

## RESEARCH ARTICLE

# Long-distance gene flow in *Acacia senegal*: Hope for disturbed and fragmented populations

Stephen F. Omondi<sup>1</sup>  | Eunice W. Githae<sup>2</sup> | Damase P. Khasa<sup>3</sup>

<sup>1</sup>Department of Forest Genetics and Tree Improvement, Kenya Forestry Research Institute, Nairobi, Kenya

<sup>2</sup>Department of Biological Sciences, Chuka University, Chuka, Kenya

<sup>3</sup>Centre for Forest Research and Institute for Systems and Integrative Biology, Université Laval, Sainte-Foy, Québec, Canada

**Correspondence**

Stephen F. Omondi, Department of Forest Genetics and Tree Improvement, Kenya Forestry Research Institute, P.O. Box 20412, 00200 Nairobi, Kenya.  
Email: [stephenf.omondi@gmail.com](mailto:stephenf.omondi@gmail.com)

**Funding information**

International Foundation for Science, Grant/Award Number: D5452-1

**Abstract**

Even though pollen and seed dispersals are some of the important factors that determine tree species survival across landscapes, gene dispersal data of important tropical dryland tree species such as *Acacia senegal* that are undergoing various population disturbances remain scarce. Understanding patterns of gene dispersal in these ecosystems is important for conservation, landscape restoration and tree improvement. We investigated pollen and seed mediated gene flow in two *A. senegal* populations of contrasting state (less disturbed and heavily disturbed) using nine microsatellites and 128 genotyping-by-sequencing single nucleotide polymorphism (SNPs) multilocus genotypes of two growth stages (juvenile and adult trees) and their spatial locations. We performed parentage assignments using likelihood approach and undertook spatial genetic structure (SGS) analyses for the two growth stages through correlation among kinship coefficients and geographical distances between pair of individuals. The SNPs showed higher resolving power and assignment rates than microsatellites; however, a combination of the two marker-types improved the assignment rate and provided robust parentage assessments. We found evidence of long-distance (up to 210m) pollination events for both populations; however, the majority of seed dispersal was found closer to the putative maternal parent. On average, parentage analysis showed high amounts of pollen (40%) and seed (20%) immigration in both populations. Significant positive SGS was found only for the adult cohorts in the less disturbed population for distance classes 20 and 40 m, indicating historical short-distance seed dispersals. Our results suggest long-distance gene flow within the species and we recommend conservation of remnant and isolated populations or individual trees to promote genetic connectivity.

**KEYWORDS**

*Acacia senegal*, anthropogenic disturbance, gene dispersal, genetic structure, mating systems

**TAXONOMY CLASSIFICATION**

Conservation ecology, Conservation genetics, Ecological genetics, Evolutionary ecology, Population genetics

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Gene flow in plants is mediated by both pollen and seed dispersal, which vary greatly in terms of mechanism and extent among species (Kremer et al., 2012; Marchini et al., 2016). It is one of the main factors determining the genetic architecture of populations together with drift, selection, and mutation (Burczyk et al., 2004; Grasty et al., 2019). Pollen and seed dispersal, however, have distinct effects on the rate of demographic expansion, and the rate at which genes move across species range. In this light, pollen and seed movements are key parameters in shaping the dynamics and evolution of plant populations (Hardy, 2009). Overall, patterns of pollen and seed-mediated gene flow and variation in the mating system directly influence levels of genetic diversity, connectivity, and population structure. These factors are essential to plant populations' evolutionary potential (Cristóbal-Pérez et al., 2021; Eckert et al., 2010). Population genetics theory predicts disruption of genetic connectivity when populations become small and fragmented or geographically isolated (Klein et al., 2013). Furthermore, loss of allelic diversity via increased levels of genetic drift is anticipated to result in reduced levels of genetic diversity within populations and increased genetic divergence among populations (Aguilar et al., 2019). In many cases, the spatial scale of effective propagule dispersal in trees depends on a variety of physical and biological processes that determine the amount and availability of pollen and seeds. These include propagules movement, viability and the probability of successful pollination leading to viable seed development and seedling establishment rates (Kremer et al., 2012). Different combinations of these processes may yield effective dispersal distances spanning from a few meters to several kilometers, generally following markedly leptokurtic patterns (González-Robles et al., 2021; Kremer et al., 2012).

Reduced plant population size and increased isolation through natural or anthropogenic disturbances may influence mating patterns, with reduced numbers of existing mates, increased levels of self-fertilization, and reduced levels of reproductive success and adaptive fitness in progenies due to inbreeding depression (Jacquemyn et al., 2012). Historically, disturbed species habitats often have small effective population sizes, and geographically fragmented and patchily distributed populations with geographically restricted ranges (Burczyk et al., 2004; Solís-Hernández & Fuchs, 2019). In agreement with population genetic theory, the influence of these factors on gene flow is expected to be largely negative (Cristóbal-Pérez et al., 2021). Meta-analyses have shown that species in disturbed habitats are generally associated with low overall genetic diversity and increased levels of among-population genetic structure as a result of the heightened impacts of genetic drift under conditions of limited genetic connectivity and/or selection under a narrow range of environmental conditions (Leimu et al., 2006). The long-term impacts of restricted gene flow on species of disturbed habitats with small fragmented populations and geographically restricted ranges may include risk of extinction (Ellstrand & Ellam, 1993). Accurate estimations of gene dispersal both by pollen and seed are essential for assessing correlates of fitness in natural

populations and determining the role of sexual selection and ecological factors in breeding systems evolution (Burczyk et al., 2004; Cristóbal-Pérez et al., 2021). This would be useful in designing reserves and conservation strategies that maximize populations' interconnectedness.

Anthropogenic disturbances and habitat fragmentation are quite common in the tropical landscape and represent the major threat for genetic diversity maintenance and natural plant populations' viability (Quesada et al., 2013; White et al., 2002). These threats can adversely affect the exchange of genetic information within remnant populations by promoting reproductive isolation through effective population size reduction and altered pollen and seed dispersal patterns (Solís-Hernández & Fuchs, 2019). The impacts of disturbance on genetic structure and gene flow within fragmented forest landscapes are, however, poorly understood, especially in tropical dryland ecosystems (White et al., 2002). Through human disturbances, decline in population size may reduce reproductive individuals' density. Such disturbances may increase isolation and its effects on the underlying forces of gene flow, which may directly influence the genetic structure within a fragment (Aguilar et al., 2008). Simultaneously, spatial isolation between individuals and relict populations may restrict connectivity because of low levels of pollen and seed dispersal between patches and in the long term may cause loss of genetic variability in the remnant tree populations due to genetic drift and high levels of inbreeding (Sork & Smouse, 2006; Xiao et al., 2016).

*Acacia senegal* trees are highly valued for gum arabic production in the drylands of sub-Saharan Africa (Chikamai & Odera, 2002). The gum is collected as natural or artificially tapped exudates from stem and branches and traded locally and internationally hence significantly contributing to the livelihoods of local populations and the general economy of the countries involved (Fagg & Allison, 2004; Wekesa et al., 2013). Additionally, the species is also useful in agroforestry and fuel wood production (Ballal et al., 2005). Distribution of the species is mainly in the drylands of Kenya; however, some populations are threatened by human disturbances (Omondi et al., 2016). Recent studies show that some *A. senegal* populations have suffered from selective harvesting and land clearance for agricultural production and settlements (Lelon, 2008; Omondi et al., 2016). Such disturbances may affect the species' biological processes including mating systems and gene flow patterns which may be detrimental to the species evolutionary potential and population stability (Omondi et al., 2016). However, no study has investigated the effects of human disturbances on gene flow patterns within the species. Such studies are essential components of designing conservation and restoration plans of degraded landscapes.

Studies show that tropical trees are particularly vulnerable to the effects of habitat disturbances and fragmentation because they naturally occur at low density and have specialized interactions with pollinators and seed dispersers (Ward et al., 2005). It has also been shown that habitat disturbances negatively affect plant reproduction by reducing pollinator activity, pollen deposition, fruit set, outcrossing rates, and the number of progenies in disturbed landscapes

(Aguilar et al., 2006, 2008, 2019; Cunningham, 2000). However, empirical evidence shows contrasting results, in some species. Under certain conditions, habitat disturbance and fragmentation can improve gene movement and connectivity between isolated habitats which ultimately reduces genetic differentiation and maintains high levels of genetic diversity in the long-term (Dick et al., 2003; Lowe et al., 2005). Such contrasting evidence may be brought about by the quality of dispersal of different gene vectors and the degree of isolation and the size of the remnant populations (Fuchs et al., 2003; Nason & Hamrick, 1997). Although pollen flow can persist in fragmented habitats, remnant forest patches may be at risk for loss of genetic diversity because of a reduced number of local and immigrant pollen sources (Sork & Smouse, 2006). Therefore, integration of estimated gene dispersal distance with data on the genetic diversity of pollen pools is required to understand whether isolation and reduction of population size due to anthropogenic disturbances and fragmentation have negative effects on plant reproduction and mating patterns of tree populations.

The aim of the present study was to investigate the patterns of gene flow within two natural populations of *A. senegal* with contrasting disturbance levels so as to determine the effect of habitat disturbance on gene flow patterns. We also evaluated the impact of gene flow on the spatial genetic structure (SGS) of the two populations. Our study hypothesize that the heavily disturbed population of *A. senegal* will display restricted gene flow pattern characterized by short-distance pollen and seed dispersal, leading to non-random distribution of genotypes in space compared to the less disturbed population. We used mating patterns and the SGSs of the pre (mature individuals)- and post (juvenile individuals)-disturbance gene pools to illustrate the extent of gene flow, and separated them into components of identified local against unidentified putative local or immigrant genes. We achieved the aims of the present study using both microsatellite and single nucleotide polymorphism (SNP) markers. Implications of the study findings for conservation and management of the species are discussed.

