Article (refereed) - postprint

1 **Plio-Pleistocene history and phylogeography of *Acacia senegal* in dry**

2 **woodlands and savannahs of sub-Saharan tropical Africa: evidence of early**

3 **colonisation and recent range expansion**

4

5 **DW Odee[1,2*], A Telford[1], J Wilson[1], A Gaye[3] and S Cavers[1]**

6

7

8 [1] Centre for Ecology & Hydrology, Bush Estate, Penicuik, Midlothian EH26 0QB, Scotland,

9 UK

10 [2] Kenya Forestry Research Institute, P.O Box 20412-00200, Nairobi, Kenya

11 [3] Institut Sénégalais de Recherches Agricoles (ISRA), B.P. 3120, Dakar, Senegal

12

13

14 **Address for correspondence:** *Centre for Ecology & Hydrology, Bush Estate, Penicuik,

15 Midlothian EH26 0QB, Scotland, UK

16 Tel. 0131 445 4343, Fax. 0131 445 3943

17 * Email daee@ceh.ac.uk

18

19 **Running head:** Phylogeography of *Acacia senegal*

20 **Total word count:** 6 485

21

Abstract

The gum arabic tree *(Acacia senegal)* is an arid-adapted, morphologically diverse species (varieties *senegal*, *kerensis*, *rostrata* and *leiorhachis*) widespread in the extensive but relatively little-studied dry woodland and savannah biomes of sub-Saharan tropical Africa. We used variation in nuclear ribosomal DNA (nrDNA) internal transcribed spacer (ITS) sequences and chloroplast DNA markers (polymerase chain reaction-restriction fragment length polymorphism [PCR-RFLP] and chloroplast microsatellites [cpSSR]) to study the phylogeography of the species with 293 individuals from 66 populations sampled across its natural range. The predominant pattern was a phylogeographic distinction between West and Central African, and East and Southern African haplotypes. Phylogenetic analysis of ITS data indicated a more recent origin for the clade including West and Central African haplotypes, suggesting range expansion in the Sudano-Sahelian region, probably as recently as the Holocene humid period. Variety *leiorhachis* formed a single clade while the other three were mixed and some evidence for hybridization between *A. senegal* and other *Acacia* species was present. Chloroplast DNA data showed high regional and rangewide haplotypic diversity ($h_{T(cpSSR)}$ = 0.903 - 0.948) and population differentiation ($G_{ST(RFLP)}$ = 0.700 - 0.782) with a phylogeographic pattern that indicated extensive historical gene flow via seed dispersal. Haplotypes were not restricted to any of the four varieties, but showed a significant geographic structuring of genetic variation ($G_{ST(cpSSR)}$ = 0.392; R_{ST} = 0.673; R_{ST} > R_{ST} (permuted)), with the major division separating East and Southern Africa populations from those in West and Central Africa. In conjunction with paleobotanical evidence, our data suggests dispersal to West Africa, and across to the Arabian Peninsula and Indian subcontinent from source populations located in the East African region during climate oscillations of the Plio-Pleistocene.

46

47 **Keywords:** Aridity, gum arabic, hybridization, long-distance dispersal, phylogeny, refugia

48

49

Introduction

Modern dry woodland and savannah biomes, dominated by grasses and woody species, are key ecosystems in sub-Saharan tropical Africa (Plana, 2004). With a species composition and diversity that is driven by low rainfall and pronounced dry seasons (Jacobs, 2004), they evolved during the Miocene (23-5 Mya), a period associated with drastic lowering of global temperatures and increasing aridity that saw the replacement of once expansive extensive lowland rainforests with savannah woodland (Plana, 2004). From the Plio-Pleistocene onwards (from ~ 5 Mya), the species that occupy these biomes have experienced cyclical range expansions and contractions due to climatic oscillations (Maley, 2001; Plana, 2004), restricted by aridification on one hand and moist forest expansion on the other. Thus the genetic structure of widely distributed dryland tree species is likely to reflect the influence of these processes.

The interplay of major geomorphological features with climate oscillations has also played a key role in shaping the genetic patterns and phylogeography of flora and fauna since the end of the Pliocene (~ 3.5 Mya, deMenocal, 1995). For example, contraction and expansion of the Sahara desert in Central and West Africa, and interaction of climate with the Great Rift Valley and the elevated topographies of East and Southern Africa produced wide altitudinal ranges forming a complex mosaic of landscapes and localised climate regimes that functioned as refugia during extreme climate conditions (Plana, 2004; Lorenzen *et al*., 2010).

Previous phylogeographic studies of tropical African tree species have largely focussed on moist forests (montane, lowland or rainforests) and have typically interpreted results with respect to Pleistocene refugia (e.g. *Hagenia abyssinica*, Ayele *et al*., 2009;

4

74   *Milicia excelsa*, Daïnou *et al.*, 2010; *Irvingia gabonensis*, Lowe *et al.*, 2010). It has been

75   suggested that previously continuous rainforests fragmented into refugia within mountainous

76   areas and became separated by the expansion of savannah vegetation during glacial maxima

77   associated with cooler and drier conditions (Plana, 2004). Dry woodlands and savannahs are

78   likely to have experienced the same processes in reverse, and fragmented dryland habitats

79   may have acted as refugia for arid-adapted species during warmer and wetter periods (Plana,

80   2004). Relics of xeric vegetation or savannah enclaves persisting today in the Central and

81   West African rain forest since the last glacial maximum (LGM) may represent such refuges

82   (Maley, 2001).

