



OPEN

## The genetic legacy of fragmentation and overexploitation in the threatened medicinal African pepper-bark tree, *Warburgia salutaris*

Annae M. Senkoro<sup>1,2</sup>, Pedro Talhinas<sup>3</sup>, Fernanda Simões<sup>4</sup>, Paula Batista-Santos<sup>3</sup>, Charlie M. Shackleton<sup>1</sup>, Robert A. Voeks<sup>5</sup>, Isabel Marques<sup>6</sup>✉ & Ana I. Ribeiro-Barros<sup>6</sup>✉

The pepper-bark tree (*Warburgia salutaris*) is one of the most highly valued medicinal plant species worldwide. Native to southern Africa, this species has been extensively harvested for the bark, which is widely used in traditional health practices. Illegal harvesting coupled with habitat degradation has contributed to fragmentation of populations and a severe decline in its distribution. Even though the species is included in the IUCN Red List as *Endangered*, genetic data that would help conservation efforts and future re-introductions are absent. We therefore developed new molecular markers to understand patterns of genetic diversity, structure, and gene flow of *W. salutaris* in one of its most important areas of occurrence (Mozambique). In this study, we have shown that, despite fragmentation and overexploitation, this species maintains a relatively high level of genetic diversity supporting the existence of random mating. Two genetic groups were found corresponding to the northern and southern locations. Our study suggests that, if local extinctions occurred in Mozambique, the pepper-bark tree persisted in sufficient numbers to retain a large proportion of genetic diversity. Management plans should concentrate on maintaining this high level of genetic variability through both *in* and *ex-situ* conservation actions.

Medicinal plants have been used worldwide since ancient times, being particularly relevant in the developing world where *ca.* 80% of the population rely on these resources to fulfil their basic health care needs<sup>1–4</sup>. Additionally, at the global level the importance of bio-based compounds continues to grow and phytochemical research towards the identification of new active compounds of medical and nutritional importance is among top research priorities (e.g.<sup>5–14</sup>).

Sub-Saharan Africa harbours a vast repository of plant biodiversity, with 45,000 known vascular plant species<sup>15</sup>, many of which are used in traditional medicine<sup>16–20</sup>. However, efforts to safeguard this biodiversity are often compromised by anthropogenic pressures, with proximal drivers being land transformation, synergistic impacts of fires, grazing, climate change and harvesting (c.f.<sup>17,21–27</sup>), and growing commercialisation of medicinal plant in high demand (c.f.<sup>17,28,29</sup>). The last is motivated by preferences for certain species due to cultural identity, traditions, and lower costs in comparison with modern pharmaceuticals, even under circumstances of access to modern medical facilities<sup>21,30</sup>. On the other hand, the conservation status of many endemic and native species is

<sup>1</sup>Department of Environmental Science, Rhodes University, Grahamstown 6140, South Africa. <sup>2</sup>Departamento de Ciências Biológicas, Universidade Eduardo Mondlane CP 257, Maputo, Moçambique. <sup>3</sup>Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisbon, Portugal. <sup>4</sup>Instituto Nacional de Investigação Agrária E Veterinária, Av. da República, Quinta Marquês, Edifício Sede, 2780-157 Oeiras, Portugal. <sup>5</sup>Department of Geography and the Environment, California State University, 800 N State College Blvd, Fullerton, CA 92831, USA. <sup>6</sup>Forest Research Centre (CEF), Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, 1349-017 Lisbon, Portugal. ✉email: isabelmarques@isa.ulisboa.pt; aribeiro@isa.ulisboa.pt

poorly understood<sup>31,32</sup> and many natural populations may be at risk. Current exploitation rates, often in tandem with other pressures like fire, invasive species, browsing and land transformation, threaten wild populations unless management methods are established, including community-based approaches<sup>17,21,30</sup>.

Under the current scenario of climate change and human population growth, the use of genomic tools is valuable to understand species evolution and adaptation in natural ecosystems<sup>33,34</sup>. The importance of phylogenetic data, genetic diversity, and population structure analyses to characterize the biodiversity of wild species has been well-established in numerous studies (e.g.<sup>35–39</sup>). Microsatellites (Single Sequence Repeats, SSR) are amongst the most efficient and widely used markers for these studies as they are codominant and highly polymorphic loci<sup>40</sup>. Although these markers are species specific, the increasing accessibility to next-generation sequencing<sup>41</sup> has enabled the development of SSRs for the so-called orphan, neglected or wild crop relative species (e.g.<sup>42–45</sup>), although sequencing large plant genomes still remains a challenge<sup>46</sup>.

The pepper-bark tree, *Warburgia salutaris* (Bertol.) Chiov. (Family Canellaceae) is one of the most widely used and traded medicinal plants in southern Africa<sup>47</sup>. This slow growing species is part of an early diverging group of basal angiosperms, thought to be native to eastern and southern Africa<sup>48</sup>. However, subsequent studies confined the distribution of *W. salutaris* to only a sub-region of southern Africa, i.e. South Africa<sup>21,49</sup>, Eswatini (previously known as Swaziland)<sup>24,50</sup>, Zimbabwe<sup>29,51–53</sup>, Malawi<sup>54</sup> and Mozambique<sup>55</sup>. This species is commonly used to treat several ailments such as common colds, throat and mouth sores, or coughs<sup>47,48</sup>.

In the past, sustainable harvesting of medicinal plants was regulated through traditional practises such as taboos, restrictions and harvesting tools<sup>30</sup>. However, with commercial demand increasing, *W. salutaris* groves were repeatedly raided by harvesters that often debarked the whole tree, especially mature plants<sup>56</sup>. Severe harvesting resulted in high tree mortality in many areas and in the extinction of many local populations<sup>21,24,57</sup> and consequently, *W. salutaris* is considered threatened throughout its range<sup>50,58,59</sup>, and listed as an *Endangered Species* in the IUCN Red List<sup>57</sup>. The most extreme case is that of Zimbabwe, where the species is listed as extinct in the wild<sup>29,60</sup>. That resulted in the import of bark supplies in the late 1990s from Mozambique and South Africa<sup>53</sup> being later trafficked from the same countries<sup>29</sup>. For instance, in South Africa, 43% of *W. salutaris* bark in the Johannesburg main market originated from Mozambique, with annual traded amounts estimated at 500–1000 kg<sup>28</sup>. As a result, populations of *W. salutaris* in Mozambique are currently restricted to fragmented patches in the Lebombo Mountains, Tembe River and Futi Corridor (Fig. 1)<sup>48</sup>. According to the Red List classification for Mozambique, this species is considered *Vulnerable VU A2 cd*<sup>58</sup>. Despite this critical situation, only a few studies on the populations dynamics of *W. salutaris* are available; of the 60 research and review papers available in the Web of Science on *W. salutaris* on 05 February 2020, only seven addressed this topic<sup>21,24,48,61–63</sup> while the vast majority are focused on the medicinal applications of this species. Nevertheless, amplified fragment length polymorphisms (AFLPs) have been used to solve genetic relationships between *W. ugandensis*, *W. salutaris* and *W. stuhlmanni* showing a high degree of genetic variation among individuals within populations as well as between populations<sup>62</sup>.

In this work, we have developed SSRs markers for *W. salutaris* to investigate the genetic legacy of exploitation in this slow growing species and to contribute to future re-introduction actions. For that, we have used its best known area of occurrence, Mozambique (Fig. 1) to address the following questions: (1) How is genetic diversity distributed within and among individuals across geographical areas?; (2) Is the genetic structure associated with the geographical distribution?; and (3) Is there any evidence of inbreeding or lack of gene flow between populations?