## 2 | MATERIALS AND METHODS

### 2.1 | Study sites

Two natural populations of *A. senegal* within Lake Baringo woodland, Solit (00°14'N and 36°35'E) and Kimalel (00°21'N and 36°03'E), with contrasting anthropogenic disturbance status were studied. We surveyed the two populations prior to the study and classified them as heavily disturbed (Solit) or less disturbed (Kimalel) based on population disturbance index (PDI) and density of adult trees (Table 1) as reported in Omondi et al. (2016). The PDI comprised of presence/absence of seven disturbance indicator variables including; number of cut tree stumps (percentage proportion of cut stumps against the total number of mature stem), number of old and new charcoal kilns, presence of grazing livestock, level of seedlings browsing (percentage proportion of the number of browsed seedlings against the total number of seedlings), number of recent settlements, number of fenced plots, and presence of cultivated agricultural land (Table 1). For each population, the PDI was calculated as the sum of all present indicators. The choice of indicator variables used in this study was informed by existing literature on factor affecting vegetation dynamics within Lake Baringo ecosystem (Omondi et al., 2016 and references therein) and personal observations. Density of adult *A. senegal* trees was determined by performing inventory of 15 plots measuring 20 × 20 m (400 m<sup>2</sup>) per population. Mean number of adult trees (diameter at breast height, dbh > 5 cm) per hectare was determined for each population. In this study, a tree was considered to be in disturbed habitat when PDI of the population was greater than 5 (presence of more than 60% of disturbance indicators) and the density of adult tree less than 50 per ha. Population with PDI of 1.141 and high density of adult trees was considered as less disturbed while the population with PDI of 6.52 and low adult tree density was considered as heavily disturbed (Table 1).

**TABLE 1** Disturbance indicators, density of adult trees within the two study populations of *Acacia senegal*; presence or absence of each disturbance indicator is indicated in the bracket where one (1) means presence and zero (0) means absent.

| Indicator                          | Population        |                |
|------------------------------------|-------------------|----------------|
|                                    | Heavily disturbed | Less disturbed |
| Number of cut tree stumps          | 67% (0.67)        | 2.1% (0.021)   |
| Number of charcoal kilns           | 16 (1)            | 0 (0)          |
| Presence of grazing livestock      | Yes (1)           | Yes (1)        |
| Level of seedling browsing         | 85% (0.85)        | 12% (0.12)     |
| Number of recent settlements       | 18 (1)            | 0 (0)          |
| Number of fenced plots             | 28 (1)            | 0 (0)          |
| Presence of cultivated land        | Yes (1)           | 0 (0)          |
| Population disturbance index (PDI) | 6.52              | 1.141          |
| Density (Number of adult trees/ha) | 22 (008)          | 414 (061)      |

Source: Omondi et al. (2016).

## 2.2 | Sampling of plant material

We investigated two successive generations (adults and juveniles) in each population to estimate the level of realized gene flow within the two populations. Leaf samples from all adults (>5 cm dbh) and juvenile (<5 cm dbh) individuals were collected from a 100×200 m quadrat within each population. All the sampled individuals were georeferenced using global positioning system (GPS) and the distributions were as shown in Figure 1. We collected leaf samples from a total of 152 adults and 186 juveniles from the less disturbed population and 83 adults and 59 juveniles from heavily disturbed population.

## 2.3 | DNA isolation and microsatellite genotyping

Total genomic DNA was isolated from the leaf samples using a modified cetyltrimethyl ammonium bromide (CTAB) method

described by Hanaoka et al. (2013). We performed polymerase chain reaction (PCR) amplifications using 12 polymorphic microsatellite loci developed by Assoumane et al. (2009) and cross amplified in *A. senegal* (Omondi et al., 2010) for all the sampled individuals. The PCR amplification was performed in 10 μL volume containing 10 ng of DNA, 2× Multiplex PCR Master Mix (Qiagen), 0.01 μM fluorescently labeled forward primer, and 0.15 μM reverse primer. PCR conditions were as follows: one cycle at 95°C for 15 min, 67°C for 1.5 min, and 72°C for 1 min, followed by eight cycles at 94°C for 30 s, 65°C for 1.5 min (2°C decrease after every cycle), and 72°C for 1 min, 24 cycles at 94°C for 30 s, 51°C for 1.5 min and 72°C for 1 min, and a single final extension at 60°C for 30 min. Capillary fragment electrophoresis of PCR products were scored against an internal standard (600 Liz size standards) on an ABI 3500 genetic analyser (Applied Biosystems, California, Inc.), and genotype data captured using GeneMapper 5.0 software (Applied Biosystems). We screened 12 microsatellite markers for suitability to perform parentage analysis by testing for conformance to Hardy-Weinberg

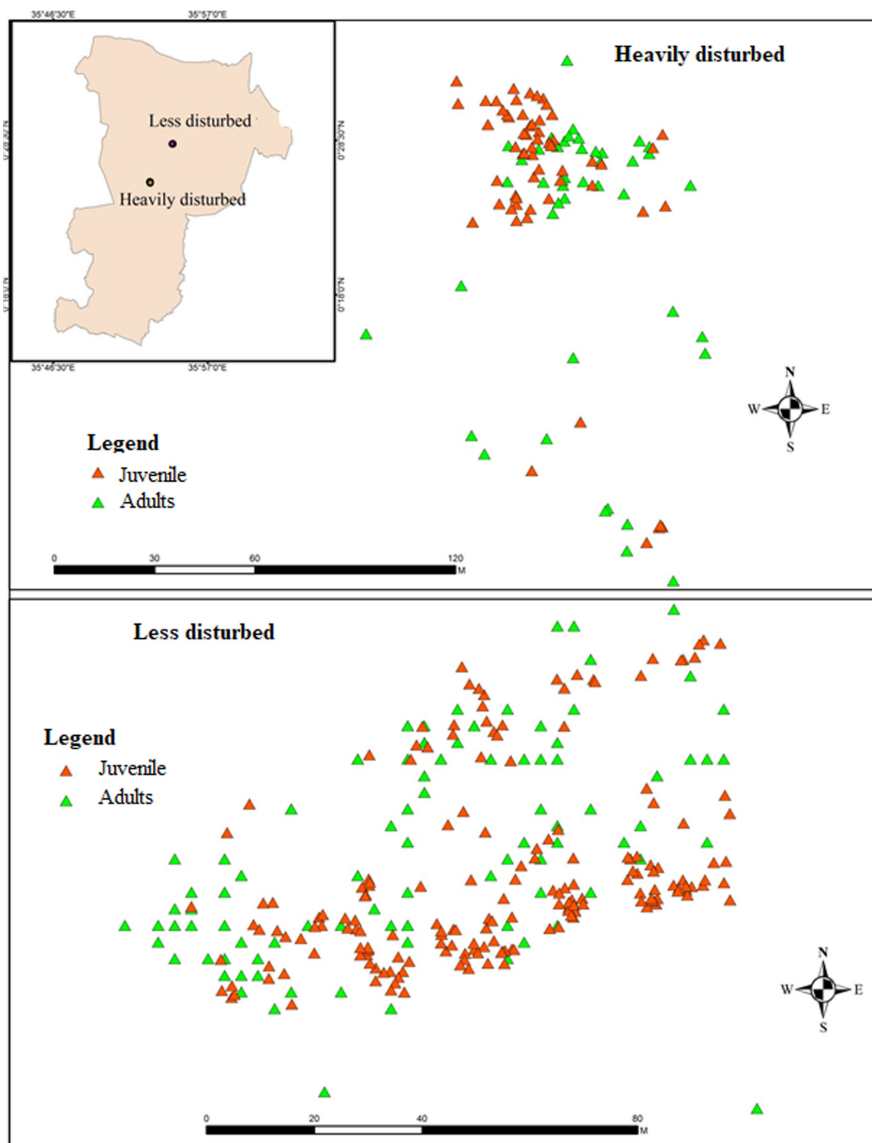


FIGURE 1 Sampling sites and distribution of sampled individuals (adults and juveniles) within the two populations of *Acacia senegal*.

equilibrium (HWE), presence of null alleles, and linkage disequilibrium. Out of the 12 loci, three showed significant departure from HWE, linkage equilibrium, and presence of high frequency of null alleles. The three loci were discarded and the remaining nine loci were used in the downstream analyses.

## 2.4 | SNP discovery and genotyping

We used the Institut de Biologie Intégrative et des Systèmes (IBIS) genotyping by sequencing tool of Université Laval, Quebec, Canada to undertake the genotyping (Torkamaneh et al., 2016). We adapted an in-house script to identify SNPs (Jérôme Laroch, personal communication). The tags were assigned to individuals by their unique nucleotide barcodes and were trimmed to 64 bp by removing the restriction sites. Restriction site-associated DNA tags with at least five reads in an individual were included in a sequence alignment across individuals. Qualifying alleles had to be detected in at least half of the sample size (216). To identify loci that were appropriate for the panel, we limited discovery to biallelic SNPs with no more than one polymorphism per 64-bp sequence. We employed Tassel software version 4.0 (Bradbury et al., 2007) to discover and genotype SNPs from the sequenced tags. Through this, 149,229 SNPs were identified for the 432 *A. senegal* samples.