83

84   A few savannah-type or dry woodland tree species have been studied, mainly in the

85   Brazilian cerrado (e.g. *Caryocar brasilensis*, Collevatti *et al.*, 2003; *Hymenaea stigonocarpa*,

86   Ramos *et al.*, 2007; *Astronium urundeuva*, Caetano *et al.*, 2008) and others in the seasonally

87   dry tropical forests (SDTFs, Pennington *et al.*, 2004; Pennington *et al.*, 2009). Species in the

88   latter show patterns of diversity, endemism and phylogeny that indicate historically stable and

89   dispersal-limited systems, caused partly by the widespread patchy distribution of this biome

90   and its persistence over evolutionary timescales regarded as Pleistocene refugia (Pennington

91   *et al.*, 2004; Collevatti *et al.*, 2003; Caetano *et al.*, 2008;). Although SDTFs share some of the

92   ecological characteristics of dry woodlands and savannahs of sub-Saharan Africa (Pennington

93   *et al.*, 2009; Lock, 2006), direct comparisons are limited by phytogeographical differences. In

94   Africa, there have been three major phylogeographic studies of African tree species found in

95   these biomes. A study of *Acacia nilotica* populations showed genetic differences that broadly

96   matched subspecific designations and a phylogeographic separation of North and West Africa

97   from East and Southern Africa (Wardill *et al.*, 2005). The other two studies, of the baobab

98   tree (*Adansonia digitata* L., Malvaceae; Pock Tsy *et al*., 2009) and the shea tree (*Vitellaria*

99   *paradoxa* C.F.Gaertn, Sapotaceae; Allal *et al*., 2011) showed strong phylogeographic

100   structure, distinguishing Eastern and Western populations within the Sudano-Sahelian region.

101   However, the ecological distribution of the baobab tree only partially matches that of the dry

102   woodlands and savannahs, and the shea tree has a limited distribution in the Sudano-Sahelian

103   zone. In contrast to plants, there are numerous studies of vertebrates that show development

104   of intraspecific differentiation that coincides with formation of the woodland and savannah

105   biomes in Africa (Jacobs, 2004), e.g. plains zebra, *Equus quagga* (Lorenzen *et al*., 2008) and

106   common eland antelope, *Taurotragus oryx* (Lorenzen *et al*., 2010).

107

108   Here, we study the phylogenetics and phylogeography of *Acacia senegal* (L.) Willd.

109   (Leguminosae, Mimosoideae) commonly known as the 'gum arabic' tree. Although renowned

110   for its commercial gum, which has been traded for centuries (Fagg and Allison 2004), *Acacia*

111   *senegal* also provides other key ecosystem services such as sand stabilization, shade, fodder

112   and forage for livestock and wildlife as well as N_2-fixation, which enhances soil fertility and

113   sustains food crop production in the gum arabic agroforestry systems. *Acacia senegal* is

114   drought tolerant and can occur in extremely dry habitats, but also has a wide climatic

115   (temperature and rainfall) and altitudinal range (Supplementary Table S1). It has an inter-

116   continental distribution extending from Africa eastwards to the Arabian Peninsula, Pakistan

117   and India (Figure 2; Fagg and Allison, 2004). Reports from common garden experiments have

118   shown wide variations in survival, physiology, growth and gum production traits among

119   provenances, reflecting the adaptive variation that exists across its natural distribution range

120   (Sprent *et al*., 2010).

121   According to a recent phylogeny of African acacias, the genus *Acacia* (Syn.

122 *Senegalia*) is estimated to have inhabited the open habitats (woodlands and savannahs) of

123 Africa since the Miocene epoch (~ 21 Mya, Bouchenak-Khelladi *et al*., 2010). This also

124 marks the period of rapid diversification within the Leguminosae family (e.g see Richardson

125 *et al*., 2001; Lavin *et al*., 2005). As important drivers of the evolutionary processes of

126 diversification, hybridization and introgression were likely to have contributed to the

127 adaptation of the African acacias to the new, sometimes extreme environments experienced in

128 the sub-Saharan regions during the Plio-Pleistocene (Maley, 2001; Jacobs, 2004; Plana, 2004;

129 Bouchenak-Khelladi *et al*., 2010). *Acacia senegal* belongs to a group of more than 20 closely

130 related species referred to as the *Acacia senegal* complex characterised by spicate

131 inflorescences and prickles that are mostly in threes, the central one hooked downwards and

132 the laterals hooked upwards (Ross 1979; Fagg and Allison, 2004). It forms a natural hybrid

133 (*A. laeta* R.Br. ex Benth.) with *A. mellifera* (Vahl) Benth. (El Amin, 1976). We focus on

134 *Acacia senegal* documented to have four putative varieties (*A. senegal* var. *senegal*, *A.*

135 *senegal* var. *kerensis* Schweinf, *A. senegal* var. *rostrata* Brenan and *A. senegal* var.

136 *leiorhachis* Brenan) whose delimitations rely largely on growth and morphological characters

137 (Ross 1979; Fagg and Allison, 2004). In Eastern Africa, regarded as the centre of diversity for

138 the species, the intraspecific delimitation can be less resolved, especially among varieties

139 *senegal*, *kerensis* and *rostrata* due to variation and overlaps in morphological characters

140 (Ross 1979; Ross 1981; Fagg and Allison, 2004, Table 1). However, var. *leiorhachis* has

141 some unique growth, pod and phenological features that separate it from the other varieties.