## Results

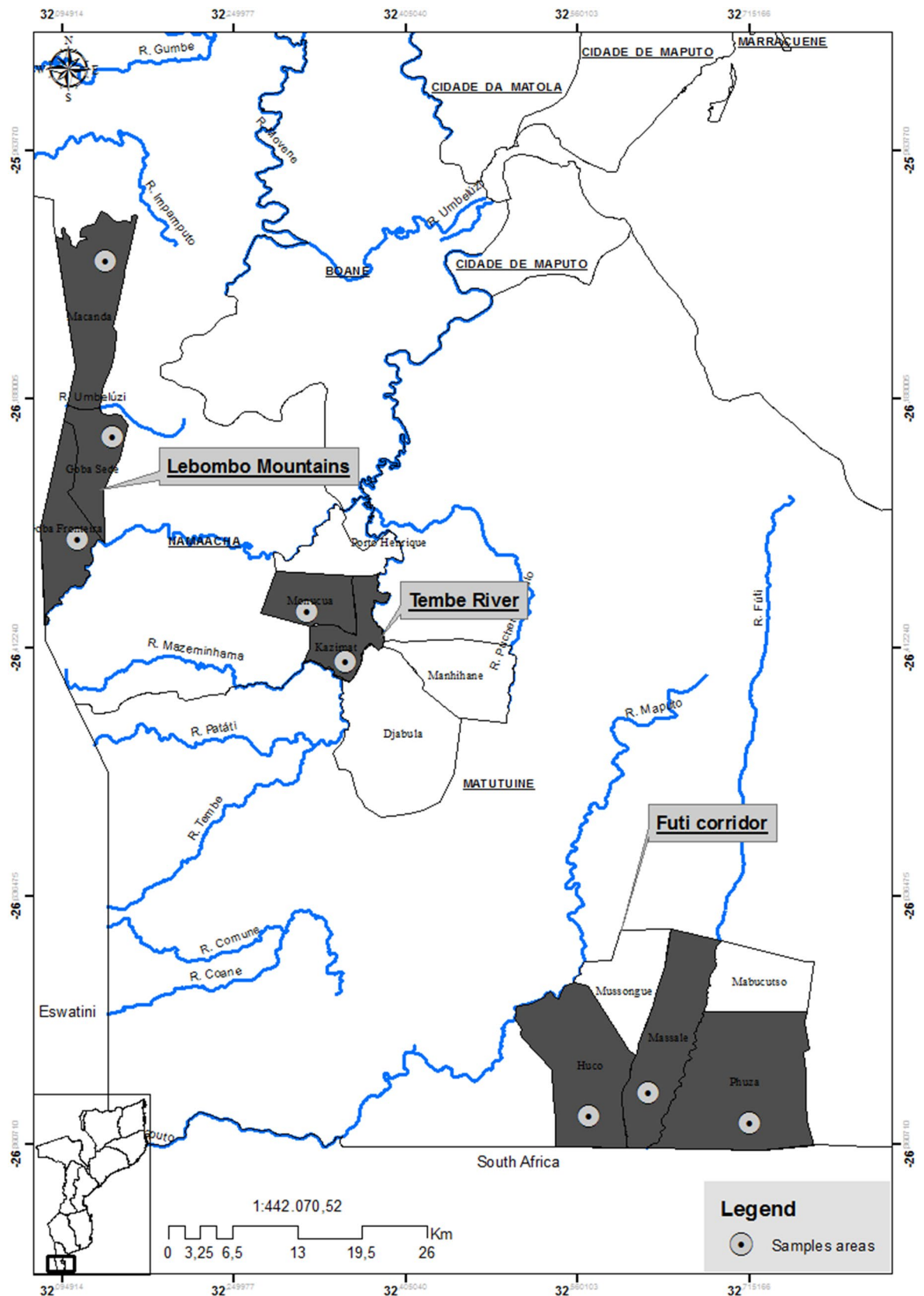
**Genetic diversity.** For each locus, the numbers of alleles varied from three (13-N1132836, 16-N1150626 and 18-N1173706 locus) to nine (31-N2284857 and 43-N1009973 locus) with an average of  $5.8 \pm 2.3$  alleles per locus and a total of 58 alleles considering all loci (Table 1). The average observed and expected heterozygosity per loci varied from  $0.299 \pm 0.186$  (16-N1150626) to  $0.852 \pm 0.062$  (10-N1110523), and from  $0.249 \pm 0.109$  (16-N1150626) to  $0.812 \pm 0.048$  (31-N2284857), respectively.

From the three sampling areas of *W. salutaris* 156 alleles were found in the 48 individuals sampled, being the number of alleles higher in LM than in the other two areas (Table 2). The average Shannon's diversity index (*I*) was also higher in LM than in TR and FC. Observed and expected heterozygosity had similar average values in LM and TR being slightly lower in FC. The polymorphic information content (*PIC*) had high average values while inbreeding coefficients (*F<sub>IS</sub>*) were low and showing negative values in the three sampling areas.

**Population genetic structure and differentiation.** The Bayesian clustering program STRUCTURE found the highest  $\ln P(D)$  and  $\Delta K$  values for  $K=2$  (Fig. 2; Fig. S1). One cluster was predominantly found across LM and TR areas, while a second one characterized the FC area. Nevertheless, some individuals in this last area showed signs of genetic admixture between the two genetic groups (\* indicated in Fig. 2).

The first two coordinates of the principal coordinate analysis (PCoA) explained 22.9% of the total variation, and populations were spatially separated into the two main groups found by STRUCTURE (Fig. 3). The neighbour-joining tree revealed several small clusters although mostly with a very low support (< 30% BS) and overall, with no association between the clusters found and the three geographic areas (Fig. 4) as reported in the other analyses. However, a clear cluster grouped all the FC geographical area.

The pairwise population *F<sub>ST</sub>* values varied from 0.049 (TR vs. LM) to 0.114 (FC vs. TR) revealing moderate levels of genetic differentiation between FC and TR and between FC and LM and lower levels between TR and LM (Table 3).



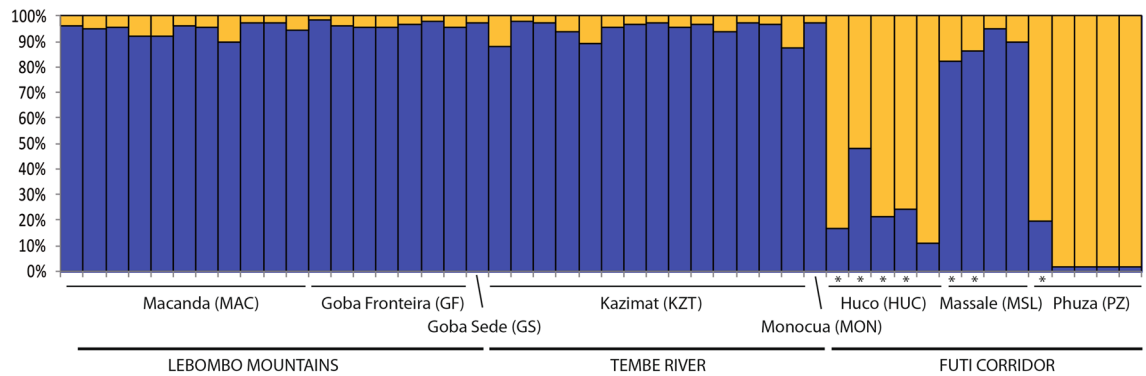
**Figure 1.** Location of the Lebombo Mountains, Tembe River, and Futi Corridor areas and their respective villages in southern Mozambique. Maps were generated with Idrisi Selva v.17.02 environment (Clark Labs, Clark University, [www.clarklabs.org](http://www.clarklabs.org)).

Locus	Repeat motif	Accession number	Primer Sequence 5'-3'	Size range	Na	Ho	He
1-N1002135	(ATG)5	MT515706	F: TATGTTGGGAGAGGG TGAGG R: GTTTAACGACTGCAT CATCCCA	132-174	6	0.487 ± 0.139	0.394 ± 0.101
7-N1082598	(AAT)9	MT515707	F: GTTGATCATAGACAC GCCAAGG R: GTCGTCAACCTAGA GGTCC	161-182	7	0.633 ± 0.085	0.700 ± 0.029
10-N1110523	(TTA)9	MT515708	F: AACCAATTGGCACCTC AAGTC R: GTTGAAGTTGAGGGA AGGGATG	244-262	7	0.852 ± 0.060	0.786 ± 0.023
12-N1126672	(TTG)7	MT515709	F: GTTAAATCTGGACCC ACTTGCC R: GGGTGAATTAGTGAA CGTCTTG	161-180	7	0.805 ± 0.125	0.718 ± 0.074
13-N1132836	(AAG)7	MT515710	F: GTTCCTGCTCCGAGA CCTAGAA R: TCATGAAGAAATCGC AACCA	138-144	3	0.304 ± 0.087	0.296 ± 0.086
16-N1150626	(TGG)5	MT515711	F: GTCTTTGGCGAAATC AGTTGGT R: GAAGGTTTCCAGGTT GGTGA	149-159	3	0.299 ± 0.186	0.249 ± 0.109
18-N1173706	(AAG)6	MT515712	F: GAGTCGCTCGATAT GGACT R: GTTATCCAATGGCCA AGAAACC	164-170	3	0.398 ± 0.105	0.421 ± 0.078
31-N2284857	(TTC)12	MT515713	F: GTCTCTTGCTATCAT GCGGTCA R: CAGATTGGAGAATCC AGACCA	207-263	9	0.771 ± 0.138	0.812 ± 0.078
33-N3477883	(TGA)6	MT515714	F: GTACAAGATTTCATGT GACCGGC R: GCAAGGCATCATATT CACGA	184-200	4	0.550 ± 0.171	0.472 ± 0.124
43-N1009973	(AT)10	MT515715	F: GTTGCCTCATCGAT CTGTA R: GTGCGAACTATGATC GGACGAA	146-185	9	0.439 ± 0.102	0.778 ± 0.027

