared if this difference is positive because single-locus estimates include all apparent selfing due to biparental inbreeding, whereas multilocus estimates exclude much of the apparent selfing due to biparental inbreeding (Shaw *et al.* 1981), as an observed outcross at any locus overrides the apparent selfing manifested at other loci (Ritland 2002). (ii) By comparing the coefficient of inbreeding of the progenies (F_p) and the inbreeding coefficient expected at equilibrium (F_{eq}), following Brown (1979). If biparental inbreeding occurs F_{eq} will be less than F_p (Neel *et al.* 2001). (iii) Evaluating the correlation of paternity (fraction of siblings that share the same male parent), (iv) Correlation of selfing among loci (r_s). With biparental inbreeding, this quantity is less than one (Ritland 2002). (v) Single-locus versus multilocus correlated paternity: if single-locus is greater than multilocus, there is an effect of population substructure on male similarity between outcrosses (Ritland 2008). Methods (iii), (iv), and (v) were performed using MLTR.

Results

SSR polymorphism and extent of outcrossing rate

The SSR primer pair sb4-72 failed to amplify the DNA of most of the families in Kobo population. Hence, this population was analysed separately using PCR with the remaining four pair of primers and excluded from the combined analysis. The number of observed alleles per locus for all families ranged from 16 (sb5-256) to 56 (sb1-1) showing that the SSR markers used were highly polymorphic. Multilocus outcrossing rate (t_m) ranged from 0.33 (Humera/Donkey) to 0.65 (Abuare) (table 2). The average multilocus outcrossing rate (t_m) of all populations excluding Kobo (which had 0.19) was 0.51. However, t_m was in the extreme range of 0 to 1.00 in some families. This variation in outcrossing rate among wild sorghum populations collected from different geographical regions was significant ($P = 0.016$).

Multilocus family outcrossing rate in the progenies was found to have no correlation with plant height ($r = -0.046$, $P = 0.61$), head width ($r = -0.068$, $P = 0.45$) and

Table 2. Estimates of the population-level mating system parameters of wild sorghum.

Estimate[a]	Ghibe	Bisidimo	Mandura	Humera	Hagereselam	Abuare	Kobo	Combined[b]
t_m	0.57 (0.07)	0.48 (0.07)	0.45 (0.05)	0.33 (0.07)	0.61 (0.06)	0.65 (0.05)	0.19 (0.07)	0.51 (0.03)
t_s	0.24 (0.04)	0.23 (0.04)	0.27 (0.03)	0.23 (0.05)	0.34 (0.06)	0.38 (0.04)	0.14 (0.07)	0.21 (0.01)
$t_m - t_s$	0.33 (0.04)	0.25 (0.04)	0.19 (0.03)	0.10 (0.02)	0.26 (0.04)	0.28 (0.03)	0.05 (0.02)	0.30 (0.01)
$r_{p(m)}$	0.73 (0.13)	0.93 (0.09)	0.14 (0.04)	0.71 (0.12)	0.55 (0.11)	0.54 (0.11)	0.21 (0.23)	0.78 (0.04)
$r_{p(s)}$	0.49 (0.16)	0.72 (0.17)	−0.07 (0.04)	0.72 (0.18)	0.24 (0.08)	0.42 (0.13)	−0.03 (0.23)	0.51 (0.06)
r_s	0.13 (0.06)	1.10 (0.18)	0.34 (0.07)	0.42 (0.14)	0.27 (0.12)	0.24 (0.08)	0.58 (0.18)	1.09 (0.20)

Figures in parentheses indicate standard errors;
[a] t_m, multilocus outcrossing rate; t_s, single locus outcrossing rate; $r_{p(m)}$, correlation of multilocus paternity; $r_{p(s)}$, correlation of single locus paternity; r_s, correlation of selfing among loci; [b]combined value does not include Kobo population.

head width ($r = 0.026$, $P = 0.775$). Moreover, the correlation between population outcrossing rate and panicle compactness was negative, but not significant ($r = -0.374$, $P = 0.465$). The population outcrossing rate had highly significant correlation with gene diversity (expected heterozygosity) of the maternal plants ($r = 0.714$, $P = 0.01$), but no correlation with altitude ($r = 0.013$, $P = 0.978$). The inbreeding coefficient of the maternal plants (F_M) had no correlation with plant density ($r = 0.52$, $P = 0.291$), or altitude ($r = 0.59$, $P = 0.217$). Correlation of outcrossed paternity (r_P) had negative, but not significant correlation with plant density ($r = -0.358$, $P = 0.48$).