142 The key growth and morphological characters often used to distinguish the varieties include

143 growth form, type of stem or trunk texture, inflorescence axis pubescence and shape of pods

144 (Table 1).

146   *Acacia senegal* is mainly insect pollinated, and predominantly outcrossing. Flowering

147   occurs annually (or biannually in some regions) followed by a good seed crop, which is

148   dispersed by wind and to some extent animals, especially ungulates (Fagg and Allison, 2004).

149   Cytological studies on *A. senegal* var. *senegal* indicated that it is a diploid, 2n = 26 (Bukhari,

150   1997). The species is reproductively active within 2-4 years of establishment. Previous

151   genetic studies of Kenyan populations using both chloroplast and nuclear microsatellites

152   found high genetic diversity within populations and low genetic differentiation among

153   populations, indicating extensive gene flow (Omondi *et al*., 2010). These attributes

154   demonstrate effective pollen and seed dispersal mechanisms with potential for long-distance

155   dispersal.

156

157   Given the historical range changes that *A. senegal* has probably experienced, the complex

158   landscape it occupies and its effective dispersal ecology, we predict the following:

159   1. The divergence events in *A. senegal* will reflect those of major climatic shifts because

160   sub-Saharan woodlands and savannahs are ancient and have been subject to

161   fluctuations in size and distribution out of phase with those of the moist forests during

162   climate oscillations of the Plio-Pleistocene.

163   2. As regional differentiation has been seen in previous studies of dryland species (e.g.

164   baobab, shea) we predict similar differentiation for *A. senegal*. Within regions we

165   expect low genetic structure due to life history characteristics that predispose the

166   species to effective dispersal.

167   3. There will be high genetic diversity and presence of basal groups in the regions of

168   Eastern and Southern Africa due to extant taxonomic diversity and a complex mosaic

169   of landscapes (lowlands to highlands covered by forests, woodlands and savannahs)

170   and climate regimes predating the Pleistocene.

171

172   To address these hypotheses, we use variation in the internal transcribed spacer (ITS) of the

173   nuclear ribosomal DNA (nrDNA) and the cpDNA (*trn*H-*psb*A and cpSSR) to (i) assess

174   phylogenetic relationships among the *A. senegal* varieties (ii) test for phylogeographic

175   structuring within *A. senegal* (iii) examine evolutionary history of *Acacia. senegal* in terms of

176   colonisation of the dryland habitats from its origins and hypothesised refugia in the light of

177   environmental and habitat distributional changes since the original diversification that

178   produced the species in the Plio-Pleistocene.

179

180   **Materials and Methods**

181   Collection and DNA extraction

182   Samples were collected from a total of 293 *Acacia senegal* individuals from wild populations,

183   experimental trials and herbarium specimens representing 66 sites across its natural range in

184   sub-Saharan Africa, the Arabian Peninsula, Pakistan and India (Table 2). Field samples were

185   collected from well known and documented sources; samples of doubtful or unverified

186   intraspecific affiliations are indicated (Table 2). Fresh leaf or stem tissue samples were dried

187   on silica gel in ziplock bags. For each sample, roughly 1 cm² of dried tissue was ground to a

188   fine powder using a Retsch Tissuelyser. DNA extraction was carried out using QIAGEN

189   DNeasy 96 Plant kits following the manufacturer's protocol. Extracted DNA was assessed for

190   quality and concentration on a 1 % agarose gel before storage at -20 °C.

191

192   PCR and sequencing

193   We used data from both nuclear and organelle genomes to test for phylogeographic patterns.

194   The internal transcribed spacer (ITS) of the nuclear ribosomal DNA (nrDNA) was used for

195   phylogenetic reconstruction. Besides its utility in phylogenetic studies in angiosperms, the

196   high sequence variation found among conspecifics and allopatric populations also makes it

197   suitable for studying intraspecific relationships (Baldwin *et al.*, 1995). The ITS region, alone

198   or concatenated with other spacer regions, has been successfully used in previous studies to

199   resolve phylogenetic and taxonomic issues among other *Acacia* taxa (e.g. Ariati *et al.*, 2006;

200   Brown *et al.*, 2008). Chloroplast (cp) DNA was also used to analyse geographical patterns of

201   diversity and population genetic structure across the distribution range of the species. The

202   cpDNA is maternally inherited in angiosperms and has a slow rate of evolution (compared to

203   nrDNA), making it ideal for studying historical patterns of gene flow, colonisation and

204   migration events (Cavers *et al.*, 2003).

205

206   For ITS sequence analysis, a subsample of 65 individuals representing the four different

207   recognised varieties and the full extent of the geographic range were selected (Table 1). The

208   5.8S subunit and flanking spacers ITS 1 and ITS 2 were amplified together with part of the

209   18S and 26S gene regions as a single fragment using primers 17SE and 26SE and PCR

210   conditions as detailed by Sun *et al*. (1994). Direct sequencing was carried out at the NERC

211   Biomolecular analysis facility at the University of Edinburgh, United Kingdom.

212

213   Three chloroplast regions were selected for screening of the whole collection after

214   preliminary screening in a subset of 24 geographically and taxonomically representative

215   samples. These were: the *trn*H-*psb*A spacer (Shaw *et al*., 2005), restricted with *Dra*I (selected

216   after testing *Apa*I, *Alu*I, *BamH*I, *Dra*I, *Mse*I and *Rsa*I) and universal microsatellite primer

217 pairs ccmp5 and 10 (Weising and Gardner, 1999, after screening primer pairs ccmp 1-10).