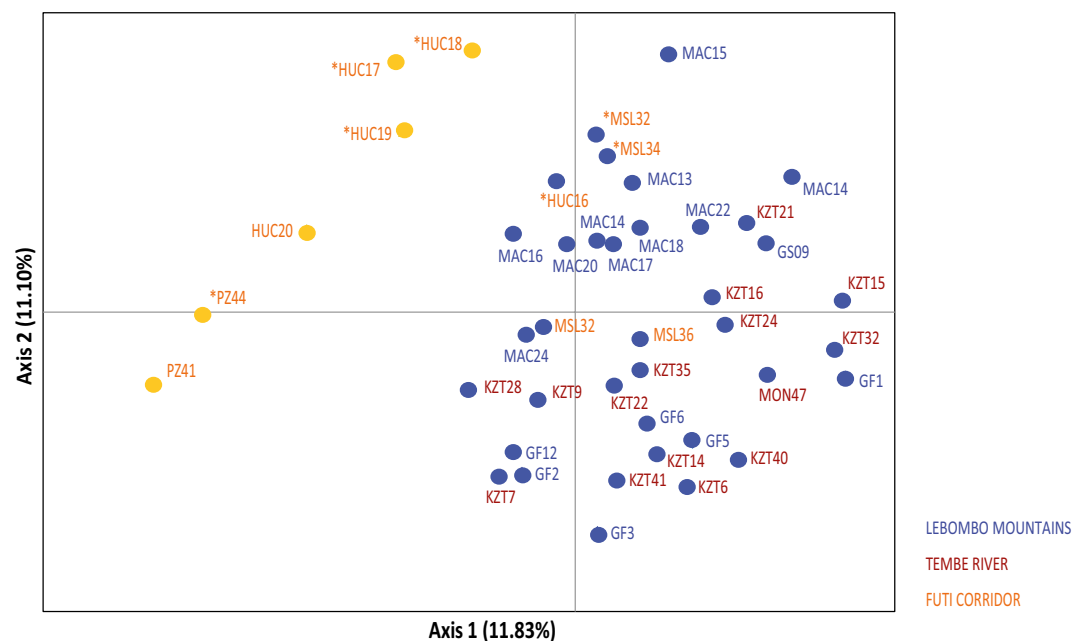
**Table 1.** Characteristics and genetic diversity statistics of the 10 polymorphic microsatellite markers developed for *Warburgia salutaris*. For each loci, the repeat motif, Genbank accession number, primer sequence, and size range (bp) is indicated. *Na* refers to the number of alleles, *Ho* to observed heterozygosity (mean ± SE) and *He* to expected heterozygosity (mean ± SE).

Locus	Lebombo Mountains (LM)					Tembe River (TR)					Futi Corridor (FC)				
	Na	I	Ho	He	PIC	Na	I	Ho	He	PIC	Na	I	Ho	He	PIC
1-N1002135	3	0.809	0.579	0.499	0.499	3	0.840	0.667	0.491	0.491	2	0.340	0.214	0.191	0.191
7-N1082598	7	1.457	0.526	0.672	0.672	9	1.764	0.800	0.758	0.758	8	1.516	0.571	0.671	0.671
10-N1110523	9	1.942	0.895	0.832	0.832	5	1.480	0.733	0.762	0.762	7	1.649	0.929	0.763	0.763
12-N1126672	7	1.716	0.842	0.795	0.795	7	1.739	1.000	0.789	0.789	4	1.061	0.571	0.569	0.569
13-N1132836	2	0.576	0.421	0.388	0.388	2	0.245	0.133	0.124	0.124	2	0.562	0.357	0.375	0.375
16-N1150626	3	0.455	0.158	0.234	0.234	2	0.637	0.667	0.444	0.444	2	0.154	0.071	0.069	0.069
18-N1173706	3	0.942	0.579	0.564	0.564	3	0.680	0.400	0.407	0.407	2	0.469	0.214	0.293	0.293
31-N2284857	14	2.429	0.947	0.895	0.895	9	1.884	0.867	0.816	0.816	6	1.487	0.500	0.727	0.727
33-N3477883	4	0.954	0.526	0.517	0.517	3	0.468	0.267	0.238	0.238	3	1.090	0.857	0.661	0.661
43-N1009973	11	1.980	0.632	0.801	0.810	9	1.827	0.400	0.798	0.798	5	1.438	0.286	0.724	0.724
Average ± SE	6.300 ± 1.274	1.326 ± 0.213	0.611 ± 0.075	0.620 ± 0.068	0.621 ± 0.216	5.200 ± 0.952	1.156 ± 0.203	0.593 ± 0.089	0.563 ± 0.081	0.563 ± 0.256	4.100 ± 0.722	0.977 ± 0.175	0.457 ± 0.089	0.504 ± 0.080	0.504 ± 0.252

**Table 2.** Genetic diversity of *Warburgia salutaris* in the three study areas. *Na* refers to the number of alleles, *I* to Shannon's diversity index, *Ho* to observed heterozygosity (mean ± SE), *He* to expected heterozygosity (mean ± SE) and *PIC* to polymorphic information content.



**Figure 2.** Population structure of *Warburgia salutaris* based on 10 SSRs and using the best assignment result retrieved by STRUCTURE ( $K=2$ ). Each individual sample is represented by a thin vertical line divided into  $K$  coloured segments that represent the individual's estimated membership fractions in  $K$  clusters. Populations and main geographical areas are indicated below following Table 4. Asterisks indicate individuals with a probably of membership lower than 90% to the main genetic cluster, as revealed by STRUCTURE.



**Figure 3.** Principal Coordinate analysis (PCoA) of the studied *Warburgia salutaris* using the scored SSR markers. Percentage of explained variance of each axis is given in parentheses. Population labels follow Table 4. Colour of symbols (circles) indicate the two genetic groups identified by STRUCTURE. Colour of labels follow the three main geographical areas as depicted in Fig. 1. Asterisks as in Fig. 2.

## Discussion

**High genetic diversity and admixture in *Warburgia salutaris*.** Assessment of genetic diversity is critical to understand the ability of a species to cope with changing conditions and environments, especially for threatened species<sup>39,64–68</sup>. In this study, we reported for the first time the development of Single Sequence Repeats (SSR) markers in *W. salutaris* by employing next generation sequencing (Illumina platform). The 10 SSR markers were validated and found to be highly polymorphic, with values similar to the ones found in other threatened species such as *Acer miaotaiense* ( $PIC=0.604$ )<sup>69</sup> or *Corylus avellana* ( $PIC=0.778$ )<sup>70</sup>. These markers are now available to extend *W. salutaris* population studies to a worldwide level. Additionally, the SSRs developed during this work might potentially be suitable to study genetic diversity in other species within the genus *Warburgia*, since only a limited number of studies is available and based on Amplified Fragment Length Polymorphism (AFLP)<sup>62,71</sup>, a time-consuming and costly technique. To the best of our knowledge, the present study represents the first genome size estimation of *W. salutaris* and only the second within the Canellaceae family having a genome size 4 × smaller than *Canella winterana* ( $2C=11.7$  pg<sup>72,73</sup>). The relatively small genome size of *W. salutaris* (see methods) is within the range of the non-expanded genomes of currently known magnoliids (Fig. S2) and may facilitate future genomic initiatives although further analyses are needed to determine its ploidy level.



**Figure 4.** Unrooted neighbour-joining tree of the studied *Warburgia salutaris* based on Nei's  $D_a$  genetic distance. Numbers associated with branches indicate bootstrap values (BS) based on 1000 replications. Only BS above 30 are shown. Colours of branches indicate the two genetic groups identified by STRUCTURE. Colour of circles near each label indicate the three main geographic areas as depicted in Fig. 1. Asterisks as in Fig. 2.

Population	Lebombo Mountains	Tembe River	Futi Corridor
Lebombo Mountains	0.000		
Tembe River	0.049	0.000	
Futi Corridor	0.084	0.114	0.000

**Table 3.** Pairwise population  $F_{ST}$  values for *Warburgia salutaris* in the three study areas.

Due to the heavy harvesting pressure to which *W. salutaris* is subjected in Mozambique<sup>28,48</sup>, genetic diversity levels were expected to be low. However, we found high levels of genetic diversity in the three surveyed areas in comparison to other narrowly distributed species, as for instance, the tropical tree *Paypayrola blanchetiana* ( $N_a$ : 2–5 alleles per locus;  $H_o$ : 0.063–0.563 in the two populations;  $H_e$ : 0.063–0.567 in the first population and 0.063–0.627 in the second)<sup>39</sup>. However, genetic diversity indices of *W. salutaris* were similar to other species where

bark has been heavily-exploited, such as *Cinchona officinalis* ( $N_a$ : 5.2–7.6 alleles per locus;  $H_o$ : 0.580–0.680;  $H_e$ : 0.616–0.717)<sup>74</sup> or even lower than *Himatanthus drasticus* ( $N_a$ : 6–24;  $H_o$ : 0–0.847,  $H_e$ : 0–0.864)<sup>34</sup>.