Inbreeding

Figure 2 shows the three coefficients of inbreeding estimated from the population and family data. The figure shows that inbreeding coefficient of the progenies (F_p) was greater than both inbreeding coefficient of the maternal plants (F_m) and equilibrium inbreeding coefficient expected from the observed outcrossing (F_{eq}) in all populations, except in Mandura where it was less than F_m. Inbreeding coefficient of the maternal plants (F_m) was in the range 0.016 ± 0.091 (± standard error) in Humera/Donkey to 0.662 ± 0.063 in Mandura, with an average of 0.496 ± 0.029. Similarly, F_p ranged from 0.436 ± 0.036 in Humera/Donkey to 0.710 ± 0.022 in Abuare, with an average of 0.691 ± 0.043.

Biparental inbreeding

Biparental inbreeding was evident in all populations with variable magnitude as measured by the different procedures. First, the single-locus outcrossing rate (t_s) was consistently smaller than the multilocus outcrossing rate (t_m) in all populations (table 2). On average $t_m - t_s = 0.24$, which accounted for 48% of the total outcrossing rate in all families, but the difference was lower in Humera (0.101 ± 0.023) and higher in Ghibe (0.328 ± 0.041) families. Second, the coefficient of inbreeding of the progenies, (F_p), was greater than F_{eq} in six of the seven populations, which reveals that there was more inbreeding in the progenies than expected from the observed outcrossing rate. Third, correlation of outcrossed paternity (r_p) was high in five of the seven populations (table 2), indicating contribution of few fathers to the pollen pool, which in turn was an indication of presence of biparental inbreeding or nonrandom mating. Fourth, the correlation of selfing among loci was less than one in all except in Bisidimo population. Fifth, the single-locus correlated paternity was less than the multilocus ones in all populations except in Humera/Donkey. On evaluation of the coefficient of inbreeding, it was found that F_p was greater than both F_m and F_{eq}.

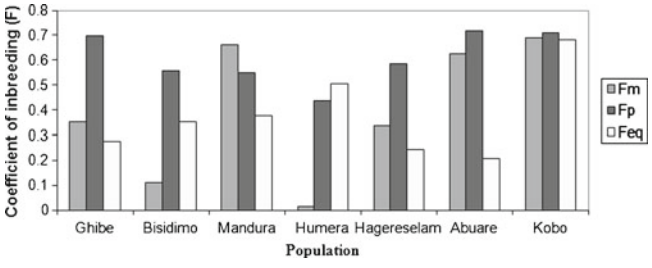

Figure 2. The three coefficients of inbreeding estimated from the population and family data (see text for interpretation of symbols F_m, F_p, and F_{eq}).

Discussion

Variation in outcrossing rate among wild sorghum populations from different geographical regions

The high and variable outcrossing rate (0.33–0.65) among the wild sorghum populations collected from different geographical regions in Ethiopia is in conformation with the results of Muraya *et al.* (2011a). It also verifies that outcrossing rate is higher than it was reported in cultivated sorghum in some of the earlier studies (Ellstrand and Foster 1983; D'je *et al.* 2004). This can be expected because one of the evolutionary consequences of domestication in self-compatible crop plants is the evolution towards inbreeding (Jain 1976). This high variation can be attributed to variation in geoclimatic factors. The factors that might have affected outcrossing rate include temperature, humidity (and moisture) and wind (speed and direction). It is known that the sorghum growing environments of Ethiopia are diverse in geoclimatic variables. The wild sorghum populations included in the present experiment were collected at elevations ranging from 631 m to 1709 m. This wide range of elevations is accompanied by modification of climatic and edaphic factors that might have favoured or disfavoured outcrossing rate and inbreeding of populations of different geographical regions. However, the correlation between outcrossing rate and altitude *per se* was not significant in this study. Moreover, unlike the results of D'je *et al.* (2004), all of the measured phenotypic traits (plant height, panicle length, panicle width, and panicle compactness) had no significant correlation with multilocus outcrossing rate. Probably the proximity to cultivated sorghum might have affected the mating systems because all of the wild sorghum populations were collected from cultivated sorghum fields as weeds, except the Ghibe population, which was found in isolated habitat. This proximity to cultivated sorghum might be associated with panicle size and shape of the cultivars, and the relative density of crop and wild in the mixture as it may also affect pollen competition. In sorghum, self pollen (within wild or within cultivated) is reported to have more seed siring efficiency than pollen from other taxa (Muraya *et al.* 2011b). In the most recent study, Mutegi *et al.* (2012) suggested the possibility of asymmetric gene flow from crop to wild sorghum due to differences in population density. Differences in farmers' cultivation practices were also suggested as factors affecting mating system (Barnaud *et al.* 2008).