218 The fragment *trn*H-*psb*A was amplified in 25 μL reactions containing: 2 μL DNA [~ 20 ng of

219 genomic template DNA], 2.5 μL of 10X buffer (New England Biolabs), 0.5 μL dNTPs

220 (Promega), 0.5 μL each primer (MWG Biotech), 0.2 μL (0.5 U) Taq (New England Biolabs),

221 0.4 μL BSA (Sigma) and 13.4 μL water. PCR was carried out on a Thermo MBS thermal

222 cycler with the following protocol: 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 45 °C for 30

223 s, 72 °C for 1 min and finally 72 °C for 10 min. Amplicons were digested with *Dra*I and

224 restriction fragment patterns were visualised as described by Cavers *et al*. (2003).

225

226   The two cpSSR loci were amplified in 25 μL reactions containing: 0.5 μL DNA [~ 20

227 ng of genomic template DNA], 2 μL of 10X PCR buffer (Promega), 1.6 μL dNTPs

228 (Promega), 0.4 μL each primer (MWG Biotech), 0.2 μL (0.5 U) Taq (New England Biolabs),

229 0.4 μL BSA and 19.5 μL water. The PCR products were electrophoresed, band sizes

230 determined and characterised for cpDNA haplotype following the procedure described by

231 Omondi *et al*. (2010).

232

233 Statistical analysis

234 ITS sequences were edited and assembled in CodonCode Aligner 3.5.7 (CodonCode

235 Corporation, Dedham, MA, USA). Sequences were lodged with GenBank under the accession

236 numbers HQ605042-HQ605077 (Supplementary Table S2, Supporting information).

237 Insertions/deletions (indels) were coded as presence or absence. At ITS loci, multiple copies,

238 paralogues or pseudogenes can confound phylogenetic inference (Álvarez and Wendel, 2003).

239 Therefore, consistency index (CI), retention index (RI) and GC content were assessed and

11

240 sequences checked for large indels which can indicate the presence of these variant

241 sequences; none of these parameters suggested the occurrence of paralogues or pseudogenes.

242   Phylogenetic analysis of ITS sequence data was conducted using parsimony

243 approaches in the computer package PAUP* v4.0b10 (Swofford, 2003). The heuristic search

244 option was employed, using tree bisection-reconnection (TBR), with 1000 random stepwise

245 addition replicates and two trees held at each cycle. Branch support values were calculated

246 using a faststep heuristic search with 10,000 bootstrap replicates. A bootstrap majority rule

247 consensus tree was calculated in PAUP. We included an ITS sequence from GenBank;

248 EF638213.1 (*Acacia* [*Senegalia*] *senega*l) from Zimbabwe. Sequences of *Vachellia* [Syn.

249 *Acacia*] *farnesiana* (EF638219.1) from Australia and *Vachellia* [Syn. *Acacia*] *collinsii*

250 (EF638216.1) from Mexico were included as outgroups. Bayesian inference was performed in

251 MrBayes version 3.1.2 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003).

252 Four Markov Chain Monte Carlo simulations (three heated, one cold) were run with sampling

253 every 100 generations for 5,000,000 generations with the first 10% discarded as burn in.

254 Trees remaining after burn in were used to calculate posterior probabilities for nodes in the

255 majority-rule consensus tree.

256

257   The geographical structure of genetic variation at the cpSSR was explored using the

258 program SAMOVA (spatial analysis of molecular variation, Dupanloup *et al.*, 2002). The

259 method uses a simulated annealing procedure to define *K* groups of populations that are

260 geographically homogenous and maximally differentiated from each other. The method

261 requires the *a priori* definition of the number of groups (*K*) and generates *F* statistics (F_{SC},

262 F_{ST} and F_{CT}) based on AMOVA. One hundred simulated annealing processes were used for

263 values of *K* from 2 to 10. Herbarium specimens were also included in SAMOVA analysis

264   where we had at least three geographically proximate individuals to constitute a quasi

265   population. The most likely number K of groups was identified from F_{CT}, the proportion of

266   total genetic variance due to differences among groups of populations. Population structure

267   was further investigated by a Bayesian-based assignment algorithm using the STRUCTURE

268   program version 2.3.3 (Pritchard *et al.*, 2000). We used the admixture model with correlated

269   allele frequencies and run analysis with and without prior sample location information. We

270   performed 10 independent repetitions for each K (ranging from 1 to 20), with 100 000 MCMC

271   repetitions and a 100 000 burn-in period. The minimum number of K was evaluated using the

272   ΔK procedure (Evanno *et al.*, 2005). The geographical distribution of ITS, *trn*H-*psb*A

273   haplotypes and SAMOVA (cpSSR) delineated groups were mapped using the ESRI software

274   ArcMap 10.

275

276   Although the cpDNA molecule is non-recombining and therefore functions as a single

277   locus, data derived from RFLP and SSR loci were treated separately due to their different

278   modes of evolution. The program PERMUT (Pons and Petit, 1996)