High levels of heterozygosity may be due to factors including the reproductive system such as self-incompatibility<sup>75</sup> or high gene flow<sup>65,76</sup>. Results from this work revealed a range of the inbreeding coefficient of -0.492 (TR) to -0.363 (LM), which is much lower than those found in e.g. *H. drasticus* (0.248–0.303)<sup>34</sup>, *Calotropis gigantea* (0.167), *C. procera* (0.177)<sup>77</sup>, or *Phoenix theophrasti* (0.9)<sup>78</sup>. The negative inbreeding values found here suggest the existence of random mating<sup>79</sup> among individuals of *W. salutaris* and might also explain the levels of heterozygosity found here. Indeed, the related species *Warburgia ugandensis* has a mixed mating system being predominantly outcrossing<sup>62</sup>. Additionally, insect pollinators of *W. salutaris* such as bees are probably able to travel over the large agricultural blocks separating the three geographical areas studied here, promoting gene flow. Genetic admixture between sites might also be facilitated by frugivorous birds that often eat the berries thereby facilitating the dispersion of seeds. In accordance, we found high levels of genetic admixture between populations with only two genetic clusters being found, one grouping the northern populations and the other one, the southern populations.

Our study suggests that, although some local populations might have been severely affected by harvesting, the pepper-bark tree might have persisted in sufficient numbers in Mozambique to allow outcrossing between sites, retaining a large proportion of genetic diversity. Although there are no records of the historical distribution of this species, the studied populations could be relicts of once much larger populations that persisted in specific locations. In addition, recent conservation efforts might have diminished trade in Mozambique, avoiding severe barking in these populations. Further research should focus on understanding the factors limiting the regeneration of *W. salutaris* trees.

**Population differentiation between geographic areas.** Population differentiation of endangered species is variable. For example, low differentiation was found between populations of *Platanthera leucophaea* ( $F_{ST} < 0.02$  over distances  $< 2 \text{ km}$ <sup>80</sup>) while in *H. drasticus* the differentiation levels were high ( $F_{ST}$  from 0.036 to 0.077 over short distances)<sup>34</sup>. In contrast, the endangered *Paeonia rockii* revealed a high differentiation between populations ( $F_{ST}$  varied from 0.780 to 0.982)<sup>81</sup>. Despite the narrow distributional area of *W. salutaris* in Mozambique, this study revealed a high genetic differentiation between the northern populations located in LM and TR and the southern populations located in FC (Fig. 1). Pairwise  $F_{ST}$  comparisons showed lower genetic differentiation between LM and TR (0.049), which are separated by only 28 km, than either between LM and FC areas (0.084, separated by 81 km) or between TR and FC (0.114, separated by 49 km). STRUCTURE analyses also found a distinct genetic cluster in the FC area, which was also supported by PCoA analyses and the NJ tree. Contrary to LM and TR areas, where *W. salutaris* occurs in slopes and forest patches, in the FC area this species occurs near seasonal pans in thicket vegetation associated with termitaria on clay soils<sup>82,83</sup>. This might imply differences in reproductive ecology, particularly regarding flowering phenology and the activity of pollinators, which would affect gene flow with the other sites, explaining the genetic structure and population differentiation found between the studied sites. Thus, the differentiated FC genetic clusters could be harbouring novel and important alleles and should be given priority in in situ and ex situ conservation strategies in Mozambique<sup>77,84,85</sup>.

**How to conserve a species widely exploited and needed?** Several populations of *W. salutaris* are threatened by fire from slash and burn agriculture, as they occur in adjacent patches or in agricultural lands<sup>48</sup>. Equally, burning of natural vegetation to improve livestock fodder, poaching, and opening of new areas for settlements are also potential threats to the species (e.g.<sup>86–88</sup>). Vegetative propagation of *W. salutaris* is possible through tissue culture<sup>63</sup> although expensive. This species is being largely cultivated ex situ in South Africa<sup>89</sup> and in small scale in Zimbabwe<sup>53</sup> and Mozambique (unpublished data), to encourage the sustainable use of the species. Home gardening would also be important for this species although that requires the involvement of local communities and understanding their perceptions towards the conservation of this species.

Considering the confined distribution and threatened status, the long-term persistence of *W. salutaris* should be secured by conserving the maximum genetic diversity of the species. As it is impossible to designate every natural wild plant habitat as a protected area, nurseries could be implemented to ensure production stability. The disclosure of genetic variation and understanding of genetic relatedness within populations is useful for their sustainable uses<sup>90</sup>. Knowledge of genetic diversity from other countries as the one reported here would also help to implement conservation strategies including re-introduction programs, selecting the most suitable material to be used. Understanding the degree of genetic variation between Mozambique and the neighbouring countries would facilitate transborder conservation actions. Further studies must also be conducted to detect and understand how reductions of natural regeneration or fitness are affected by harvesting. Finally, efforts to educate the local population and landowners on the importance of conserving the natural populations of *W. salutaris* should continue.

## Methods

**Study species.** *Warburgia salutaris* is an evergreen tree, generally 5–10 m tall, but occasionally up to 20 m<sup>57</sup>. The flowers are small ( $< 7 \text{ mm}$  in diameter), white to greenish in colour, generally solitary or in tight, few-flowered heads, borne on short, robust stalks in the axils of the leaves from autumn to winter (April–June). Flowers are bisexual, actinomorphic (having symmetrically arranged perianth parts of similar size or shape that are divisible into 3 or more equal halves). Flowers are visited by many insect species, most especially bees. The flowers develop rounded, oval berries (30 mm in diameter), usually dark-green and turning purple during ripening that occurs throughout winter into early summer (July to December). Dispersion occurs by frugivorous birds that disperse the seeds, although fruits can also drop near the maternal tree. Leaves are glossy and dark green,

Accessions	Location	ID	Lat	Long	Accessions	Location	ID	Lat	Long
GF1 (1)	Goba Fronteira	LM	- 26.23266	32.09810	KZT16 (27)	Kazimat	TR	- 26.40994	32.35490
MAC13 (2)	Macanda	LM	- 26.03522	32.12181	KZT21 (28)	Kazimat	TR	- 26.40391	32.36711
MAC14 (3)	Macanda	LM	- 26.03577	32.12150	KZT22 (29)	Kazimat	TR	- 26.40059	32.35109
MAC15 (4)	Macanda	LM	- 26.03778	32.12730	KZT24 (30)	Kazimat	TR	- 26.40206	32.36188
MAC16 (5)	Macanda	LM	- 26.03692	32.12772	KZT28 (31)	Kazimat	TR	- 26.36735	32.37323
MAC17 (6)	Macanda	LM	- 26.05158	32.11803	KZT35 (32)	Kazimat	TR	26.36737	32.37266
MAC18 (7)	Macanda	LM	- 26.05159	32.11565	KZT40 (33)	Kazimat	TR	- 26.36873	32.37078
MAC19 (8)	Macanda	LM	- 26.81118	32.64545	KZT41 (34)	Kazimat	TR	- 26.36929	32.37334
MAC20 (9)	Macanda	LM	- 26.04696	32.11979	KZT46 (35)	Kazimat	TR	- 26.36935	32.37321
MAC22 (10)	Macanda	LM	- 26.04508	32.11982	MON47 (36)	Monucua	TR	- 26.36952	32.32288
MAC24 (11)	Macanda	LM	- 26.03521	32.12181	Huc16 (44)	Huco	FC	- 26.85013	32.60338
GF2 (12)	Goba Fronteira	LM	- 26.26867	32.10719	Huc17 (45)	Huco	FC	- 26.86159	32.60604
GF5 (13)	Goba Fronteira	LM	- 26.23250	32.09818	Huc18 (46)	Huco	FC	- 26.86169	32.60353
GF6 (14)	Goba Fronteira	LM	- 26.23241	32.09815	Huc19 (47)	Huco	FC	- 26.86129	32.60282
GF12 (15)	Goba Fronteira	LM	- 26.23240	32.09822	Huc20 (48)	Huco	FC	- 26.86025	32.60309
GS09 (16)	Goba Sede	LM	- 26.23238	32.09822	MSL32 (49)	Massale	FC	- 26.83979	32.88339
MAC11 (17)	Macanda	LM	- 26.04509	32.11983	MSL33 (50)	Massale	FC	- 26.86458	32.60790
GF3 (18)	Goba Fronteira	LM	- 26.26879	32.10747	MSL34 (51)	Massale	FC	- 26.80948	32.64368
GF4 (19)	Goba Fronteira	LM	- 26.23233	32.09818	MSL36 (52)	Massale	FC	- 26.80590	32.63823
KZT6 (22)	Kazimat	TR	- 26.41303	32.36338	Pz41 (53)	Phuza	FC	- 26.78824	32.67368
KZT7 (23)	Kazimat	TR	- 26.41190	32.36422	Pz42 (54)	Phuza	FC	- 26.78817	32.67434
KZT9 (24)	Kazimat	TR	- 26.40960	32.36578	Pz43 (55)	Phuza	FC	- 26.78814	32.67383
KZT14 (25)	Kazimat	TR	- 26.40414	32.35073	Pz44 (56)	Phuza	FC	- 26.78760	32.67419
KZT15 (26)	Kazimat	TR	- 26.38806	32.35008	Pz45 (57)	Phuza	FC	- 26.81144	32.66415

**Table 4.** Sampled accessions and locations of *Warburgia salutaris* sorted by geographical area. LM Lebombo Mountains, TR Tembe River, FC Futi Corridor.

with a bitter, peppery taste. The stem is covered by a brown bark marked with corky lenticels and is bitter and peppery and is widely used medicinally. The active compounds (drimanes and sesquiterpenoids) are mostly found in the inner part of the stem and root bark.