The relatively lower outcrossing rate in Kobo families might be due to the reduced number of SSR loci used as outcrossing rate is a function of the number of loci used. On the other hand, the highest outcrossing rate in Abuare families was probably the result of better synchronization of flowering and better anther extrusion ability of the intermixed improved variety developed through selection. Although crop-specific SSR loci were not used, the highest outcrossing rate that was observed in this weedy sorghum intermixed with the improved variety is an indication of the high potential of gene flow from improved varieties to wild sorghums, which in turn is an indication of the high risk of potential transgene flow from future GM sorghum to non-GM cultivated sorghum and wild and weedy sorghum.

The significant correlation between gene diversity and outcrossing rate indicates the role of mating system in shaping the genetic diversity of wild sorghum. The negative correlation between outcrossed paternity (r_P) and plant density ($r = -0.605$, $P = 0.203$), though not significant, may show that when maternal plant density decreases, the probability that any two progenies from the same mother share the same father (being full sibs) increases owing to low concentration of pollen admixture received via wind, as was observed in bombacaceous trees (Murawski *et al.* 1990).

Biparental inbreeding

Biparental inbreeding has significant consequences for evolution of mating systems, yet is difficult to estimate in natural populations (Griffin and Eckert 2003), and its frequency is unknown (Kelly and Willis 2002). However, some procedures have been proposed to estimate biparental inbreeding based on genetic data (e.g. Brown 1979; Ritland 1989, 2002; Kelly and Willis 2002). In this study, three of the five methods applied revealed biparental inbreeding in the Ethiopian wild sorghum populations. The kind of situation observed in this experiment whereby the progenies tend to be more inbred than both their parents and the level of inbreeding expected at equillibrium (F_p is greater than both F_m and F_{eq}) is a characteristic of predominantly outbreeding species (Brown 1979; Neel *et al.* 2001).

The Ghibe wild population was found to be isolated from crop sorghum on either side of the main road from Addis Ababa to Jimma crossing a natural forest. It occurs at high density (135 plants per quadrant, 16 m^2 area) and in a continuous hedge for more than 3 km distance. However, in the current set of populations, no correlation was observed between multilocus outcrossing rate and plant density ($r = -0.084$, $P = 0.875$) which is in agreement with the results of Neel *et al.* (2001) in *Eriogonum ovalifolium* var. *vineum*. Moreover, the plants in this population had highly laxed panicles, which might have contributed to high outcrossing (e.g. D'je *et al.* 2004). The high biparental inbreeding in families derived from this population was perhaps due to limited seed dispersal. However, the biparental inbreeding observed in the rest of the families, whose parents were found intermixed as weeds or found in close proximity with cultivated sorghum, might be due to farmer practices and/or in some cases reduced synchronization of flowering between crop and wild sorghum. Each weedy sorghum plant is surrounded by a mixture of cultivated sorghum plants and other weedy plants, most of which, found in close proximity, could be closely related. Hence it may receive pollen by wind from these related plants and/or possibly from unrelated individuals located a bit farther. If there is reduced synchronization of flowering between the cultivated plants and the weeds, and because of better pollen competition, the greatest chance of mating will be between related individuals, which are in close proximity. Although Tesso *et al.* (2008) reported overlap in flowering between cultivated and wild sorghum in most regions of Ethiopia where they made the survey, the wild sorghum seed for this experiment was collected from those main plants which matured earlier than the cultivated sorghum in all locations except at Abuare, where the cultivars were also early maturing. Our interviews with farmers also support the view that wild sorghums flower and mature earlier than cultivated sorghum. This was further confirmed by a seed maintenance activity held at Melkassa in the main season, 2010, where two samples from each of 30 wild sorghum populations were grown with eight cultivated sorghum landrace populations collected from the same place where the wild sorghum populations were collected. All of the wild sorghum collections flowered earlier than the cultivated ones. However, this does not rule out overlap in flowering between cultivated sorghum and latecoming tillers of wild sorghum as wild sorghum has extended flowering time. In plants like wild and weedy sorghums whose seeds are dispersed over short distances from the maternal plants, biparental inbreeding is not unexpected.