279   (www.pierroton.inra.fr/genetics/labo/Software) was used to calculate the mean within-

280   population genetic diversity (h_S), the total gene diversity (h_T), and the proportion of diversity

281   resulting from genetic differentiation among populations (G_{ST}), as well as the corresponding

282   ordered parameters (taking into account similarities between haplotypes, N_{ST} and R_{ST} for

283   RFLP and SSR, respectively), to test for phylogeographic structure at rangewide and regional

284   geographic scales. Differentiation parameters were tested and compared with means from

285   1000 permutations. For all population-based analysis, samples were pooled where possible to

286   form populations of $n \geq 3$; highly isolated individual samples were excluded. Since ITS

287   phylogeny, SAMOVA and STRUCTURE analyses indicated an East and West African

phylogeographic pattern, PERMUT analysis was also carried out at the African geographic

289 scale. In addition to the PERMUT analysis (Pons and Petit, 1996), a hierarchical analysis of

290 molecular variance (AMOVA) and pairwise F_{ST} between geographic groups was computed

291 for both RFLP and cpSSR data using ARLEQUIN version 3.5 (Excoffier *et al.*, 2005).

292

293 **Results**

294 ITS sequences and phylogeny

295 A total of 65 sequences were obtained averaging 570 bp aligned length in all individuals. The

296 ITS region provided 109 parsimony informative characters (with outgroup), including four

297 indels ranging from 2-4 bp. Parsimony analysis of these data found 1,111,625 trees of length

298 251-576; CI = 0.831, RI = 0.843. The strict consensus tree (Figure 1) of 141,604 trees

299 resolved 17 nodes, of which 10 were supported by bootstrap values > 50 %. The *A. senegal*

300 specimens formed a monophyletic group with respect to the outgroup taxa, *V. collinsii* and *V.*

301 *farnesiana*. The basal group comprised two geographically disparate individuals (one from

302 Karofane in Niger, West Africa and the other a herbarium specimen from Botswana, Southern

303 Africa), which was highly divergent from the main branch (72 and 87 % bootstrap and

304 posterior probability support, respectively). The main branch exhibited high genetic variation

305 across the geographic range of the species with a topology displaying sequential East and

306 Southern Africa-West Africa organisation for the African sampled populations (Figures 1).

307 The basal clades comprise mainly East and Southern Africa populations. They consist of two

308 groups affiliated to var. *leiorhachis* (including Genbank sequence EF638213.1 from

309 Zimbabwe), strongly separated from the third group (98 % posterior probability) which

310 contains samples affiliated to vars. *senegal*, *kerensis* and *rostrata*. The terminal clades

311 comprise mainly West African samples affiliated to var. *senegal*, majority of which are in the

312   least resolved clade (see Figure 1); it is notable that the few samples of East African origin are

313   also mostly affiliated to var. *senegal*.

314

Geographic distribution of *trn*H-*psb*A haplotypes

316   The *trn*H-*psb*A amplicons were ~ 490 bp in size. Six RFLP *trn*H-*psb*A haplotypes were

317   detected (Figure 2b, Supplementary Table S3, Supporting information). Most populations

318   were fixed for a single haplotype. Haplotype 1 was the most frequent, highly dominant in

319   East Africa (mainly Kenya, Tanzania and Sudan), and present in Oman and Pakistan.

320   Haplotypes 2 and 3 were the most widespread, occurring throughout Africa and the Indian

321   subcontinent. Haplotypes 4, 5 and 6 were rare: haplotype 4 was restricted to Southern Africa;

322   haplotype 5 was present in Southern Africa and Oman and haplotype 6 occurred in two

323   geographically distant populations - in Maroua (Cameroon) and Jodhpur (ICRAF general

324   collection, India). The Maroua population was fixed for haplotype 6 while in Jodhpur it was

325   mixed with another haplotype, constituting only a small proportion (13 %). All of the

326   haplotypes found in India, Pakistan and Oman were shared with continental African

327   populations.

328

Geographic distribution of chloroplast microsatellite haplotypes

330   Thirty six haplotypes were identified (Supplementary Table S5, Supporting information).

331   Unlike the *trn*H-*psb*A haplotypes, most populations with at least two analysed

332   individuals/population were polymorphic (mixed). The distribution of haplotypes across the

333   range showed a pattern similar to that seen for *trn*H-*psb*A haplotypes, in that the majority

334   were regionally fixed in either East or West Africa. Haplotypes 1-21 (21 haplotypes) occurred

335   in East Africa, of which haplotypes 3, 4, 6, 7 were shared with Southern Africa. Haplotypes

15

336   19-36 (18 haplotypes) mainly occurred in West Africa. Populations in Sudan were the most

337   diverse. Populations from Oman, Pakistan and India predominantly shared haplotypes with

338   West and East Africa. Haplotypes 10 and 14 were the most prevalent, occurring mainly in the

339   East African region and accounting for 11.4 and 11.1 % of samples, respectively. Haplotypes

340   24 and 26 were the most prevalent in the West African region, accounting for 7.6 and 8.7 %

341   samples, respectively. Several private haplotypes were found, most notably in the Fallatu

342   population from Sudan (5), but also in the quasi population from Oman (3) and one each in

343   Koriema (Kenya), Kigwe (Tanzania), Maroua (Cameroon), Somo (Mauritania) and Jodhpur

344   (India, World Agroforestry Centre general collection). Populations Di (Burkina Faso), Daaba

345   and Rimoi (Kenya), Jodhpur Inde50 and Inde60 (India) were all fixed for a single haplotype.