**Study area.** The present study was carried out in the districts of Matutuine and Namaacha (Mozambique), in the three areas of known occurrence of *W. salutaris*<sup>48</sup>: (1) Lebombo Mountains (LM) also named the western area, (2) Tembe River (TR) or centre, and (3) Futi Corridor (FC) or eastern area (Fig. 1). The climate is subtropical to tropical, encompassing a wet (October–April) and dry season (May–September). The mean annual temperature ranges from 21 to over 24 °C, and the mean annual rainfall from 600 to 1000 mm<sup>88,91</sup>. In LM, *W. salutaris* is accompanied by *Acacia nigrescens* Oliv., *Acacia burkei* Benth. and *Combretum apiculatum* Sond, although *Aloe marlothii* A. Berger, *Ficus* spp. and *Euphorbia* spp. are found in shallow soils, and *Olea africana* Miller and *Combretum* spp. in steeper stony slopes<sup>92,93</sup>. In TR, *W. salutaris* is found in sand forest patches together with *Pteleopsis myrtifolia* (M.A.Lawson) Engl. & Diels, *Cleistanthus schlechteri* Pax (Hutch.), *Hymenocardia ulmoides* Oliv. and *Monodora junodii* Engl. & Diels<sup>94</sup>. The open savanna woodland links the patches of *W. salutaris*, being composed mainly by *Strychnos* spp., *Terminalia sericea* Burch. ex DC., *Acacia burkei* Benth., *Combretum molle* R. Br. ex G.Don and *Albizia versicolor* Oliv.<sup>95</sup>. In FC, *W. salutaris* occurs near seasonal pans<sup>96</sup> in thicket vegetation associated to termitaria in clay soils<sup>82</sup>. Common tree species found in this community include *Berchemia zeyheri* (Sond.) Grubov, *Pappia capensis* Eckl. & Zeyh. and *Olea europaea* subsp. *africana* (Miller) P.S. Green<sup>97</sup>. The primary economic activities of local residents are subsistence agriculture, livestock rearing, trade of non-timber forest products and migrant labour to South Africa<sup>87,98,99</sup>.

**Population sampling, DNA extraction, genome size value, and SSR development.** Based on the areas of occurrence (Senkoro et al., unpublished data), 48 individuals were sampled: 19 individuals from LM, 15 from TR and 14 from LM (Table 4). Fresh, young undamaged leaves were collected for each individual plant and frozen at -80 °C until DNA isolation. Total genomic DNA was extracted from 50 mg of ground leaves using the InnuSPEED Plant DNA Kit (Analytik Jena Innuscreen GmbH, Germany) according to the manufacturer's protocol. The average yield and purity were assessed spectrophotometrically by OD230, OD260 and OD280 readings (Nanodrop 2000, Thermo Fisher Scientific, Waltham, MA, USA) and visualized by electrophoresis in 1% agarose gels under UV light. Normalized DNA from five individuals of each population was used to develop the SSR markers at CD Genomics (cd-genomics.com/hi-ssrseq.html).

For the development of the markers, we first estimated the nuclear DNA content of *W. salutaris* by flow cytometry using fresh young leaves that were chopped using a razor blade together with an internal standard in a Petri dish containing 1 mL of Woody Plant Buffer<sup>100</sup> following the protocol described in<sup>101</sup>. *Solanum lycopersicum*



'Stupické' ( $2C = 1.96 \text{ pg}$ )<sup>102</sup> was used as internal standard. The nuclear suspension was then filtered through a  $30 \mu\text{m}$  nylon filter, and  $50 \mu\text{g/mL}$  of propidium iodide (PI; Sigma-Aldrich, St. Louis, USA) and  $50 \mu\text{g/mL}$  of RNase (Sigma-Aldrich) were added to stain the DNA only. The fluorescence intensity of nuclei was analysed using a CyFlow Space flow cytometer (Sysmex, Kobe, Japan). Four independent replicates collected from Kazimat (TR) were measured. Conversion of mass values into numbers of base pairs was done according to the factor  $1 \text{ pg} = 978 \text{ Mbp}$ <sup>103</sup>. The mean  $2C$ -value of *W. salutaris* was found to be  $2.91 \text{ pg}$  ( $\pm 0.068$ ), corresponding to an average genome size of  $2845 \text{ Mbp}$  (Fig. S2). Samples had an average coefficient of variation of  $4.18\%$ .

Genomic libraries were constructed using the KAPAHyper prep kit and sequenced by Illumina HiSeq 2500. We firstly used SSRHunter1.3 to screen the potential SSRs from the sequenced data that had at least five repeats (penta-) for 3–5 bp units. Based on the obtained sequences, primers were designed with Primer Premier 5.0 software (Table 1). Fourteen geographically representative samples of *W. salutaris* (LM, TR and FC; Fig. 1) were first used to test microsatellite amplification and to troubleshoot amplification conditions. Amplifications were performed in  $15 \mu\text{l}$  reactions containing:  $1.25 \text{ U}$  TaKaRa Hot startTaq polymerase,  $1 \times$  Buffer I,  $1 \text{ mM}$  dNTPs,  $5 \mu\text{M}$  Primer F and R and  $100 \text{ ng}$  DNA. The PCR amplification conditions were run as follows:  $95 \text{ }^\circ\text{C}$  for  $5 \text{ min}$ ,  $94 \text{ }^\circ\text{C}$  for  $30 \text{ s}$ , 30 cycles of  $56 \text{ }^\circ\text{C}$  for  $30 \text{ s}$ ,  $72 \text{ }^\circ\text{C}$  for  $30 \text{ s}$ ,  $94 \text{ }^\circ\text{C}$  for  $30 \text{ s}$ , 10 cycles of  $53 \text{ }^\circ\text{C}$  for  $30 \text{ s}$ ,  $72 \text{ }^\circ\text{C}$  for  $30 \text{ s}$  and final extension at  $60 \text{ }^\circ\text{C}$  for  $30 \text{ min}$ . We then considered 10 markers that presented  $> 20\%$  polymorphism, which were used to amplify all samples within this study (Table 1). The amplified fragments were analysed on a  $3730 \times 1$  gene analyzer (Thermo Fischer Scientific) and examined manually for microsatellite peaks. Allele sizes were determined using GeneMapper 3.2 (Applied Biosystems).

**Estimates of genetic diversity.** For each microsatellite locus, genetic polymorphism was assessed in 48 individuals by calculating the number of alleles ( $N_a$ ), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ), Shannon's diversity index ( $I$ ), and inbreeding coefficient ( $F_{is}$ ) using GenALEX software version 6.5<sup>104</sup>. The polymorphic information content ( $PIC$ ) was calculated as  $PIC = 1 - \sum P_i^2$ , where  $P_i$  is the allele frequency for each SSR marker locus<sup>105,106</sup>. Values of  $PIC$  above 0.5 were considered highly informative, between 0.5 and 0.25 moderately informative, and below 0.25 less informative<sup>107</sup>.

**Population genetic structure and differentiation.** The Bayesian program STRUCTURE v.2.3.4<sup>108</sup> was used to infer the population structure and to assign individual plants to subpopulations. Models with a putative numbers of populations ( $K$ ) from 1–5, imposing ancestral admixture and correlated allele frequencies priors, were considered. Ten independent runs with 50 000 burn-in steps, followed by run lengths of 1 000 000 iterations for each  $K$ , were computed. The number of clusters in the data was estimated using STRUCTURE HARVESTER<sup>109</sup>, which identifies the optimal  $K$  based both on the posterior probability of the data for a given  $K$  and the  $\Delta K$ <sup>110</sup>. To correctly assess the membership proportions ( $q$  values) for clusters identified in STRUCTURE, the results of the replicates at the best fit  $K$  were post-processed using CLUMPP 1.1.2<sup>111</sup>. GenALEX software version 6.5<sup>104</sup> was used to calculate the Nei's genetic distance<sup>112</sup> among individuals. A Principle Coordinate Analysis (PCoA)<sup>113</sup> was performed to detect genetic variations between *W. salutaris* individuals. POPULATION 1.2<sup>114</sup> was used to construct an unrooted neighbour-joining tree with 1000 bootstrap replicates. The Wright's  $F_{ST}$  value was computed to estimate population differentiation<sup>104</sup>. Lower genetic differentiation was considered for  $F_{ST}$  below 0.05, moderate from 0.05 to 0.15 and high genetic differentiation above 0.25<sup>115</sup>.