Implications for crop–wild gene flow and conservation

The high outcrossing observed in wild sorghum in Ethiopia can have implications for crop–wild gene flow as there are no apparent barriers for hybridization to occur between the two congeners. Moreover, it has implications for spread of crop genes (including transgenes) into the wild sorghum pool. However, it is not only outcrossing potential that is required for gene flow but also other factors like presence of sexually compatible recipient plants in the vicinity (e.g. Wilkinson *et al.* 2003; Jenczewski *et al.* 2003) and, further, ability of introgression in the recipient populations (Slatkin 1987). Recent case studies also showed that crop–wild gene flow is likely in sorghum (Tesso *et al.* 2008; Muraya *et al.* 2011a; Mutegi *et al.* 2011).

The wild/weedy sorghums have their own mechanisms to survive in the environment, such as mimicry inside crop fields in which case the farmers are not able to identify the weedy types. During flowering, cross-pollination takes place in both directions and can bring about differential consequences, perhaps the major contribution being for the perpetuation of the wild forms. On the other hand, sometimes

farmers in Ethiopia may deliberately leave wild sorghums in and around their fields to allow interpollination with the cultivated sorghum, which could lead to acquisition of beneficial characteristics by the cultivated landraces (Teshome *et al.* 1999). This kind of variation in farmers' practices in different parts of the country might have contributed to the variation in outcrossing rate of the wild sorghum populations collected from diverse environments. Therefore, maintaining life-history traits such as high outcrossing is important for preserving genetic diversity in natural populations. Because of the observed high outcrossing, crop genes (including transgenes) transferred to the wild pool can easily spread, and depending on the fitness advantage conferred by such genes/traits, this may cause genetic erosion in the wild genetic resources. This calls for effective conservation measures to be put in place.

Acknowledgements

This study serves as part of the dissertation research work for the partial fulfillment of the requirements for a Ph.D. degree of the first author. It was supported by a grant from the Biotechnology-Biodiversity Interface (BBI) programme of the United States Agency for International Development (USAID) to Prof. Allison A. Snow, Department of Evolution, Ecology and Organismal Biology, Ohio State University, Columbus, USA.

References

Adugna A., Snow A. and Sweeney P. 2011 Optimization of a high throughput, cost effective and all-stage DNA extraction protocol for sorghum. *J. Agric. Sci. Technol.* **5**, 243–250.

Ayana A., Bekele E. and Bryngelsson T. 2001 Genetic variation in wild sorghum (*Sorghum bicolor* ssp. *verticilliflorum* (L.) Moench) germplasm from Ethiopia assessed by random amplified polymorphic DNA (RAPD). *Hereditas* **132**, 249–254.

Barnaud A., Trigueros G., McKey D. and Joly H. 2008 High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? *Heredity* **101**, 445–452.

Barrett S. C. H. 1983 Crop mimicry in weeds. *Econ. Bot.* **37**, 255–282.

Brown A. H. D. 1979 Enzyme polymorphisms in plant populations. *Theor. Popul. Biol.* **15**, 1–42.

Brown A. H. D. and Allard R. W. 1970 The estimation of Wright's fixation index from genotypic frequencies. *Genetica* **41**, 399–406.

Brown S. M., Hopkins M. S., Mitchell S. E., Senior M. L., Wang T. Y., Duncan R. R. *et al.* 1996 Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor. Appl. Genet.* **93**, 190–198.

Clegg M. T. 1980 Measuring plant mating systems. *Bioscience* **30**, 814–818.

Dahlberg J. A. 2000 Classification and characterization of sorghum. In *Sorghum: origin, history, technology, and production*, (ed. C. W. Smith and R. A. Frederickson), pp. 99–130. John Wiley, New York, USA.

De Wet J. M. J. and Harlan J. R. 1975 Weeds and domesticates: evolution in the man-made habitat. *Econ. Bot.* **29**, 99–107.

Defelice M. S. 2006 Shattercane, *Sorghum bicolor* (L.) Moench ssp. *drummondii* (nees ex steud.) de wet ex davidse—black sheep of the family. *Weed Technol.* **20**, 1076–1083.

D'je Y., Heuertz M., Ater M., Lefebyre C. and Vekemans X. 2004 *In situ* estimation of outcrossing rate in sorghum landraces using microsatellite markers. *Euphytica* **138**, 205–212.