Out of the 149,229 SNPs identified, we selected a set of 662 SNP markers from the GBS analysis performed on 432 individual tree samples for parentage analysis after further filtering. We set stringent criteria for selecting SNPs as follows: the observed minimum allele frequency (MAF) of 0.3, GC score average > 0.85, and sample call rate > 0.90. In addition, the selected 662 SNPs were further tested for adherence to HWE using a Chi-square goodness-of-fit test in GENEPOP v.4 (Rousset, 2008). This tested significance of the differences between the observed and expected frequencies. SNP markers were included in the analyses only if the *p*-value of this test was larger than .1. For further screening and filtering, we tested the selected SNP markers for evidence of linkage disequilibrium through correlation of allele frequencies ( $r^2 < .11$  normally used to measure linkage disequilibrium in unordered SNPs; Pritchard & Przeworski, 2001). The markers were also subjected to assessment of the presence and frequencies of null alleles. Out of the 662 SNPs screened, only 128 passed these tests and were used in downstream analyses. During the filtering process, we also dropped 37 samples that did not have good sequences, most likely due to poor quality of DNA and therefore, 395 samples were finally used in the downstream analyses.

## 2.5 | Genetic diversity analysis

The genetic diversity parameters were determined for the two populations and marker types. We determined the following parameters; mean number of alleles (*A*), mean effective number of alleles ( $A_e$ ), observed heterozygosity ( $H_o$ ), expected heterozygosity

( $H_e$ ), and inbreeding coefficient ( $F_{IS}$ ) using GenAEx 6.503 (Peakall & Smouse, 2012).

## 2.6 | Parentage analysis

We used the same parentage analysis protocol for both microsatellite and SNP markers using the parentage assignment software, CERVUS version 3.0.7 (Kalinowski et al., 2007). The parentage assignment was evaluated for the most likely parent pairs for all the juveniles from among all possible parents using a likelihood approach. We considered all adult trees of the two populations as candidate parents. The power of our microsatellite and SNP markers to correctly detect putative parents in the populations was estimated by the non-exclusion probability. We set the software (CERVUS) to assign parentage based on LOD scores. During this analysis, we simulated 100,000 juveniles from parent genotyping data, with a mistyping error rate of 0.001, to estimate the LOD score thresholds for assignment probability. The proportion of parent sampled was held at 80% due to the wide range distribution and potential gene flow patterns of the species. We set a strict confidence level at 95%, while the relaxed confidence was at 80%. Expected heterozygosity ( $H_e$ ), polymorphism information content (PIC), and non-exclusion probability (NEP) were calculated for microsatellite and SNP markers. We also evaluated how many SNP loci would be sufficient to ensure assignment of parent pair with 95% confidence. We selected subsets of 128, 100, 80, 60, 40, and 20 most polymorphic SNP loci based on PIC, and repeated simulations using these subsets.

In this study, we assumed that the assigned parent that was physically closest to the seedling was the maternal tree, while we calculated the pollen dispersal distance as the Euclidean distance between the two assigned parent trees. In situations where the same parent was assigned to an offspring as both maternal and paternal parent, we classified the seedling as selfed. However, to avoid the potential biasness of the foregoing assumption, we used all the possible juvenile-mature combinations to calculate mean seed dispersal distance.

## 2.7 | Spatial genetic structure

We undertook the SGS analysis using a combination of microsatellite and GBS-SNP markers to test the hypothesis of random distribution of *A. senegal* genotypes in space within the two study sites. SGS autocorrelation plots were generated in GenAEx v.6.503 (Peakall & Smouse, 2006, 2012) using Queller and Goodnight's relatedness coefficient (*r*; Queller & Goodnight, 1989). We defined even-distance classes of 20 m up to the longest distance class within each site. Significance of SGS was calculated by performing 9999 permutations of the geographic locations of individuals to obtain the 95% confidence interval bars. We further subjected the tests to pairwise bootstrapping comparisons for each distance class to obtain the

95% confidence bounds around mean  $r$ -values (9999 permutations). We also performed heterogeneity test to assess difference in SGS patterns between the two sites (Smouse et al., 2008). A comparative analysis of spatial structure was done by undertaking Moran's Identity test (Moran's  $I$ ) with 10,000 permutations in SPAGeDi software version 1.5 (Hardy & Vekemans, 2002).

### 3 | RESULTS

#### 3.1 | Genetic diversity

Using both microsatellite and SNP markers, higher genetic diversity ( $H_E$  and number of effective alleles) was found in the juvenile cohort than adults in the heavily disturbed population, while there was lower genetic diversity in juvenile in the less disturbed population. However, fixation index was higher for the juvenile than adult individuals in both populations (Table 2).

#### 3.2 | Comparison of marker characteristics between SNPs and microsatellites

Mean number of alleles per locus was 14.7 in the less disturbed population and 13.5 in the heavily disturbed population (overall range 6–17) for microsatellites and two for SNPs in both populations (Table 3). The mean observed and expected heterozygosity, and PIC were greater for microsatellite compared to SNPs. Similarly, non-exclusion probabilities were lower for SNPs compared to microsatellites (Table 3). The combined marker types showed the lowest values for non-exclusion probabilities.

#### 3.3 | Parentage analysis and assignments

By using the 128 SNPs, CERVUS assigned 76.3% (87) and 71.7% (38) of the juveniles in the less disturbed and heavily disturbed populations, respectively, to at least one parent with 95% confidence

(Table 3). When we used a relaxed confidence level (80%), 99% and 100% of the juvenile in the less disturbed and heavily disturbed populations, respectively, were assigned to at least one parent. Only two juveniles could not be assigned to any putative parent in the less disturbed population due to fewer genotyped loci than the minimum number of loci set during simulation. We used similar probabilities for the nine microsatellites loci; however, only 69.1% and 65.8% of the juveniles in the less disturbed and heavily disturbed populations, respectively, were assigned to at least one parent at 95% confidence level. At a relaxed confidence level (80%), 80% of juveniles in the heavily disturbed population were assigned at least one parent while all (100%) of the juveniles in the less disturbed population were assigned to at least one parent. A combination of the marker types improved the assignment rates in the two populations. Using this combination, CERVUS could assign 79% and 86% juveniles to at least one parent with 95% confidence in the less disturbed and heavily disturbed populations, respectively. At the same confidence level, the combined marker types assigned 56.1% and 60.4% of the juveniles in the less disturbed and heavily disturbed populations, respectively, to a parent pair (two putative parents). While comparing the parent pairs assigned to each juvenile, only 80% of all the assignments in both populations by microsatellite were similar to those assigned by SNPs. We observed almost the same amount (82%) of comparison of assignments in the combined marker types with SNPs. In the entire analysis, CERVUS assigned one parent twice to one juvenile as both the maternal and paternal parent in three occasions in the heavily disturbed population.

#### 3.4 | Comparison of parentage assignments from different numbers of SNPs

In this comparison, we undertook simulation of parentage assignment with 128 SNPs and with a successive smaller subset of loci down to a minimum of 20 using PIC of the loci. The mean PIC ranged from 0.337 for the whole data set to 0.372 among the 20 most heterozygous loci. Based on this test, the power of assignments of these sets increased with the number of loci until 60 SNPs then reached a

**TABLE 2** Genetic diversity indices for nine microsatellite and 120 SNP markers studied for adults and juveniles in two populations of *Acacia senegal*.

|             | Heavily disturbed population |       |        |       | Less disturbed population |       |        |       |
|-------------|------------------------------|-------|--------|-------|---------------------------|-------|--------|-------|
|             | Juveniles                    |       | Adults |       | Juveniles                 |       | Adults |       |
| Sample size | 56                           |       | 80     |       | 114                       |       | 148    |       |
| Marker type | SNP                          | Micro | SNP    | Micro | SNP                       | Micro | SNP    | Micro |
| $N_e$       | 1.824                        | 5.35  | 1.480  | 5.01  | 1.912                     | 6.34  | 1.985  | 6.71  |
| $H_O$       | 0.414                        | 0.756 | 0.272  | 0.765 | 0.478                     | 0.742 | 0.483  | 0.755 |
| $H_E$       | 0.451                        | 0.818 | 0.318  | 0.797 | 0.476                     | 0.839 | 0.496  | 0.848 |
| $F$         | 0.083                        | 0.067 | 0.155  | 0.033 | -0.0004                   | 0.111 | 0.026  | 0.101 |

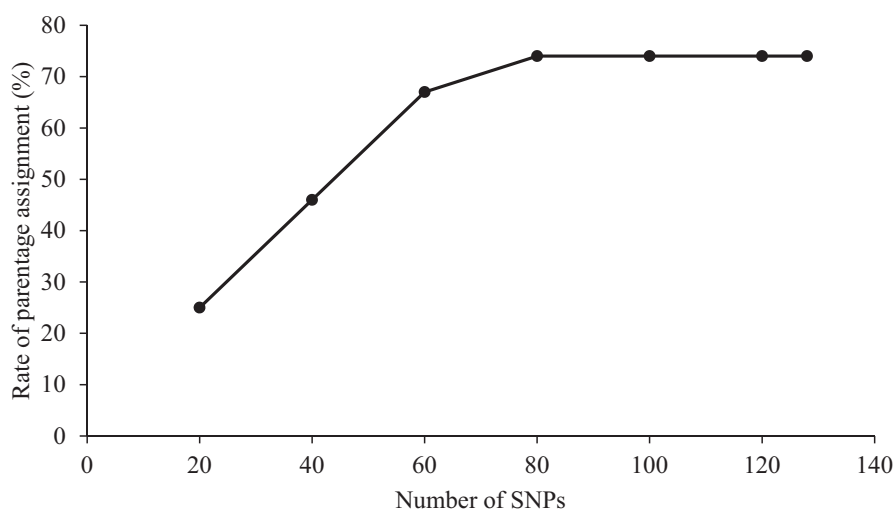
Abbreviations:  $F$ , fixation index;  $H_E$ , expected heterozygosity;  $H_O$ , observed heterozygosity; Micro, microsatellite;  $N_e$ , mean number of effective alleles.