Received: 27 May 2020; Accepted: 28 September 2020

Published online: 12 November 2020

## References

- Mukherjee, P. W. *Quality Control of Herbal Drugs: An Approach to Evaluation of Botanicals* (Business Horizons Publishers, New York, 2002).
- Bodeker, C. et al. *WHO Global Atlas of Traditional, Complementary and Alternative Medicine* (World Health Organization, Geneva, 2005).
- Bandaranayake, W. M. Quality control, screening, toxicity, and regulation of herbal drugs. In *Modern Phytomedicine, Turning Medicinal Plants into Drugs* (eds Ahmad, I. et al.) 25–57 (Wiley-VCH GmbH & Co. KGaA, Weinheim, 2006).
- World Health Organization (WHO). *Traditional and Complementary Medicine Policy. Management Science for Health*. <https://apps.who.int/medicinedocs/documents/s19582en/s19582en.pdf> (2012). Accessed 13 May 2020.
- Mulauzi, R. B., Ndhkala, A. R., Kulkarni, M. G., Finnie, J. F. & van Staden, J. Antimicrobial properties and phenolic contents of medicinal plants used by the Venda people for conditions related to venereal diseases. *J. Ethnopharmacol.* **135**, 330–337 (2011).
- Wangchuck, P., Keller, P. A., Pyne, S. G., Willis, A. C. & Kamchonwongpaisan, S. Antimalarial alkaloids from Bhutanese traditional medicine plant *Corydalis dubia*. *J. Ethnopharmacol.* **143**, 310–313 (2012).
- Mangal, M., Sagar, P., Singh, H., Raghava, G. P. S. & Agarwal, S. M. NPACT: naturally occurring plant-based anti-cancer compound-activity-target database. *Nucleic Acids Res.* **41**, D1124–D1129 (2013).
- Munodawafa, T., Chagonda, L. S. & Moyo, S. R. Antimicrobial and phytochemical screening of some Zimbabwean medicinal plants. *J. Biol. Active Prod. Nat.* **3**, 323–330 (2013).
- Gechev, T. S. et al. Natural products from resurrection plants: potential for medical applications. *Biotechnol. Adv.* **32**, 1091–1101 (2014).
- Lalisan, J. A., Nuñez, O. M. & Uy, M. M. Brine Shrimp (*Artemia salina*) bioassay of the medicinal plant *Pseudelephantopus spicatus* from Iligan City, Philippines. *Int. Res. J. Biol. Sci.* **3**, 47–50 (2014).
- Du, K. et al. Labdane and clerodane diterpenoids from *Colophospermum mopane*. *J. Nat. Prod.* **78**, 2494–2504 (2015).
- Moyo, B. & Mukanganyama, S. Antiproliferative activity of *T. welwitschii* extract on Jurkat T cells in vitro. *BioMed Res. Int.* **2015**, 817624 (2015).

13. Idris, O. A., Wintola, O. A. & Afolayan, A. J. Phytochemical and antioxidant activities of *Rumex crispus* L. in treatment of gastrointestinal helminths in eastern Cape Province, South Africa. *Asian Pac. J. Trop. Biomed.* **7**, 1071–1078 (2017).
14. Iqbal, J. *et al.* Plant-derived anticancer agents: a green anticancer approach. *Asian Pac. J. Trop. Biomed.* **7**, 1129–1150 (2017).
15. Linder, H. P. The evolution of plant diversity. *Front. Ecol. Evol.* **2**, 38 (2014).
16. Ribeiro, A., Romeiras, M. M., Tavares, J. & Faria, M. T. Ethnobotanical survey in Canhane village, district of Massingir, Mozambique: medicinal plants and traditional knowledge. *J. Ethnobiol. Ethnomed.* **6**, 33 (2010).
17. Moyo, M., Adeyemi, O., Aremu, A. O. & van Staden, J. Medicinal plants: an invaluable, dwindling resource in sub-Saharan Africa. *J. Ethnopharmacol.* **174**, 595–606 (2015).
18. Moura, I., Maquia, I., Rija, A. A., Ribeiro, N. & Ribeiro-Barros, A. I. Biodiversity studies in key species from the African Mopane and Miombo woodlands. In *Genetic Diversity* (ed. Bitz, L.) 92–109 (Intech, Rijeka, 2017).
19. Alaribe, F. N. & Motaung, K. S. C. M. Medicinal plants in tissue engineering and regenerative medicine in the African continent. *Tissue Eng. A* **25**, 827–829 (2019).
20. Burlando, B., Palmero, S. & Cornara, L. Nutritional and medicinal properties of underexploited legume trees from West Africa. *Crit. Rev. Food Sci. Nutr.* **59**, S178–S188 (2019).
21. Botha, J., Witkowski, E. T. F. & Shackleton, C. M. The impact of commercial harvesting on *Warburgia salutaris* ('pepper-bark tree') in Mpumalanga, South Africa. *Biodivers. Conserv.* **13**, 1675–1698 (2004).
22. Kouki, J., Arnold, K. & Martikainen, P. Long-term persistence of aspen—a key host for many threatened species—is endangered in old growth conservation areas in Finland. *J. Nat. Conserv.* **12**, 41–52 (2004).
23. Giam, X., Bradshaw, C. A. J., Tan, H. T. W. & Sodhi, N. S. Future habitat loss and the conservation of plant biodiversity. *Biol. Conserv.* **143**, 1594–1602 (2010).
24. Dluđu, M. N., Dlamini, P. S., Sibandze, G. S., Vilane, V. S. & Dlamini, C. S. The distribution and conservation status of the endangered pepperbark tree *Warburgia salutaris* (Canellaceae) in Swaziland. *Oryx* **51**, 441–454 (2017).
25. Feng, G., Mao, L., Benito, B. M., Swenson, N. G. & Svenning, J.-C. Historical anthropogenic footprints in the distribution of threatened plants in China. *Biol. Conserv.* **210**, 3–8 (2017).
26. Pramanik, M., Paudel, U., Mondal, B., Chakraborti, S. & Deb, P. Predicting climate change impacts on the distribution of the threatened *Garcinia indica* in the Western Ghats, India. *Clim. Risk Manag.* **19**, 94–105 (2018).
27. Dudley, A., Butt, N., Auld, T. D. & Gallagher, R. V. Using traits to assess threatened plant species response to climate change. *Biodivers. Conserv.* **28**, 1905–1919 (2019).
28. Krog, M., Falcão, M. & Olsen, C. S. *Medicinal plant markets and trade in Maputo, Mozambique* (Danish Center for Forest, Landscape and Planning, KVL, 2006).
29. Mukamuri, B. B. & Kozanayi, W. Commercialization and institutional arrangements involving tree species harvested for bark by smallholder farmers in Zimbabwe. In *Bark Use, Management and Commerce in Africa* (eds Cunningham, A. B. *et al.*) 247–254 (The New York Botanical Garden Press, New York, 2014).
30. Cunningham, A. B. *African Medicinal Plants: Setting Priorities at the Interface Between Conservation and Primary Health Care* (UNESCO, 1993).
31. Maquia, I. *et al.* Diversification of African tree legumes in Miombo-Mopane woodlands. *Plants* **6**, 182 (2019).
32. Ribeiro-Barros, A. I. *et al.* Actinorhizal trees and shrubs from Africa: distribution, conservation and uses. *Antonie Van Leeuwenhoek* **112**, 31–46 (2019).
33. Scribner, K. T., Meffe, G. F. & Groom, M. J. Conservation genetics: the use and importance of genetic information. In *Principles of Conservation Biology* (eds Groom, M. J. *et al.*) 375–415 (Sinauer Associates, Inc. Publishers, Sunderland, 2006).
34. Baldauf, C., Ciampi-Guillard, M., dos Santos, F. A. M., de Souza, A. P. & Sebbenn, A. M. Tapping latex and alleles? The impacts of latex and bark harvesting on the genetic diversity of *Himatanthus drasticus* (Apocynaceae). *For. Ecol. Manag.* **310**, 434–441 (2013).
35. de Sousa, V. A. *et al.* Genetic diversity and biogeographic determinants of population structure in *Araucaria angustifolia* (Bert.) O. Ktze. *Conserv. Genet.* **21**, 217–229 (2020).
36. Hu, J. *et al.* Comparative analysis of genetic diversity in sacred lotus (*Nelumbo nucifera* Gaertn.) using AFLP and SSR markers. *Mol. Biol. Rep.* **39**, 3637–3647 (2012).
37. Maquia, I. *et al.* Genetic diversity of *Brachystegia boehmii* Taub. and *Burkea africana* Hook. f. across a fire gradient in Niassa National Reserve, northern Mozambique. *Biochem. Syst. Ecol.* **48**, 238–247 (2013).
38. Bessega, C., Cony, M., Saidman, B. O., Aguilo, A. & Vilardi, J. C. Genetic diversity and differentiation among provenances of *Prosopis flexuosa* DC (Leguminosae) in progeny trial: implication for arid land restoration. *For. Ecol. Manag.* **44**, 59–68 (2019).
39. Braun, M., Esposito, T., Huettel, B. & Harand, A. P. Development and characterization of eleven microsatellite loci for the tropical understory tree *Paypayrola blanchetiana* Tul. (Violaceae). *Mol. Biol. Rep.* **46**, 529–2532 (2019).
40. Zane, L., Bargelloni, L. & Patarnello, T. Strategies for microsatellite isolation: a review. *Mol. Ecol.* **11**, 1–16 (2002).
41. Jackson, S. A., Iwata, A., Lee, S.-H., Schmutz, J. & Shoemaker, R. Sequencing crop genome: approaches and applications. *New Pathol.* **191**, 915–925 (2011).
42. Camacho, L. M. D. *et al.* Development, characterization and cross-amplification of microsatellite markers for *Chrysoleaena obovata*, an important Asteraceae from Brazilian Cerrado. *J. Genet.* **96**, 47–53 (2017).
43. Dantas, L. G., Alencar, L., Huettel, B. & Pedrosa Harand, A. Development of ten microsatellite markers for *Alibertia edulis* (Rubiaceae), a Brazilian savanna tree species. *Mol. Biol. Rep.* **46**, 4593 (2019).
44. Eom, S. H. & Na, J.-K. Leaf transcriptome data of two tropical medicinal plants: *Sterculia lanceolata* and *Clausena excavate*. *Data Brief* **25**, 104297 (2019).
45. Mercati, F. *et al.* Transcriptome analysis and codominant markers development in caper, a drought tolerant orphan crop with medicinal value. *Sci. Rep.* **9**, 10411 (2019).
46. Kyriakidou, M., Tai, H. H., Anglin, N. L., Ellis, D. & Strömvik, M. V. Current strategies of polyploid plant genome sequence assembly. *Front. Plant Sci.* **9**, 1660 (2018).
47. van Wyk, B.-E. & Wink, W. *Medicinal Plants of the World* (Briza Publications, Pretoria, 2004).
48. Senkoro, A. M., Shackleton, C. M., Voeks, R. A. & Ribeiro, A. I. Uses, knowledge, and management of the threatened pepperbark tree (*Warburgia salutaris*) in southern Mozambique. *Econ. Bot.* **73**, 304–324 (2019).
49. Hannweg, K., Sippel, A., Hofmeyr, M., Swemmer, L. & Froneman, W. Strategies for the conservation of *Warburgia salutaris* (family: Canellaceae), a red data list species - development of propagation methods. *Acta Hort.* **1125**, 33–40 (2016).
50. Dlamini, T. S. & Dlamini, G. M. Swaziland in *Southern African Plant Red Data Lists* (ed. Golding, J. S.) 121–134 (National Botanic Institute, 2002).
51. Maroyi, A. Community attitudes towards the reintroduction programme for the endangered pepperbark tree *Warburgia salutaris*: implications for plant conservation in south-east Zimbabwe. *Oryx* **46**, 213–218 (2012).
52. Veeman, M. M., Cocks, M. L., Muwonge, F., Chonge, S. K. & Campbell, B. M. Markets for three bark products in Zimbabwe: a case study of markets for *Adansonia digitata*, *Berchemia discolor* and *Warburgia salutaris*. In *Bark Use, Management and Commerce in Africa* (eds Cunningham, A. B. *et al.*) 227–245 (The New York Botanical Garden Press, New York, 2014).
53. Veeman, T. S., Cunningham, A. B. & Kozanayi, W. The economics of production of rare medicinal species introduced in southwestern Zimbabwe: *Warburgia salutaris*. In *Bark Use, Management and Commerce in Africa* (eds Cunningham, A. B. *et al.*) 179–188 (The New York Botanical Garden Press, New York, 2014).