346

347   ## Population genetic structure and phylogeographic patterns

348   The SAMOVA analysis showed little change in differentiation among groups from the lowest

349   to the highest F_{CT} values of 0.77 and 0.81 for K= 2 and 7, respectively. The biggest change

350   occurred between K = 2 and 3. We retained K = 3 (F_{CT} 0.79, $P < 0.001$) because single

351   population group membership appeared from K = 4. It also detected a key substructure that

352   distinguished the Southern Africa, some East African and Arabian Peninsula populations from

353   the core members of the East African group 1(also supported by strong pairwise F_{ST} values,

354   Supplementary Table S7).The three SAMOVA groups primarily separated into East (group 1)

355   and Southern (group 3) from West (group 2) African regions (Table 3 and Figure S1,

356   Supporting information). There were a few exceptions: two West African populations

357   (Diamenar, Senegal; and Tourba, Chad) were assigned to the East African group while

358   populations from Oman and India were grouped with either East or West African groups; two

359   Kenyan populations (Kulamawe and Ntumburi) were grouped with the quasi population from

360   South Africa. The STRUCTURE and the ΔK analysis identified $K = 2$ as the most likely

361   minimum number of clusters but only with prior sample location information (Supplementary

362   Figures S2a and b). As with SAMOVA groups, populations mainly grouped into East and

363   West African clusters, with Southern Africa populations assigned to the East African cluster,

364   while those from Arabian Peninsula and Indian subcontinent assigned either to East or West

365   African clusters. Most populations had >74 % ancestry, but with evidence of strong

366   admixture, particularly Ngane, Tourba and some Sudanese populations. Cluster 1 (=

367   SAMOVA groups 1 and 3) had lower differentiation ($F_{ST} = 0.188$) than Cluster 2 (=

368   SAMOVA group 2, $F_{ST} = 0.283$, Figure S1, Supporting information).

369

370   Structuring of the two cpDNA markers was fairly similar: populations in SAMOVA

371   group 1 predominantly had *trn*H-*psb*A haplotypes 1 and 2 and were located in East Africa;

372   populations in SAMOVA group 2 predominantly had *trn*H-*psb*A haplotypes 2 and 3 and were

373   located in West Africa (Table S4, Figures 2a and 2b). Overall, cpDNA variation did not

374   segregate with variety, particularly in the East and Southern Africa species range, where the

375   four varieties co-occur; regional location was more important than taxonomy in determining

376   haplotype. Hierarchical AMOVA showed contrasting marker resolution at the rangewide

377   scale (F_{CT}, *trn*H-*psb*A = 0.156, cpSSR = 0.789, Supplementary Table S7, Supporting

378   information). Differentiation was greatest between groups 2 vs 3 (F_{ST}, *trn*H-*psb*A = 0.850,

379   cpSSR = 0.964) compared to groups 1 vs 2 or 1 vs 3.

380

381   Levels and structure of genetic diversity

382   At the *trn*H-*psb*A locus, most diversity indices had generally similar levels of magnitude

383   among regions, except East and Southern Africa due to small sample size. Differentiation

384   among populations was greater in West than East Africa (G_{ST}, 0.764 vs 0.703, Table 3). The

385   contribution of phylogenetic relationships between haplotypes to among population

386   differentiation was not significant at the various geographic scales ($N_{ST} > N_{ST}$ (permuted), $P >$

387   0.05). Similar trends were also observed when only Africa populations were analysed (Table

388   S4, Supporting information). At the rangewide scale, the cpSSR data showed high within-

389   population diversity (h_S, 0.576 – 0.641) and high total diversity (h_T, 0.903 – 0.948).

390   Population differentiation neglecting haplotype order was $G_{ST} = 0.392$, but taking

391   microsatellite evolution into account was $R_{ST} = 0.673$ ($R_{ST} > R_{ST}$ (permuted), $P < 0.01$; Table

392   3), indicating clear phylogeographic structuring (Pons and Petit, 1996). However, this varied

393   considerably among the regions – with strong phylogeographic structure present West Africa

394   (SAMOVA group 2: $R_{ST} = 0.694$, $R_{ST} > R_{ST}$ (permuted), $P < 0.05$), and no significant

395   structure in East Africa (SAMOVA group 2: $R_{ST} = 136$, $R_{ST} < R_{ST}$ (permuted), $P > 0.05$) and

396   East and Southern Africa (SAMOVA group 3: $R_{ST} = 0.471$, $R_{ST} < R_{ST}$ (permuted), $P > 0.05$).

397   In contrast, strong phylogeographic structure was obtained in both East and West African

398   populations when only African dataset was analysed (Table S4, Supporting information).

399

Discussion

401   Taking *A. senegal* as a whole, our data show a significant geographic structuring of genetic

402   variation, with the major division separating East and Southern African populations from

403   those in West and Central Africa. Patterns in the nuclear ITS and chloroplast data were

404   largely concordant. Genetic data poorly reflected the taxonomic subdivision of *A. senegal*

405   vars. *kerensis* and *senegal*, suggesting few barriers to hybridization among these varieties

406   where they co-occur. The occurrence of highly divergent ITS haplotypes suggests

407   hybridization among *Acacia* species may be more frequent than has been observed to date.