**TABLE 3** Comparison of marker characteristics for SNPs, microsatellite, and a combination of both SNPs and microsatellite markers studied for adult cohorts in two populations of *Acacia senegal*.

| Parameter                        | Less disturbed population |                        |                        | Heavily disturbed population |                       |                        |
|----------------------------------|---------------------------|------------------------|------------------------|------------------------------|-----------------------|------------------------|
|                                  | SNPs                      | Micro                  | SNPs + micro           | SNPs                         | Micro                 | SNPs + micro           |
| Number of adult individuals      | 148                       | 148                    | 148                    | 80                           | 80                    | 80                     |
| Number of loci                   | 128                       | 9                      | 137                    | 128                          | 9                     | 137                    |
| Mean number of alleles per locus | 2                         | 14.7                   | 3.152                  | 2                            | 13.5                  | 2.91                   |
| Mean proportion of loci typed    | 0.997                     | 1.0                    | 0.997                  | 0.998                        | 1.0                   | 0.991                  |
| Mean $H_E$                       | 0.443                     | 0.848                  | 0.529                  | 0.315                        | 0.797                 | 0.365                  |
| Mean $H_O$                       | 0.483                     | 0.755                  | 0.508                  | 0.303                        | 0.765                 | 0.314                  |
| Mean PIC                         | 0.342                     | 0.828                  | 0.414                  | 0.257                        | 0.769                 | 0.314                  |
| NEP (first parent)               | $1.4 \times 10^{-6}$      | $8.858 \times 10^{-4}$ | $6.15 \times 10^{-9}$  | $5.85 \times 10^{-4}$        | $5.05 \times 10^{-3}$ | $2.30 \times 10^{-5}$  |
| NEP (second parent)              | $3.55 \times 10^{-11}$    | $1.729 \times 10^{-5}$ | $1.48 \times 10^{-13}$ | $2.0 \times 10^{-8}$         | $1.65 \times 10^{-6}$ | $2.15 \times 10^{-10}$ |
| NEP (parent pair)                | $1.37 \times 10^{-17}$    | $7.23 \times 10^{-9}$  | $1.04 \times 10^{-21}$ | $1.09 \times 10^{-13}$       | $3.4 \times 10^{-7}$  | $3.75 \times 10^{-17}$ |
| Assignment at 95% confidence     | 76.3%                     | 69.1%                  | 79.0%                  | 71.7%                        | 65.8%                 | 86.0%                  |

Abbreviations:  $H_E$ , mean expected heterozygosity;  $H_O$ , observed heterozygosity; micro, microsatellite; NEP, non-exclusion probabilities; PIC, polymorphic information content; SNPs, single nucleotide polymorphism.

**FIGURE 2** Percentage of CERVUS parentage assignments made using different numbers of the most heterozygous single nucleotide polymorphism (SNPs).



plateau at 80 SNPs (Figure 2). Therefore, selecting at least 80 SNP loci would be needed to ensure that parentage would always be assigned with 95% confidence.

### 3.5 | Assessment of pollen and seed dispersal

In this analysis, a combination of SNP and microsatellite markers were utilized. The gene flow within the species spanned distances ranging from 2.1 to 208.6m (distance from the juvenile to one of the assigned parents or between putative parents). In these assignments, we regarded the closest parent assigned to an offspring as the maternal parent. In this case, *A. senegal* seed movement in the less disturbed population ranged from 2.1m to about 96m with a

mean seed dispersal distance of 39m (Table S1). However, seed dispersal beyond 70m within this population was rare (9.6%). In the heavily disturbed population, seed dispersal distance ranged from 5.1 to 171.2m with a mean dispersal distance of 38.1m. A summary of frequency of assignment of maternal parents to a juvenile at different distance classes is as shown in Figure 3. In the case of pollen, we recorded up to 120 and 208.6m pollen dispersal distances in the less disturbed and heavily disturbed populations, respectively. The mean pollen dispersal distance determined for the less disturbed and heavily disturbed populations was 43.8 and 89.9m, respectively (Table S2). In the less disturbed population, we found greater proportion (72%) of pollen dispersal frequencies at distances less than 60m; however, about 52% of pollen dispersal in the heavily disturbed population was more than 80m.

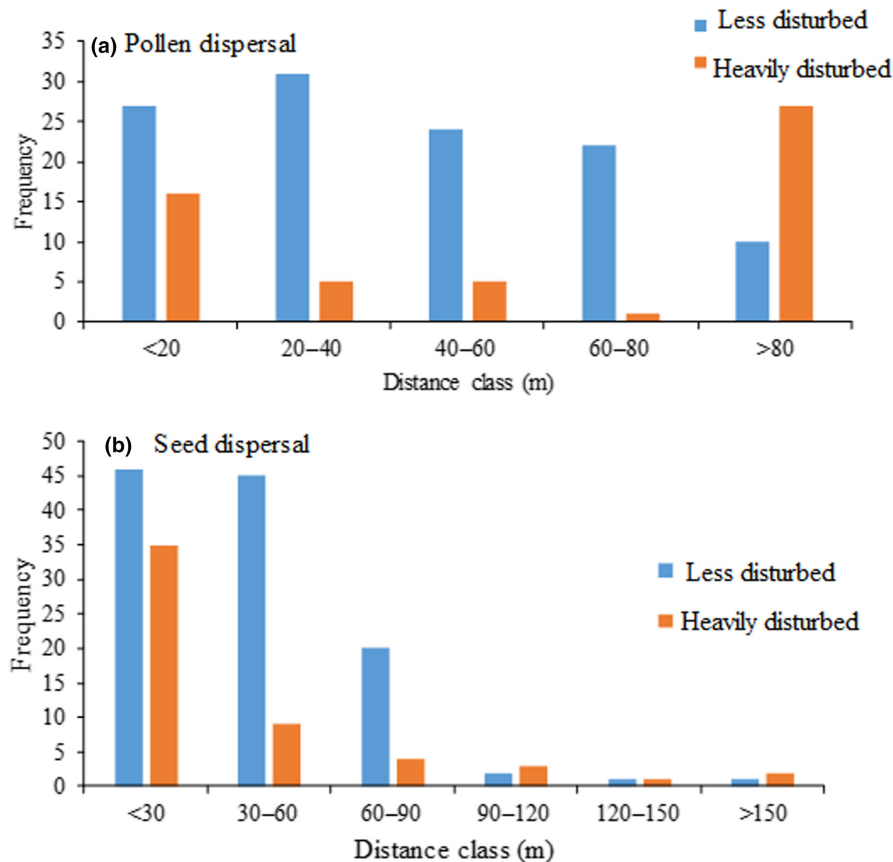


FIGURE 3 Frequency of pollen and seed dispersal events in different distance classes in *Acacia senegal* in the heavily disturbed ( $n=53$ ) and less disturbed ( $n=114$ ) populations.

### 3.6 | Spatial genetic structure

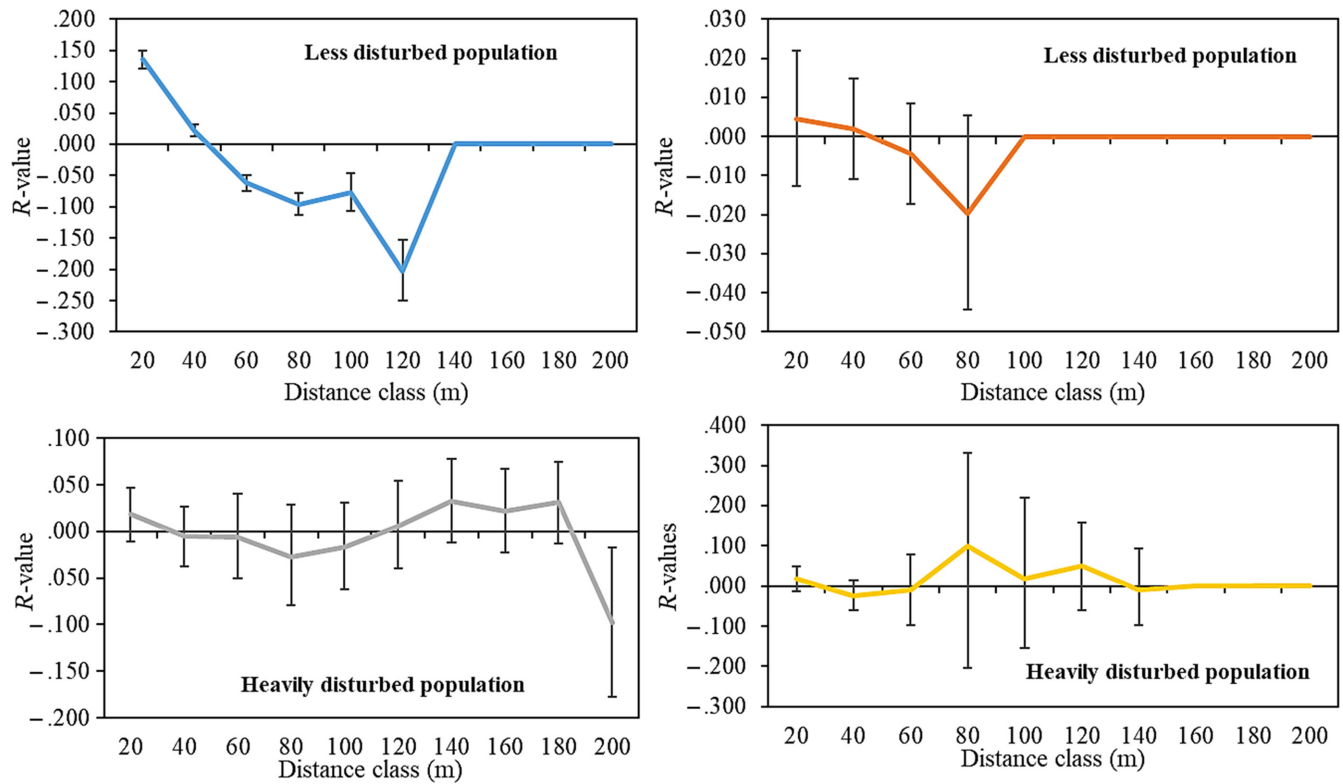
To assess the presence of SGS within the cohorts, we divided the spatial distances into 20m distance classes and a combination of SNP and microsatellite markers were used. Among all the groups analyzed, only the adult cohort of the less disturbed population showed significant spatial structure in all distance classes ( $p < .05$ ; Figure 4). However, we found mean positive values only for the first two distance classes (0–20 and 20–40m). The mean relatedness ( $r$ ) values for the less disturbed population adult cohort ranged from 0.136 to  $-0.203$ . We identified similar significant spatial structure with Moran's  $I$  in SPAGeDi for the distance classes less than 40m; however, higher Moran's  $I$  of 0.153 was found at 20m distance class (20m, Moran's  $I=0.153$ ; 30m, Moran's  $I=0.036$ ; 40m, Moran's  $I=0.02$ ; for all these distance classes  $p < .05$ ). The heavily disturbed population also showed significant spatial structure in the adult cohort at distance class 200m ( $r = -.097$ ,  $p = .03$ ), however, no spatial structure was detected in any distance class in this cohort using Moran's  $I$  analysis. The correlogram of the spatial structure analysis for all the populations and cohorts based on autocorrelation are as shown in Figure 4. The SGS of both the less disturbed and heavily disturbed populations juvenile cohorts were not significant in all the classes ( $p > .05$ ).