54. Msekandiana, G. & Mlangeni, E. Malawi in *Southern African Plant Red Data Lists* (ed. Golding, J. S.) 135–156 (National Botanic Institute, 2002).
55. Jansen, P. C. M. & Mendes, O. *Plantas medicinais: seu uso tradicional em Moçambique* (Imprensa do Partido, 1990).
56. Coates-Palgrave, M. *Trees of Southern Africa* (Struik Publishers, Cape Town, 2002).
57. Hilton-Taylor, C., Scott-Shaw, R., Burrows, J. & Hahn N. *Warburgia salutaris*. *The IUCN Red List of Threatened Species 1998*. <https://doi.org/10.2305/IUCN.UK.1998.RLTS.T30364A9541142.en>. Accessed 13 May 2020.
58. Izidine, S. & Bandeira, S. O. Mozambique in *Southern African Plant Red Data Lists* (ed. Golding, J. S.) 43–60 (National Botanic Institute, 2002).
59. Mapaura, A. & Timberlake, J. R. Zimbabwe in *Southern African Plant Red Data Lists* (ed. Golding, J. S.) 158–182 (National Botanic Institute, 2002).
60. Maroyi, A. Ethnobotanical study of two threatened medicinal plants in Zimbabwe. *Int. J. Biodivers. Sci. Ecosyst. Serv. Manag.* **4**, 1–6 (2008).
61. William, V. L., Witkowski, E. T. F. & Balkwill, K. Relationship between bark thickness and diameter at the breast height for six tree species used medicinally in South Africa. *S. Afr. J. Bot.* **73**, 449–465 (2007).
62. Muchugi, A. *et al.* Genetic structuring of important medicinal species of genus *Warburgia* as revealed by AFLP analysis. *Tree Genet. Genomes* **4**, 787–795 (2008).
63. Hannweg, K., Hofmeyer, M. & Grove, T. The pepperbark initiative: are we closer to efficiently propagating *Warburgia salutaris*? [https://www.sanparks.org/assets/docs/conservation/scientific\\_new/savanna/ssnm2015/the-pepperbark-initiative-are-we-any-closer-to-efficiently-propagating-warburgia-salutaris.pdf](https://www.sanparks.org/assets/docs/conservation/scientific_new/savanna/ssnm2015/the-pepperbark-initiative-are-we-any-closer-to-efficiently-propagating-warburgia-salutaris.pdf) (2015). Accessed 13 May 2020.
64. Tonguç, M. & Griffiths, P. D. Genetic relationships of *Brassica* vegetables determined using database derived simple sequence repeats. *Euphytica* **137**, 193–201 (2004).
65. Zhao, X., Ma, Y., Sun, W., Wen, X. & Milne, R. High genetic diversity and low differentiation of *Michelia coriacea* (Magnoliaceae), a critically endangered endemic in Southeast Yunnan, China. *Int. J. Mol. Sci.* **13**, 4396–4411 (2012).
66. Paliwal, R. *et al.* Development of genomic simple sequence repeats (g-SSR) markers in *Tinospora cordifolia* and their application in diversity analyses. *Plant Genes* **5**, 118–125 (2016).
67. Tian, Z., Zhang, F., Liu, H., Gao, Q. & Chen, S. Development of SSR markers for a Tibetan medicinal plant, *Lancea tibetica* (Phrymaceae), based on RAD sequencing. *Appl. Plant Sci.* **4**, 1600076 (2016).
68. Li, Z.-Z., Gichira, A. W., Wang, Q.-F. & Che, J. M. Genetic diversity and population structure of the endangered basal angiosperm *Brasenia schreberi* (Cabombaceae) in China. *PeerJ* **6**, e5296 (2018).
69. Li, X. *et al.* De novo transcriptome assembly and population genetic analyses for an endangered Chinese endemic *Acer miaotaiense* (Aceraceae). *Genes* **9**, 378 (2018).
70. Martins, S. *et al.* Western European wild landraces hazelnuts evaluated by SSR markers. *Plant Mol. Biol. Rep.* **33**, 1712–1720 (2015).
71. Gacheri, N., Wanjala, B. W., Jamnadass, R. & Muchugi, A. Analysis of the impact of domestication of *Warburgia ugandensis* (Sprague) on its genetic diversity based on amplified fragment length polymorphism. *Afr. J. Biotechnol.* **15**, 1673–1680 (2016).
72. Soltis, D. E., Soltis, P. S., Bennett, M. D. & Leitch, I. J. Evolution of genome size in the angiosperms. *Am. J. Bot.* **90**, 1596–1603 (2003).
73. Leitch, I. J. & Hanson, L. DNA C-values in seven families fill phylogenetic gaps in the basal angiosperms. *Bot. J. Linn. Soc.* **140**, 175–179 (2002).
74. Cueva-Agila, A. *et al.* Genetic characterization of fragmented populations of *Cinchona officinalis* L. (Rubiaceae), a threatened tree of the northern Andean cloud forests. *Tree Genet Genomes* **15**, 81 (2019).
75. Yang, S., Xue, S., Kang, W., Qian, Z. & Yi, Z. Genetic diversity and population structure of *Miscanthus lutarioriparius*, an endemic plant of China. *PLoS ONE* **14**, e0211471 (2019).
76. George, S., Sharma, J. & Yadon, V. L. Genetic diversity of the endangered and narrow endemic *Piperia yadonii* (Orchidaceae) assessed with ISSR polymorphisms. *Am. J. Bot.* **96**, 2022–2030 (2009).
77. Muriira, N. G., Muchugi, A., Yu, A., Xu, J. & Li, A. Genetic differentiation and strong population structure in *Calotropis* plants. *Sci. Rep.* **8**, 7832 (2018).
78. Vardareli, N., Doğaroglu, T., Doğaç, E., Taşkın, V. & Taşkı, B. G. Genetic characterization of tertiary relict endemic *Phoenix theophrasti* population in Turkey and Phylogenetic relations of the species with other palm species revealed by SSR markers. *Plant Syst. Evol.* **305**, 415–429 (2019).
79. Sun, R., Lin, F., Huang, P. & Zheng, Y. Moderate genetic diversity and genetic differentiation in the relict tree *Liquidambar formosana* Hance revealed by genic simple sequence repeat markers. *Front. Plant Sci.* **7**, 1411 (2016).
80. Paul, J., Budd, C. & Freeland, J. R. Conservation genetics of an endangered orchid in eastern Canada. *Conserv. Genet.* **1**, 195–204 (2013).
81. Yuan, J.-H., Cheng, F.-Y. & Zhou, S.-L. The phylogeographic structure and conservation genetics of the endangered tree peony, *Paeonia rockii* (Paeoniaceae), inferred from chloroplast gene sequences. *Conserv. Genet.* **12**, 1539–1549 (2011).
82. Matthew, W. S., van Wyk, A. E., van Rooyen, N. & Botha, G. A. Vegetation of the Tembe Elephant Park, Maputaland, South Africa. *S. Afr. J. Bot.* **67**, 573–594 (2001).
83. Joseph, G. S. *et al.* Large termitaria act as refugia for tall trees, deadwood and cavity-using birds in a miombo woodland. *Landsc. Ecol.* **26**, 439–448 (2011).
84. Ge, X.-J., Zhang, L.-B., Yuan, Y.-M., Hao, G. & Chiang, T.-Y. Strong genetic differentiation of the East Himalayan *Megacodon stylophorus* (Gentianaceae) detected by Inter-Simple Sequence Repeats (ISSR). *Biodivers. Conserv.* **14**, 849–861 (2005).
85. Zheng, W., Wang, L., Meng, L. & Liu, J. Genetic variation in the endangered *Anisodus tanguticus* (Solanaceae), an alpine perennial endemic to the Qinghai-Tibetan Plateau. *Genetica* **132**, 123–129 (2008).
86. Halafo, J. *Estudo da Planta Warburgia salutaris na floresta de Licuati: estado de conservação e utilização pelas comunidades locais*. (Universidade Eduardo Mondlane, 1996).
87. Governo do Distrito de Matutuine (GDM). *Plano estratégico do desenvolvimento do distrito de Matutuine (2009–2013)* (Governo do Distrito de Matutuine, 2008).
88. Ministério de Administração Estatal (MAE). *Perfil do distrito de Namaacha, província de Maputo. Portal do Governo de Moçambique*. <https://www.portaldogoverno.gov.mz/Informacao/distritos/> (2005). Accessed 13 May 2020.
89. Mbambezeli, G. Mbambezeli, G., Warburgia salutaris (Bertol.f.) Chiow. <https://pza.sanbi.org/> (2004).
90. Pan, L. *et al.* Genetic diversity among germplasms of Pitaya based on SSR markers. *Sci. Hortic.* **225**, 171–176 (2017).
91. Ministério de Administração Estatal (MAE). *Perfil do distrito de Matutuine, província de Maputo. Portal do Governo de Moçambique*. <https://www.portaldogoverno.gov.mz/Informacao/distritos/> (2005). Accessed 13 May 2020.
92. Kirkwood, D. *Southeastern Africa: Mozambique, Swaziland, and So. Tropical and Subtropical Moist Broadleaf Forests*. <https://www.worldwildlife.org/ecoregions/at0119> (2014). Accessed 13 May 2020.
93. Burrows, J., Burrows, S., Lotter, M. & Schmidt, E. *Trees and Shrubs: Mozambique* (Print Matters Heritage, Cape Town, 2018).
94. Izidine, S. A. *Licuati Forestry Reserve, Mozambique: Flora, Utilization and Conservation* (University of Pretoria, Pretoria, 2003).
95. Moll, E. J. Terrestrial plant ecology. In *Studies on the Ecology of Maputaland* (eds Bruton, M. N. & Cooper, K. H.) 52–68 (Cape and Transvaal Printers (Pty) Ltd, Cape Town, 1980).