408 The early and multiple evolutionary divergence events within the species support the

409 hypothesis that East Africa was the centre of diversification, and that the current wide

410 distribution has arisen largely following past colonisation, migration and range expansion

411 events from Eastern Africa since the late Pleistocene.

412

413 ## Phylogenetic relationships: intraspecific taxonomy, migration and hybridization

414 In our analysis of ITS data, var. *leiorhachis* was distinct and contained a significant amount of

415 variation, suggesting its divergence from the other varieties (*senegal*, *kerensis* and *rostrata*)

416 was ancient. Throughout its recorded distribution var. *leiorhachis* is found in association with

417 or in close proximity to other varieties, therefore genetic distinctiveness at the nuclear ITS

418 locus appears to be maintained even in the face of potential gene flow. The lack of

419 differentiation among the varieties at chloroplast loci may therefore be due to retained

420 ancestral variation or to chloroplast capture if hybridization among varieties occurs. However,

421 it is reported that var. *leiorhachis* is phenologically asynchronous with other varieties, even

422 where they co-occur as proximate populations (Fagg and Allison, 2004). Given the

423 geographically close proximity of the varieties in East Africa, it seems likely that the

424 divergence is ecologically-driven and further detailed studies on the distribution should be

425 undertaken. In contrast, variety *rostrata* is largely confined to Southern Africa. Rather than

426 being ecologically-driven, as seems to be the case for var. *leiorhachis*, it seems more likely

427 that the origins of var. *rostrata* lie in dispersal to Southern Africa and subsequent independent

428 evolution due to drift.

429

430 Of the other varieties, var. *kerensis* is restricted to East Africa, whilst var. *senegal* has

431 the widest geographic distribution, occurring throughout the range from East to West Africa

19

432 and to India. The terminal clade mainly comprised West African var. *senegal*, suggesting a

433 relatively recent expansion of the range, probably from an origin in Eastern Africa via Sudan

434 and Central Africa to the West and via the Horn of Africa and the Arabian Peninsula to the

435 East. According to the recent phylogenetic chronogram of African acacias by Bouchenak-

436 Khelladi *et al*. (2010), var. *leiorhachis* diverged from var. *rostrata* approximately 1 Mya.

437 Although such estimates are often associated with wide range margins, it suggests that the *A.*

438 *senegal* species complex predates the Pleistocene and that the early splits within var.

439 *leiorhachis* and diversifications within *A. senegal* coincide with a period marked by

440 prolonged glacial cycles (100 ky periodicity) and extreme aridity (Maley, 2001; Plana, 2004).

441 On the other hand, the unresolved terminal clade (Figure 1) depicts a rapid expansion in the

442 Late Pleistocene by var. *senegal*, which could be as recent as the last glacial maximum. This

443 period encompasses major glacial maxima with prolonged aridity phases (deMenocal, 1995;

444 Plana, 2004), which offers the necessary driver of restriction and expansion of vegetation that

445 could explain the relatively close relationships among populations across such a wide

446 geographic range.

447

448   In a number of cases, hybridization appears to have been important. In the basal group

449 that comprised an individual from Karofane in Niger (West Africa) and one from Botswana

450 (Southern Africa), similar, highly divergent, sequences were found despite wide geographic

451 separation. It seems at least possible that this is the outcome of interspecific hybridization – of

452 the 20 other related species that form the *A. senegal* species complex, most co-occur or share

453 its ecological range (Ross, 1981). Alternatively, the phylogenetic association between the

454 West African and Southern African individuals could also suggest chance retention of

455 ancestral variation. In addition, *trn*H-*psb*A haplotypes were variably distributed among ITS

456 clades, which could indicate introgressive hybridization (Figures 1; Supplementary Table S6).

457 Byrne *et al*. (2002) also reported haplotype sharing among clades/sub-specific taxa within *A*.

458 *acuminata* in the mesic and arid zones of Western Australia, which they largely attributed to

459 retention of ancestral polymorphism or incomplete lineage sorting. Both introgression or

460 hybridizations and incomplete lineage sorting may account for the observations in *A. senegal.*

461 Besides the reported natural hybrid (*A. laeta*) with *A. mellifera*, allotetraploids have recently

462 been found in some members of the *A. senegal* complex indicating that hybridization may not

463 be uncommon (Assoumane A *et al*., unpublished). Whilst sharing of haplotypes between the

464 regions may also suggest homoplasy, the unexpected shared ITS and chloroplast haplotypes

465 between individuals from East Africa's var. *leiorhachis* populations and one of the Jodhpur

466 Indian accessions (see H4, Figures 1 and Table 2,) is likely to be due to human-mediated

467 dispersal. These regions have had a long history of human migrations and trade links, which

468 may have included gum arabic as a commodity and possible germplasm exchanges or

469 introductions such as was the case for the widely domesticated drumstick tree (*Moringa*

470 *oleifera*, Moringaceae; Muluvi *et al*., 1999).