When comparing the two study populations using heterogeneity tests, we found positive significant difference ( $p < .05$ ) in spatial population structure (heterogeneous spatial structure) at all distance class sizes for adult cohort (Figure 5), however, no significance was observed for the juvenile cohort.

## 4 | DISCUSSION

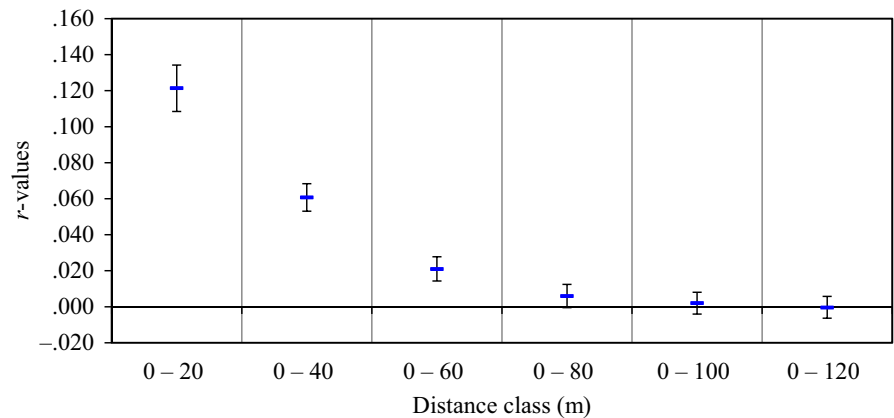
### 4.1 | Marker type and data quality

In trying to understand the different maker characteristics, we generated three different sets of genotypic data for the same set of putative parents and progenies using 128 SNPs, nine nuclear microsatellites and a combination of SNPs and microsatellite marker types. The data sets varied in call rate, mistyping and ability to assign parents to offspring. A similar comparative study on *Eucalyptus nitens* using a SNP chip and two sets of microsatellite panels revealed the same results as the one reported in the present study (Telfer et al., 2015). In their work, Telfer et al. (2015) suggested that data quality is the main contributing factor to the difference. For instance, the near Hardy-Weinberg disequilibrium in some of the microsatellite loci observed in the present study could imply that genotypes scored as homozygous could be heterozygous for null alleles, and that one would most likely be observing true null alleles rather than failed amplifications. Such challenges are common mainly when multiplex PCR analyses are performed. This introduces variable amplification efficiencies for some alleles or loci leading to low signal intensity fragment peaks that do not reach the minimum threshold to call an allele. Genotyping errors have been reported to significantly affect the power of a molecular marker; however, low error rates have been reported for SNPs than microsatellite markers as found in the present study (Hauser et al., 2011; Kaiser et al., 2017; Kalinowski et al., 2007; Labuschagne et al., 2015;



**FIGURE 4** Overlaid SGS autocorrelation plots for up to 200m, even-distance classes in 20m interval, for less disturbed population (adults = 148, juveniles = 114) and heavily disturbed population (adults = 80, juveniles = 53);  $R$ -values, average pairwise relatedness among individuals in the shortest distance class. Error bars represent the upper and lower 95% confidence intervals (CIs) around  $R$ -values estimated by bootstrap resampling.

**FIGURE 5** Combined  $r$ -values of the adult cohort at multiple distance class sizes for the two study sites.



Walling et al., 2010). As pointed out by Labuschagne et al. (2015), each locus adds linearly to the multilocus error rate and therefore the optimum number of SNP loci to be used should be determined prior to the actual analysis. Based on our analysis, 60 informative SNPs with high PIC and  $H_E$  were as good as 128 SNPs in parentage assignment. Large sets of SNPs appear to provide more information than microsatellite loci; however, a combination of the two marker types improved slightly the assignments of the parents to the offspring in our study. Despite the power of SNPs, microsatellites still have clear benefit over SNPs when the amount of information that can be gathered from a single locus is considered and the fact that fewer loci are required to perform parentage assignment (Telfer

et al., 2015). Caution, however, needs to be taken when fewer loci are genotyped, because in such cases an undetected null allele will have a significant impact on the exclusionary power of the marker panel (Labuschagne et al., 2015).

## 4.2 | Pollen and seed dispersal

Despite earlier studies suggesting that *A. senegal* predominantly has long-distance pollen and seed dispersal, we have managed to reveal that the species can also exhibit short-distance gene flow. However, none of the earlier studies employed parentage analysis as in the

present study. Our study showed that majority of gene flow was occurring within the populations; nonetheless, this was happening at different scales in the two populations. Since about 80% of the juveniles were assigned to at least one parent in the two populations, our results suggest that about 20% of the juveniles were realized from seed immigration from outside the two populations. Our finding supports the predominant wind dispersal mechanism of the species that promotes seed dispersal within the population (Bunney et al., 2017). The study revealed higher proportion of seed immigration in the less disturbed than heavily disturbed population. This difference could be attributed to the closeness of the less disturbed population to other populations of *A. senegal* to the north that could act as source of immigrant seed. Generally, mature seeds of *A. senegal* dehisce from the dry pod during dry and windy days and fall sometimes on the ground close to the mother-tree leading to regeneration close to the maternal tree (Orwa et al., 2009). Nevertheless, some seeds that remain stuck on the pods are blown far away from the maternal tree, supporting migration or immigration among closely located populations. Seed immigration between far populations could also be probable due to the involvement of domestic ungulates and pastoral nomadism which is common within the study area (Andersen et al., 2016). It is documented that small livestock, mainly goats and sheep that are commonly found grazing within the study area, feed on dry pods together with seeds of *A. senegal* and move them from one location to another (Fagg & Allison, 2004). The animals ingest the seeds and disperse them through excretion. Furthermore, during the dry seasons when the seed are ready for dispersal, there are no agricultural crops on the farmlands and the livestock freely move among populations of *A. senegal* in search of pasture and water providing opportunities for migration/immigration and long-distance seed dispersal (Andersen et al., 2016; Bunney et al., 2017; Omondi et al., 2016). Successful animal ingestion and seed dispersal of several other *Acacia* species have been reported (Stone et al., 2003).

Our study has shown that large proportion of the juveniles in the two populations had a local pollen source suggesting that close to 40% of the pollen actually immigrated from adjacent or far off populations. This rate of pollen immigration was higher than most pollen dispersal so far reported for other African *Acacia* species. Generally, life history traits, for example, large tree size and longevity tend to encourage large amounts of pollen production and extensive dispersal (Millar et al., 2014). In fact, most African *Acacia* species such as *A. senegal* do not have large amounts of floral nectaries that may favor pollen dispersal by birds (Millar et al., 2014). As a result, insects such as ants, moths, wasps, beetles and bees that are known as generalist pollinators, are thought to be the pollen dispersers in most *Acacia* species (González-Robles et al., 2021; Stone et al., 2003). These insects can influence the tail of dispersal curves and the pollen dispersal directly or through pollen carryover when moving across habitat matrices among populations (Gamba & Muchhala, 2023; Lander et al., 2010).

*Acacia senegal* trees are mainly shrubby and of small stature, but in some cases, they grow up to 15 m high (Raddad et al., 2005). The species also has a prolific flowering tendency and produces small

amounts of nectar and pollen that are foraged by bees, which are capable of expanding pollen distribution to several kilometers while foraging across populations (Fagg & Allison, 2004; Harich et al., 2016). The foraging behavior of honeybees indicate that they visit many flowers in one tree then move to nearby trees and finally fly far away to other trees, perhaps a different population, enhancing pollen migration/immigration (Sujii et al., 2021). The abundance of pollen also influences the distance that can be traveled by a honeybee. For instance, during periods of pollen scarcity, the bees may move longer distances in search of pollen and nectar (Pasquet et al., 2008; Peters et al., 2022; Steffan-Dewenter & Kuhn, 2003). Large proportions of pollen immigration in the two *A. senegal* populations is in fact consistent with the general patterns observed in many studies on common and wide spread tropical tree species that are bee-pollinated (Byrne et al., 2008; Dick et al., 2003, 2008; Millar et al., 2014; Sebbenn et al., 2011; Wandrag et al., 2017; Ward et al., 2005).