96. Ministerio de Turismo (MITUR). *Proposta de demarcação do Corredor de Futi*. [https://conservationcorridor.org/cpb/Transfrontier\\_Conservation\\_Areas\\_2002.pdf](https://conservationcorridor.org/cpb/Transfrontier_Conservation_Areas_2002.pdf) (2002). Accessed 13 May 2020.
97. van Rooyen, N. Die plantegroei van die Roodeplaatdam-natuurreservaat II. Die plantgemeenskappe. *South-Africa Tydskrif Plant* **2**, 115–125 (1983).
98. Instituto Nacional de Estatística (INE). *Estatísticas do distrito de Namaacha* (Instituto Nacional de Estatística, 2013).
99. Instituto Nacional de Estatística (INE). *Estatísticas do distrito de Matutuine* (Instituto Nacional de Estatística, 2013).
100. Loureiro, J., Rodriguez, E., Doležel, J. & Santos, C. Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Ann. Bot.* **100**, 875–888 (2007).
101. Guilengue, N., Alves, S., Talhinhos, P. & Neves-Martins, J. Genetic and genomic diversity in a Tarwi (*Lupinus mutabilis* Sweet) germplasm collection and adaptability to Mediterranean climatic conditions. *Agronomy* **10**, 2 (2020).
102. Doležel, J., Sgorbati, S. & Lucretti, S. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiol. Plant.* **85**, 625–631 (1992).
103. Doležel, J. & Bartoš, J. Plant DNA flow cytometry and estimation of nuclear genome size. *Ann. Bot.* **95**, 99–110 (2005).
104. Peakall, R. & Smouse, P. E. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **6**, 288–295 (2006).
105. Benor, S., Zhang, M., Wang, Z. & Zhang, H. Assessment of genetic variation in tomato (*Solanum lycopersicum* L.) inbred lines using SSR molecular markers. *J. Genet. Genomics* **35**, 373–379 (2008).
106. Chesnokov, Y. V. & Artemyeva, A. M. Evaluation of the measure of polymorphism information of genetic diversity. *Agric. Biol.* **50**, 571–578 (2015).
107. Botstein, D., White, R. L., Skolnick, M. S. & Davis, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **62**, 314–331 (1980).
108. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
109. Earl, D. A. & vonHoldt, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
110. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
111. Jakobsson, M. & Rosenberg, N. A. CLUMPP: a cluster matching and permutation program for dealing with labels switching and multimodality in analysis of population structure. *Bioinformatics* **15**, 1801–1806 (2007).
112. Nei, M. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583–590 (1978).
113. Gower, J. C. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**, 325–338 (1966).
114. Langella, O. *Populations*, 1.2.30 (CNRS UPR9034, 1999).
115. Wright, S. *Evolution and the Genetics of Populations: Variability Within and Among Natural Populations* Vol. 4 (The University of Chicago Press, Chicago, 1978).

## Acknowledgements

We are grateful to the Organization for Women in Science for the Developing World (OWSD), Swedish International Development Cooperation Agency (SIDA), Russell E. Train Education for Nature Program, World Wildlife Fund (Agreement #SS20), Camões, I.P. (Portugal), and the Portuguese Science and Technology Foundation through the research units UIDB/00239/2020 (CEF) and UID/04129/2020 (LEAF), and the contribution to the International Rice Research Institute. We would also like to thank our field guides Antonio Tembe, Ernesto Bié, Filimone Cossa, Luis Cossa, Massale Tembe, Ricardo Mundlovu and Sifiso Masuko for their support during sample collection. We are also grateful to our colleagues, Domingos Maguengue, Ivete Maquia, and Ana Gomes during the initial phase of the work and to José Alfredo Amanze for providing the study area map.

## Author contributions

A.M.S., C.M.S., R.A.V., and A.I.R.B. conceived the work and the experimental design. A.M.S., C.M.S., and R.A.V. performed the field data survey and sample collection. A.M.S., P.B.S., F.S., and P.T. performed the laboratorial analysis. P.T. performed the flow cytometry data analysis. A.M.S., F.S., I.M. and A.I.R.B. performed the microsatellite data analysis. A.M.S., I.M., and A.I.R.B. wrote the first draft and final version of the manuscript, which has been thoroughly reviewed by all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-76654-6>.

**Correspondence** and requests for materials should be addressed to I.M. or A.I.R.-B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020