Doggett H. 1988 *Sorghum*, 2nd edition. Longman, London, UK.

Ejeta G. and Grenier C. 2005 Sorghum and its weedy hybrids. In *Crop ferality and volunteerism*, (ed. J. Gressel), pp. 123–135. Taylor and Francis, Boca Raton, USA.

Ellstrand N. C. 2003 *Dangerous liaisons? When cultivated plants mate with their wild relatives*. Johns Hopkins University Press, Baltimore, USA.

Ellstrand N. C. and Foster F. 1983 Impact of population structure on the apparent outcrossing rate of grain sorghum (*Sorghum bicolor*). *Theor. Appl. Genet.* **66**, 323–327.

Enjalbert J. and David L. 2000 Inferring recent outcrossing rates using multilocus individual heterozygosity: application to evolving wheat populations. *Genetics* **156**, 1973–1982.

Felber F., Kozlowski G., Arrigo N. and Guadagnuolo R. 2007 Genetic and ecological consequences of transgene flow to the wild flora. In *Green gene technology. Advances in biochemical engineering/biotechnology*, (ed. A. Fiechter and C. Sautter), pp. 173–205. Springer, Berlin and Heidelberg, Germany.

Fyfe J. L. and Bailey N. T. J. 1951 Plant breeding studies in leguminous forage crops. 1. Natural cross breeding in winter beans. *J. Agric. Sci.* **41**, 371–378.

Gepts P. and Papa R. 2003 Possible effects of (trans) gene flow from crops on the genetic diversity from landraces and wild relatives. *Environ. Biosafety Res.* **2**, 89–103.

Goudet J. 2002 FSTAT, version 2.932. Institute of Ecology, Biology Building, UNIL, Lausanne, Switzerland.

Griffin C. A. M. and Eckert C. G. 2003 Experimental analysis of biparental inbreeding in a self-fertilizing plant. *Evolution* **57**, 1513–1519.

Hamrick J. L. J. and Godt W. 1990 Allozyme diversity in plants. In *Plant population genetics, breeding, and genetic resources*, (ed. H. D. Brown, M. T. Clegg, A. L. Kahler and B. S. Weir), pp. 43–63. Sinauer Publishers, Sunderland, USA.

Holm L. G., Plucknett D. L., Pancho J. V. and Herberger J. P. 1977 *The world's worst weeds: distribution and biology*. University of Hawaii Press, Honolulu, USA.

Idury R. M. and Cardon L. R. 1997 A simple method for automated allele binning in microsatellite markers. *Genome Res.* **11**, 1104–1109.

Jain S. K. 1976 The evolution of inbreeding in plants. *Ann. Rev. Ecol. Sys.* **7**, 469–495.

James C. 2011 Global status of commercialized biotech/GM crops: ISAAA Brief No. 43. ISAAA, Ithaca, NY. (http://www.isaaa.org/resources/publications/briefs/43/).

Jenczewski E., Ronfort J. and Chèvre A. M. 2003 Crop-to-wild gene flow, introgression and possible fitness effects of transgenes. *Environ. Biosafety Res.* **2**, 9–24.

Kelly J. K. and Willis J. H. 2002 A manipulative experiment to estimate biparental inbreeding in monkey flowers. *Int. J. Plant Sci.* **163**, 575–579.

Mann J. A., Kimber C. and Miller F. R. 1983 The origin and early cultivation of sorghums in Africa. *Texas Agric. Exp. Stn. Bull.* 1454.

Murawski D. A., Hamrick J. L., Hubbell S. P. and Foster R. B. 1990 Mating systems of two Bombacaceous trees of a neotropical moist forest. *Oecologia* **82**, 501–506.

Muraya M. M., Mutegi E., Geiger H. H., de Villiers S. M., Sagnard F., Kanyenji B. M. *et al.* 2011a Wild sorghum from different eco-geographic regions of Kenya display a mixed mating system. *Theor. Appl. Genet.* **122**, 1631–1639.

Muraya M. M., Mutegi E., Geiger H. H., de Villiers S. M., Sagnard F., Kanyenji B. M. *et al.* 2011b Investigation of pollen

competition between wild and cultivated sorghums (*Sorghum bicolor* (L.) Moench) using simple sequence repeats markers. *Euphytica* **178**, 393–401.