### 4.3 | Seed dispersal distance

In both studied populations, we found a strong negative relationship between the probability of maternity and the distance between a seedling and a putative mother tree, as might be expected given barochory as a dominant mode of seed dispersal in *Acacias* (Tybirk, 1997). The within-neighborhood average distance between a putative mother tree and a juvenile was short for both the less disturbed and the heavily disturbed populations. This therefore suggests that the majority of seedlings grew close to the mother tree. Although in the case of heavily disturbed population, there was a greater proportion of long-distance seed dispersal. The seed dispersal patterns were different for the two populations. For instance, there was more frequent short-distance seed dispersal in the less disturbed population than in the heavily disturbed population. This may be explained by the contrasting density of the two populations as was also suggested for *Boswellia papyrifera* by Bekele (2016). In the present study, seed dispersal is more constrained in the less disturbed than the heavily disturbed population. However, this may only be true, unlike in our case, when considering wind as the predominant agent of dispersal and that density acts as a barrier to seed movement while sparse or open structure allows the pods to be blown to long distances (Curtu et al., 2015). Similar results have been reported for some dryland species such as *Swietenia humilis* (White et al., 2002), however, when animal dispersal is also involved, the effect of population density may not be distinct. *Acacia senegal* is dispersed by both wind and ungulates, however, the activities of animals within the two populations were minimal during the study period, suggesting that the frequent short-distance seed dispersal found here may have been impacted by tree density as reported for other species (e.g. *Chamaecrista fasciculata*, Fenster, 1991; *Heliconia acuminata*, Cortes et al., 2013). Furthermore, most of the community members living around the two population have abandoned the pastoralism lifestyle and adopted more sedentary lifestyle with some level of peasantry farming. This change of lifestyle has limited free livestock movement among populations, probably affecting *A. senegal* seed

dispersal. Constrained seed dispersal was reported for *A. tortilis*, a species that heavily depends on animal dispersal for its seed, in the Negev desert (Ward & Rohner, 1997), although Bekele (2016) suggested animal dispersal as the main contributing factor to the relatively high gene flow among populations of *Boswellia papyrifera* in the drylands of Ethiopia corroborating the theory that free animal movement enhances seed dispersal.

#### 4.4 | Pollen dispersal distance

The relationship between distance and probability of mating was significant in both study populations. The average distance between mates within the same area was different between the two population. Although pollen flow may reveal a certain level of stochasticity, the greater average distance between mates observed in heavily disturbed population can be partly related to a lower population density that is, larger distances among individuals within the same area and an opened population structure allowing for more extensive pollinator movement within the population (Jones et al., 2022; Peters et al., 2022). The less disturbed population showed dense distribution pattern of *A. senegal* compared to the heavily disturbed population, which may have promoted the frequent short-distance pollination events in the less disturbed population. This was evidenced by the high frequency of pollination distance of less than 60m. By contrast, the heavily disturbed population that was characterized by sparse density showed high frequencies of pollination distance greater than 80m. Though not studied in the present work, our findings are consistent with density-dependent foraging behavior of bees where a decrease in plant density is accompanied by an increase in flight distance (Byrne et al., 2008; Carper et al., 2022). Similar density dependent pollination distance was reported for *Quercus robur* and *Q. petraea* (Chybicki & Burczyk, 2010). In addition, both the populations were characterized by the presence of other *Acacia* species that may be sharing pollinators (S. F. Omondi, personal observation). Assuming that intraspecific pollination dominates over interspecific pollination, the random distribution of *A. senegal* trees in heavily disturbed population, could have contributed to extensive pollen dispersal than the less disturbed population.

One of the striking results of our study is that no paternal parent was detected for over 40% of the juveniles within the area around the putative mother tree for both populations. This was similar to results reported for other species such as *Quercus* species (Chybicki & Burczyk, 2010). From our results, pollen mediated gene flow in *A. senegal* appears to follow a bi-modal pattern, where substantial local pollination is favored, counterbalanced by long-distance pollen flow (Peters et al., 2022).

#### 4.5 | Spatial structure

The significant SGS for the adult cohorts in the less disturbed population suggests the occurrence of a historical high frequency of short seed dispersal distances and short-distance pollination that may have

resulted in significant relatedness among individuals at shorter distance classes. However, because there was no difference in the SGS among juvenile cohorts in the two populations, our study suggests that the contemporary combined pollen and seed dispersal dynamics within the two populations are relatively similar. Similar findings were reported for *Koompassia malaccensis* which, like *A. senegal* is insect pollinated and wind dispersed (Noreen et al., 2016). Adult trees in the highly disturbed population showed no SGS while adults in the less disturbed population did, suggesting divergent historical genetic events. Based on allele frequencies, the Moran's *I* test demonstrated that only individuals within the lowest spatial distance were closely related. However, individuals within the medium spatial distance may also be related. Such hypothesis can only be tested with fine spatial scale studies that incorporate more individuals across more spatial distances and from many sites or populations.

Because both maternal and paternal genotypes were unknown, it could not be determined absolutely whether the long-distance gene flow reported for this study is due to seed or pollen dispersion, however pollen dispersal is the most probable because of the gene dispersal agents involved. Similar long-distance pollen dispersals (>1000m) have been reported for other tropical trees pollinated by bees and wasps (Nason & Hamrick, 1997). In fact, paternity studies in angiosperms have revealed that individuals from outside the population frequently pollinate larger portions of the seed crop of a target population (Devlin & Ellstrand, 1990; Gaiotto et al., 2003). It is possible that such long-distance gene flow events are more frequent than expected in most tropical tree populations.

## 5 | CONCLUSION

The study shows clearly that *A. senegal* pollen and seeds can move over long distances and at high rates (Ward et al., 2005). These may be much higher than expected for natural undisturbed populations. Similar results were suggested by Trapnell and Hamrick (2005) on their study of orchid populations. In the present study, *A. senegal* have shown levels of pollen flow that are sufficient to prevent or reduce inbreeding and genetic drift. This suggests that even small fragments of the species population have conservation value as stepping stones that can contribute significantly to maintaining populations' genetic connectivity (Nason & Hamrick, 1997; Sabine et al., 2022; White et al., 2002). Additionally, the primary determinant of pollen flow between isolated trees is pollinator behavior and as pointed out by White et al. (2002), all pollinators have a distance threshold beyond which they cannot travel. In that case, the level to which landscape disturbance affects different plant species will be highly variable (Aldrich & Hamrick, 1998). Although the genetic impacts of anthropogenic disturbances and population fragmentation are undoubtedly complex and will vary among species, this study has clearly demonstrated that remnant of fragmented populations and isolated trees of *A. senegal* will provide a buffer to the deleterious genetic consequences of habitat destruction and may be vital to the future long-term viability of the species (Salmona et al., 2023).

## AUTHOR CONTRIBUTIONS

**Stephen F. Omondi:** Conceptualization (lead); data curation (equal); formal analysis (lead); funding acquisition (equal); investigation (lead); methodology (equal); project administration (lead); resources (equal); software (lead); supervision (equal); validation (equal); visualization (equal); writing – original draft (lead); writing – review and editing (equal). **Eunice W. Githae:** Conceptualization (supporting); data curation (supporting); formal analysis (supporting); funding acquisition (equal); investigation (supporting); methodology (equal); project administration (supporting); resources (equal); software (supporting); supervision (supporting); validation (supporting); visualization (supporting); writing – original draft (equal); writing – review and editing (equal). **Damase P. Khasa:** Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (supporting); methodology (supporting); project administration (equal); resources (equal); software (supporting); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal).

## ACKNOWLEDGMENTS

This study was supported and funded by Kenya Forestry Research Institute (KEFRI), the International Foundation for Science (IFS), Stockholm, Sweden to SFO (grant no. D5452-1), Queen Elizabeth II Diamond Jubilee Scholarship to SFO and Natural Science and Engineering Research Council of Canada (NSERC) to DPK. We are also thankful to Professor Yousry El-Kassaby (University of British Columbia) for his critical review of an earlier version of the manuscript.

## CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

## DATA AVAILABILITY STATEMENT

All the data necessary for confirming the conclusions presented in this article is given fully in DRYAD public repository. Dryad, Dataset, <https://doi.org/10.5061/dryad.4tmpg4ff6>.

## BENEFITS GENERATED

Benefits from this research accrue from the sharing of our data and results on public databases as described above. A research collaboration was developed with scientists from the three institutions and all collaborators are included as coauthors and the results have been shared broadly, including outside the scientific community. The authors are committed to international scientific partnerships, as well as institutional capacity building.

## ORCID

Stephen F. Omondi  <https://orcid.org/0000-0003-1931-2672>

## REFERENCES

Aguilar, R., Ashworth, L., Galetto, L., & Aizen, M. A. (2006). Plant reproductive susceptibility to habitat fragmentation: Review and synthesis through a meta-analysis. *Ecology Letters*, 9, 968–980.

- Aguilar, R., Cristóbal-Pérez, E. J., Balvino-Olvera, F. J., Aguilar-Aguila, M. J., Aguirre-Acosta, N., Ashworth, L., Lobo, J. A., Mart'en-Rodríguez, S., Fuchs, E. J., Sanchez-Montoya, G., Bernardello, G., & Quesada, M. (2019). Habitat fragmentation reduces plant progeny quality: A global synthesis. *Ecology Letters*, 22, 1163–1173. <https://doi.org/10.1111/ele.13272>
- Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y., & Loco, J. (2008). Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. *Molecular Ecology*, 17, 5177–5188.
- Aldrich, P. R., & Hamrick, J. L. (1998). Reproductive dominance of pasture trees in a fragmented tropical forest mosaic. *Science*, 281, 103–105.
- Andersen, G. L., Krzywinski, K., Gjessing, H. K., & Pierce, R. H. (2016). Seed viability and germination success of *Acacia tortilis* along land-use and aridity gradients in the eastern Sahara. *Ecology and Evolution*, 6(1), 256–266.
- Assoumane, A. A., Vaillant, A., Mayaki, A. Z., & Verhaegen, D. (2009). Isolation and characterization of microsatellite markers for *Acacia senegal* (L.) Willd., a multipurpose arid and semi-arid tree. *Molecular Ecology Resources*, 9, 1380–1383.
- Ballal, M. E., El Siddig, E. A., Elfadl, M. A., & Luukkanen, O. (2005). Gum arabic yield in differently managed *Acacia senegal* stands in western Sudan. *Agroforestry Systems*, 63(3), 237–245. <https://doi.org/10.1007/s10457-005-4162-x>
- Bekele, A. A. (2016). *Conservation genetics of the frankincense tree*. [PhD thesis]. Wageningen University, Wageningen, Netherlands.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23, 2633–2635.
- Bunney, K., Bond, W. J., & Henley, M. (2017). Seed dispersal kernel of the largest surviving megaherbivore—the African savanna elephant. *Biotropica*, 49, 395–401. <https://doi.org/10.1111/btp.12423>
- Burczyk, J., DiFazio, S. P., & Adams, W. T. (2004). Gene flow in forest trees: How far do genes really travel. *Forest Genetics*, 11, 1–14.
- Byrne, M., Elliott, C. P., Yates, C. J., & Coates, D. J. (2008). Maintenance of high pollen dispersal in Eucalyptus wandoo, a dominant tree of the fragmented agricultural region in Western Australia. *Conservation Genetics*, 9, 97–105.
- Carper, A. L., Warren, P. S., Adler, L. S., & Irwin, R. E. (2022). Pollen limitation of native plant reproduction in an urban landscape. *American Journal of Botany*, 109(12), 1969–1980. <https://doi.org/10.1002/ajb2>
- Chikamai, B. N., & Odera, J. A. (2002). *Commercial plant gums and resins in Kenya*. Executive Printers.
- Chybicki, I. J., & Burczyk, J. (2010). Realized gene flow within mixed stands of *Quercus robur* L. and *Q. petraea* (Matt.) L. revealed at the stage of established seedling. *Molecular Ecology*, 19, 2137–2151.
- Cortes, M. C., Uriarte, M., Lemes, M. R., Gribel, R., Kress, W. J., Smouse, P. E., & Bruna, E. M. (2013). Low plant density enhances gene dispersal in the Amazonian understory herb *Heliconia acuminata*. *Molecular Ecology*, 22, 5716–5729.
- Cristóbal-Pérez, E. J., Fuchs, E. J., Martén-Rodríguez, S., & Quesada, M. (2021). Habitat fragmentation negatively affects effective gene flow via pollen, and male and female fitness in the dioecious tree, *Spondias purpurea* (Anacardiaceae). *Biological Conservation*, 256(2021), 109007. <https://doi.org/10.1016/j.biocon.2021.109007>
- Cunningham, S. A. (2000). Depressed pollination in habitat fragments causes low fruit set. *Proceedings of the Royal Society of London Series of Biological Sciences*, 267, 1149–1152.
- Curtu, A. L., Craciunescu, I., Enescu, C. M., Vidalis, A., & Sofletea, N. (2015). Fine-scale spatial genetic structure in a multi-oak-species (*Quercus* spp.) forest. *Journal of Biogeosciences and Forestry*, 8, 324–332.

- Devlin, B., & Ellstrand, N. C. (1990). The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. *Evolution*, 44, 248–259.
- Dick, C. W., Etchelecu, G., & Usterlitz, F. A. (2003). Pollen dispersal of tropical trees (*Dinizia excelsa*: Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. *Molecular Ecology*, 12, 753–764.
- Dick, C. W., Hardy, O. J., Jones, F. A., & Petit, R. J. (2008). Spatial scales of pollen and seed-mediated gene flow in lowland tropical rain forest trees. *Tropical Plant Biology*, 1, 20–33.
- Eckert, C., Kalisz, S., Geber, M., Sargent, R., Elle, E., Cheptou, P.-O., Goodwillie, C., Johnston, M. O., Kelly, J. K., Moeller, D. A., Porcher, E., Ree, R. H., Vallejo-Marin, M., & Winn, A. A. (2010). Plant mating systems in a changing world. *Trends in Ecology & Evolution*, 25, 35–43.
- Ellstrand, N. C., & Ellam, R. D. (1993). Population genetic consequences of small population size: Implications for plant conservation. *Annual Review of Ecology and Systematics*, 24, 217–242.
- Fagg, C. W., & Allison, G. E. (2004). *Acacia senegal and gum Arabic trade*. *Tropical forestry papers no. 42*. Oxford Forestry Institute.
- Fenster, C. B. (1991). Gene flow in *Chamaecrista fasciculata* (Leguminosae). I. Gene dispersal. *Evolution*, 45, 398–409.
- Fuchs, E., Lobo, J. A., & Quesada, M. (2003). Effects of forest fragmentation and flowering phenology on the reproductive success and mating patterns on the tropical dry forest tree, *Pachira quinata* (Bombacaceae). *Conservation Biology*, 17, 149–157.
- Gaiotto, F. A., Grattapaglia, D., & Vencovsky, R. (2003). Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart). *Journal of Heredity*, 94, 399–406.
- Gamba, D., & Muchhala, N. (2023). Pollinator type strongly impacts gene flow within and among plant populations for six Neotropical species. *Ecology*, 104(1), e3845. <https://doi.org/10.1002/ecy.384512>
- González-Robles, A., Salido, T., Manzaneda, A. J., Valera, F., & Rey, P. J. (2021). Extensive pollen-mediated gene flow across intensively managed landscapes in an insect-pollinated shrub native to semiarid habitats. *Molecular Ecology*, 30, 3408–3421. <https://doi.org/10.1111/mec.15950>
- Grasty, M. R., Thompson, P. G., Hendrickson, E. C., Pheil, A. E., & Cruzan, M. B. (2019). Data from: Fine-scale habitat heterogeneity and vole runways influence seed dispersal in *Plagiobothrys nothofulvus*. *Dryad Digital Repository*. <https://doi.org/10.5061/dryad.612jm6403>
- Hanaoka, S., Omondi, S. F., & Machua, J. (2013). *Basic molecular techniques for tree breeding – Experimental protocols*. Sankeisha Co. Ltd.
- Hardy, O. J. (2009). How fat is the tail? *Heredity*, 103, 437–438.
- Hardy, O. J., & Vekemans, H. (2002). SPAGeDi: A versatile computer program to analyze spatial genetic structure at the individual or population levels. *Molar Ecology Notes*, 2, 618–620.
- Harich, F. K., Treydte, A. C., Ogutu, J. O., Roberts, J. E., Savini, C., Bauer, J. M., & Savini, T. (2016). Seed dispersal potential of Asian elephants. *Acta Oecologia International Journal of Ecology*, 77, 144–151. <https://doi.org/10.1016/j.actao.2016.10.005>
- Hauser, L., Baird, M., Hilborn, R., Seeb, L. W., & Seeb, J. E. (2011). An empirical comparison of SNPs and microsatellites for parentage and kinship assignment in a wild sockeye salmon (*Oncorhynchus nerka*) population. *Molecular Ecology Resources*, 11, 150–161.
- Jacquemyn, H., De Meester, L., Jongejans, E., & Honnay, O. (2012). Evolutionary changes in plant reproductive traits following habitat fragmentation and their consequences for population fitness. *Journal of Ecology*, 100, 76–87.
- Kaiser, S. A., Taylor, S. A., Chen, N., Sillett, T. S., Bondra, E. R., & Webster, M. S. (2017). A comparative assessment of SNP and microsatellite markers for assigning parentage in a socially monogamous bird. *Molecular Ecology Resources*, 17, 183–193.
- Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16, 1099–1106.
- Klein, E. K., Bontemps, A., & Oddou-Muratorio, S. (2013). Seed dispersal kernels estimated from genotypes of established seedlings: Does density-dependent mortality matter? *Methods in Ecology and Evolution*, 4, 1059–1069.
- Kremer, A., Ronce, O., Robledo-Arnuncio, J. J., Guillaume, F., Bohrer, G., Nathan, R., & Bridle, J. R. (2012). Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*, 15, 378–392.
- Labuschagne, C., Nupen, L., Kotze, A., Grobler, P. J., & Dalton, D. L. (2015). Assessment of microsatellite and SNP markers for parentage assignment in ex situ African penguin (*Spheniscus demersus*) populations. *Ecology and Evolution*, 5, 4389–4399.
- Lander, T., Boshier, D., & Harris, S. (2010). Fragmented but not isolated: Contribution of single trees, small patches and long-distance pollen flow to genetic connectivity for *Gomortega keule*, an endangered Chilean tree. *Biological Conservation*, 143, 2583–2590.
- Leimu, R., Mutikainen, P., Koricheva, J., & Fischer, M. (2006). How general are positive relationships between plant population size, fitness and genetic variation? *Journal of Ecology*, 94, 942–952.
- Lelon, J. K. (2008). *Uptake of micronutrients by Acacia senegal varieties and its possible effects on gum arabic quality*. [PhD thesis]. University of Nairobi, Kenya.
- Lowe, A. J., Boshier, D., Ward, M., Bacles, C. F. E., & Navarro, C. (2005). Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity*, 95, 255–273.
- Marchini, G. L., Cole Sherlock, N., Ramakrishnan, A. P., Rosenthal, D. M., & Cruzan, M. B. (2016). Rapid purging of genetic load in a metapopulation and consequences for range expansion in an invasive plant. *Biological Invasions*, 18, 183–196.
- Millar, M., Coates, D., & Byrne, M. (2014). Extensive long-distance pollen dispersal and highly outcrossed mating in historically small and disjunct populations of *Acacia woodmaniorum* (Fabaceae), a rare banded iron formation endemic. *Annals of Botany*, 114(5), 961–971.
- Nason, J. D., & Hamrick, J. L. (1997). Reproductive and genetic consequences of forest fragmentation: Two case studies of neotropical canopy trees. *Journal of Heredity*, 88(4), 264–276.
- Noreen, A. M. E., Niissalo, M. A., Lum, S. K. Y., & Webb, E. L. (2016). Persistence of long-distance, insect-mediated pollen movement for a tropical canopy tree species in remnant forest patches in an urban landscape. *Heredity*, 117, 472–480. <https://doi.org/10.1038/hdy.2016.64>
- Omondi, F. S., Kireger, E., Dangasuk, G. O., Chikamai, B., Odee, D. W., Cavers, S., & Khasa, D. P. (2010). Genetic diversity and population structure of *Acacia senegal* (L) Willd. in Kenya. *Journal of Tropical Plant Biology*, 3, 59–70.
- Omondi, S. F., Ongamo, G. O., Kanya, J. I., Odee, D. W., & Khasa, D. P. (2016). Genetic consequences of anthropogenic disturbances and population fragmentation in *Acacia senegal*. *Conservation Genetics*, 17(6), 1235–1244.
- Orwa, C., Mutua, A., Kindt, R., Jamnadass, R., & Anthony, S. (2009). *Agroforestry database: A tree reference and selection guide version 4.0*. <http://www.worldagroforestry.org/sites/treedbs/treedatabases.asp>
- Pasquet, R. S., Peltier, A., Hufford, M. B., Oudin, E., Saulnier, J., Paul, L., Knudsen, J. T., Herren, H. R., & Gepts, P. (2008). Long-distance pollen flow assessment through evaluation of pollinator foraging range suggests transgene escape distances. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 13456–13461.
- Peakall, R., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295.