Mutegi E., Sagnard F., Semagn K., Deu M., Muraya M., Kanyenji B. *et al.* 2011 Genetic structure and relationships within and between cultivated and wild sorghum (*Sorghum bicolor* (L.) Moench) in Kenya as revealed by microsatellite markers. *Theor. Appl. Genet.* **122**, 989–1004.

Mutegi E., Sagnard F., Labuschagne M., Herselman L., Semagn K., Deu M. *et al.* 2012 Local scale patterns of gene flow and genetic diversity in a crop–wild–weedy complex of sorghum (*Sorghum bicolor* (L.) Moench) under traditional agricultural field conditions in Kenya. *Conserv. Genet.* **13**, 1059–1071.

Neel M. C., Ross-Ibarra J. and Ellstrand N. C. 2001 Implications of mating patterns for conservation of the endangered plant *Eriogonum ovalifolium* var. *vineum* (Polygonaceae). *Am. J. Bot.* **88**, 1214–1222.

Pedersen J. F., Toy J. J. and Johnson B. 1998 Natural outcrossing of sorghum and sudangrass in the central great plains. *Crop Sci.* **38**, 937–939.

Prasanth V., Chandra S., Jayashree B. and Hoisington D. 2006 AlleloBin-a program for allele binning of microsatellite markers based on the algorithm of Idury and Cardon (1997). International Crops Research Institute for the Semi, Arid Tropics (ICRISAT), Patancheru, India.

Ritland K. 1983 Estimation of mating systems. In *Isozymes in plant genetics and breeding – Part A*. (ed. S. D. Tanksley and T. J. Orton), pp. 289–302. Elsevier Science Publishers, Amsterdam.

Ritland K. 1989 Correlated matings in the partial selfer, *Mimulus guttatus*. *Evolution* **43**, 848–859.

Ritland K. 2002 Extensions of models for the estimation of mating systems using n independent loci. *Heredity* **88**, 221–228.

Ritland K. 2008 Multilocus mating system program: MLTR, version 3.2 (documentation file). Available with the program at: http://genetics.forestry.ubc.ca/ritland/programs.html.

Ritland K. and Jain S. K. 1981 A model for the estimation of outcrossing rate and gene frequencies using n independent loci. *Heredity* **47**, 35–52.

Schmidt M. R. and Bothma G. C. 2006 Risk assessment for transgenic sorghum in Africa: crop-to-crop gene flow in *Sorghum bicolor* (L.) Moench. *Crop Sci.* **46**, 790–798.

Shaw D. V. and Allard R. W. 1982 Estimation of outcrossing rates in douglas-fir using isozyme markers. *Theor. Appl. Genet.* **62**, 113–120.

Shaw D. V., Kahler A. L. and Allard R. W. 1981 A multilocus estimator of mating system parameters in plant populations. *Proc. Natl. Acad. Sci. USA* **78**, 1298–1302.

Slatkin M. 1987 Gene flow and the geographic structure of natural populations. *Science* **236**, 787–792.

Small E. 1984 Hybridization in the domesticated-weed-wild complex. In *Plant biosystematics* (ed. W. F. Grant), pp. 195–210. Academic Press, Toronto, USA.

Teshome A., Fahrig L., Torrance J. K., Lambert J. D., Arnason Z. J. and Baum B. R. 1999 Maintenance of sorghum (*Sorghum bicolor*, Poaceae) landrace diversity by farmers' selection in Ethiopia. *Econ. Bot.* **53**, 79-88.

Tesso T., Kapran I., Grenier C., Snow A. A., Sweeney P., Pedersen J. *et al.* 2008 The potential for crop-to-wild gene flow in sorghum in Ethiopia and Niger: a geographic survey. *Crop Sci.* **48**, 1425–1431.

VSN International Ltd. 2007 GenStat Discovery Edition 3. VSN International Ltd., Hernel Hempstead, UK.

Wilkinson M. J., Elliott L. J., Allainguillaume J., Shaw M. W., Norris C., Welters R. *et al.* 2003 Hybridization between *Brassica napus* and *B. rapa* on a national scale in the United Kingdom. *Science* **302**, 457–459.

Zhao Z. 2007 The Africa biofortified sorghum project–applying biotechnology to develop nutritionally improved sorghum for Africa. In *Biotechnology and sustainable agriculture 2006 and beyond*, (ed. Z. Xu, J. Li, Y. Xue and W. Yang), pp. 273–277. Springer, Dordrecht, The Netherlands.

Received 20 July 2012, in revised form 14 September 2012; accepted 9 October 2012
Published on the Web: 15 March 2013