- Peakall, R., & Smouse, P. E. (2012). GenAIEx 6.5: Genetic analysis in excel. Population genetic software for teaching and research – An update. *Bioinformatics*, 28, 2537–2539.
- Peters, B., Keller, A., & Leonhardt, S. D. (2022). Diets maintained in a changing world: Does land-use intensification alter wild bee communities by selecting for flexible generalists? *Ecology and Evolution*, 12(5), e8919. <https://doi.org/10.1002/ece3.8919>
- Pritchard, J. K., & Przeworski, M. (2001). Linkage disequilibrium in humans: Models and data. *American Journal of Human Genetics*, 69, 1–14.
- Queller, D. C., & Goodnight, K. F. (1989). Estimating relatedness using genetic markers. *Evolution*, 43, 258–275.
- Quesada, M., Herrerias-Diego, Y., Lobo, J. A., Sánchez-Montoya, G., Rosas, F., & Aguilar, R. (2013). Long-term effects of habitat fragmentation on mating patterns and gene flow of a tropical dry forest tree, *Ceiba aesculifolia* (Malvaceae: Bombacoideae). *American Journal of Botany*, 100(6), 1095–1101.
- Raddad, E. Y., Salihi, A. A., Elfadi, M. A., Kaarakka, V., & Luukkanen, O. (2005). Symbiotic nitrogen fixation in eight *Acacia* Senegal provenances in dryland clays. *Plant Soil*, 275, 261–269.
- Rousset, F. (2008). Genepop'007: A complete implementation of the gen pop software for windows and Linux. *Molecular Ecology Resources*, 8, 103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Sabine, F., Hoppler-Wiedmer, A., Zengerer, V., Egger, G., Schletterer, M., & Scheidegger, C. (2022). Gene flow in a pioneer plant metapopulation (*Myricaria germanica*) at the catchment scale in a fragmented alpine river system. *Scientific Reports*, 12, 8570. <https://doi.org/10.1038/s41598-022-12172-x>
- Salmona, J., Dresen, A., Ranaivoson, A. E., Manzi, S., Le Pors, B., Hong-Wa, C., Razanatsoa, J., Andriaholinirina, N. V., Rasoloharijaona, S., Vavitsara, M.-E., & Besnard, G. (2023). How ancient forest fragmentation and riparian connectivity generate high levels of genetic diversity in a microendemic Malagasy tree. *Molecular Ecology*, 32, 299–315. <https://doi.org/10.1111/mec.16759>
- Sebbenn, A. M., Carvalho, A. C. M., Freitas, M. L. M., Moraes, S. M. B., Gaino, A. P. S. C., da Silva, J. M., Jolivet, C., & Moraes, M. L. T. (2011). Low levels of realized seed and pollen gene flow and strong spatial genetic structure in a small, isolated and fragmented population of the tropical tree *Copaifera langsdorffii* Desf. *Heredity*, 106, 134–145. <https://doi.org/10.1038/hdy.2010.33>
- Smouse, P. E., Peakall, R., & Gonzales, E. (2008). A heterogeneity test for fine-scale genetic structure. *Molecular Ecology*, 17(14), 3389–3400. <https://doi.org/10.1111/j.1365-294x.2008.03839.x>
- Solis-Hernández, W., & Fuchs, E. J. (2019). Effective gene flow patterns across a fragmented landscape in southern Costa Rica for *Symphonia globulifera* (Clusiaceae); a species with mobile seed and pollen dispersers. *Revista de Biología Tropical*, 67(2), S95–S111.
- Sork, V. L., & Smouse, P. E. (2006). Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology*, 21(6), 821–836.
- Steffan-Dewenter, I., & Kuhn, A. (2003). Honeybee foraging in differentially structured landscapes. *Proceedings of the Royal Society of London B*, 270, 569–575.
- Stone, G. N., Raine, N. E., Prescott, M., & Willmer, P. G. (2003). Pollination ecology of *Acacias* (Fabaceae, Mimosoideae). *Australian Systematic Botany*, 16, 103–118.
- Sujii, P. S., Tambarussi, E. V., Grando, C., Silvestre, E. A., Viana, J. P. G., Pedro, H. S., Brancalion, P. H. S., & Zucchi, M. I. (2021). High gene flow through pollen partially compensates spatial limited gene flow by seeds for a Neotropical tree in forest conservation and restoration areas. *Conservation Genetics*, 22, 383–396. <https://doi.org/10.1007/s10592-021-01344-3>
- Telfer, E. J., Stovold, G. T., Li, Y., Silva-Junior, O. B., Grattapaglia, D. G., & Dungey, H. S. (2015). Parentage reconstruction in *Eucalyptus nitens* using SNPs and microsatellite markers: A comparative analysis of marker data power and robustness. *PLoS One*, 10(7), e0130601. <https://doi.org/10.1371/journal.pone.0130601>
- Torkamaneh, D., Laroche, J., & Belzile, F. (2016). Genome-wide SNP calling from genotyping by sequencing (GBS) data: A comparison of seven pipelines and two sequencing technologies. *PLoS One*, 11(8), e0161333. <https://doi.org/10.1371/journal.pone.0161333>
- Trapnell, D. W., & Hamrick, J. L. (2005). Mating patterns and gene flow in the neotropical epiphytic orchid, *Laelia rubescens*. *Molecular Ecology*, 14, 75–84.
- Tybirk, K. (1997). *Reproductive biology and evolution of the genus Acacia*. International Group for the Study of Mimosoideae. International Group for the Study of Mimosoideae. Bulletin No. 20, 45–53.
- Walling, C. A., Pemberton, J. M., Hadfield, J. D., & Kruuk, L. E. B. (2010). Comparing parentage inference software: Reanalysis of a red deer pedigree. *Molecular Ecology*, 19, 1914–1928.
- Wandrag, E. M., Dunham, A. E., Duncan, R. P., & Rogers, H. S. (2017). Seed dispersal increases local species richness and reduces spatial turnover of tropical tree seedlings. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 10689–10694. <https://doi.org/10.1073/pnas.1709584114>
- Ward, D., & Rohner, C. (1997). Anthropogenic causes of high mortality and low recruitment in three acacia tree taxa in the Negev Desert, Israel. *Biodiversity and Conservation*, 6(6), 877–893.
- Ward, M., Dick, C. W., Gribel, R., & Lowe, A. J. (2005). To self, or not to self. A review of outcrossing and pollen-mediated gene flow in neotropical trees. *Heredity*, 95, 246–254.
- Wekesa, C., Luvanda, A. M., Muga, M. O., & Chikamai, B. N. (2013). Market chain analysis of gum arabic trade in Kenya. *Octa Journal of Environmental Resources*, 1(2), 93–106.
- White, G. M., Boshier, D. H., & Powell, W. (2002). Increased pollen flow counteracts fragmentation in tropical dry forest: An example from *Swietenia humilis* Zuccarini. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 2038–2042.
- Xiao, Y., Li, X., Cao, Y., & Dong, M. (2016). The diverse effects of habitat fragmentation on plant-pollinator interactions. *Plant Ecology*, 217(7), 857–868. <https://doi.org/10.1007/s11258-016-0608-7>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Omondi, S. F., Githae, E. W., & Khasa, D. P. (2023). Long-distance gene flow in *Acacia senegal*: Hope for disturbed and fragmented populations. *Ecology and Evolution*, 13, e10292. <https://doi.org/10.1002/ece3.10292>