471

472 Genetic diversity, population structure and phylogeography of *Acacia senegal*

473 (refugia and gene flow barriers)

474 Recognising the utility and limitations of different markers on interpretations of diversity and

475 differentiation indices within and among species (e.g. see Petit *et al.*, 2005; Meirmans and

476 Hedrick, 2010), the two chloroplast markers used in this study showed considerable

477 complementarity. Most populations were fixed for single chloroplast RFLP *trn*H-*psb*A

478 haplotypes but contained multiple cpSSR haplotypes. Although RFLP *trn*H-*psb*A haplotypes

479 had less resolution than cpSSR haplotypes at the rangewide scale, both showed similar

480   regional genetic structuring and phylogeographic patterns (Table 3, Figures 2, Figure S1). The

481   levels of within-population and total diversity estimated from cpSSR data (h_S = 0.576, h_T =

482   0.948) were comparable to those reported by Byrne *et al*. (2002) within *A. acuminata*

483   populations (*cf.* h_S = 0.442; h_T = 0.920) in Western Australia. Of particular interest is

484   comparisons with the boabab tree which also showed a similar East-West African

485   phylogeographic pattern; they had contrasting within-population, total diversity and

486   population differentiation with RFLP data (e.g. *A. senegal*: h_S, 0.155; h_T , 0.711; G_{ST}, 793 vs

487   boabab: h_S, 0.017; h_T, 0.58; G_{ST}, 0.970; Pock Tsy *et al*., 2009). These differences are likely to

488   be due to the contrasting life histories and modes of seed or fruit dispersal between the two

489   species. *A. senegal* is pollinated mainly by bees and seeds dispersed by wind (Fagg and

490   Allison, 2004). Baobab is pollinated by bats and seed dispersed by large mammals and

491   humans (Pock Tsy *et al*., 2009). Baobab's long domestication history, significant human-

492   mediated dispersal and reported longevity of ~1300 years are particularly important

493   distinguishing factors.

494

495   The distribution patterns of cpDNA haplotypes suggest that colonisation, migration or

496   expansion events may have happened more than once originating from an East African source

497   population. The Horn of Africa is an important centre of speciation in the *A. senegal* complex

498   and also has the highest concentration of African *Acacia* species (Ross, 1981; Fagg and

499   Allison, 2004). In *A. senegal*, this is the only region where all the four varieties are reportedly

500   distributed, in some cases as co-occurring populations (Fagg and Allison, 2004).

501   Diversification of the species may have been driven by an interaction between the climate

502   changes of the late Pleistocene and the complex topography of Eastern Africa, which provides

503   potential refugia in areas that have historically been sheltered against the impacts of extreme

504   climate oscillations. During enhanced and prolonged aridity phases the *Acacia* range may

505   have shifted or migrated to higher altitudes, tracking moist forest (Hamilton, 1982), and

506   causing fragmentation. These sites could have acted as source populations for (re-)

507   colonisations or migrations into other regions during favourable climatic conditions. The

508   genetic and morphological diversity of *A. senegal* in this region, as well as fossil pollen data,

509   supports this hypothesis. *Acacia*-type pollen, predating the Pleistocene, has been recorded

510   from the Rift Valley floor - at the Lokichar Basin (near one of our sampling sites, see Table 1)

511   in the southwestern part of Lake Turkana Basin, Kenya (Vincens *et al.*, 2006). The Lake

512   Turkana Basin still harbours several *Acacia* spp., including the predominant *A. senegal*. There

513   is also pollen evidence of montane *Acacia* woodlands (up to 4040 masl) in the East African

514   highlands predating the arid period at ~10 000 years BP (Hamilton, 1982). These pollen

515   records suggest that East Africa may have harboured refugia for *Acacia* taxa predating the

516   Pleistocene. The haplotype-rich populations of Sudan are found in this region. Of these

517   populations, Fallatu forest had an exceptionally high number of private cpSSR haplotypes (5),

518   possibly suggesting an ancient refugium. The Sudanese populations occur in the so-called

519   gum belt from which the bulk of the internationally traded gum is produced in the traditional

520   'gum orchard' agroforestry systems. Although it is possible that domestication has resulted in

521   incorporation of extra diversity, importation of seed would be uncharacteristic because the

522   traditional gum production systems are normally established with local collections or through

523   natural regenerations (Fagg and Allison, 2004). Alternatively, if the region has long acted as a

524   refugium, the potential for hybridization and chloroplast capture from congeners is higher and

525   may offer an explanation. In Southern Africa, the presence of a private *trn*H-*psb*A haplotype

526   suggests long term persistence and isolation of the population in this region. The expansion

527   and contraction of moist rainforest across East Africa is likely to have been the driving force

528 behind isolation of Southern African populations (Cowling *et al*., 2008). Similar patterns have

529 been found in phylogeographic studies of other savannah-adapted species, such as the plains

530 zebra (Lorenzen *et al*., 2008). Further sampling of *A. senegal* in Southern Tanzania and

531 Mozambique would help to test this hypothesis.

532

533   In West Africa, population differentiation was higher and had a stronger

534 phylogeographic structure than that of East Africa (cpSSR data, G_{ST} = 0.458 vs 0.290; RFLP,

535 0.764 vs 0.703, Table 3; also see Supplementary Figure S2, STRUCTURE analysis, F_{ST} =

536 0.283 vs 0.188), which also indicates high admixture of *A. senegal* organelle lineages. This

537 suggests that dispersal from refugia occurred over relatively short distances, probably due to

538 the influence of geographic barriers, as invoked to explain the genetic structuring of other dry

539 woodland and savannah species in the Sudano-Sahelian region (e.g. shea tree, Allal *et al*.,

540 2011). For example, in the Chad basin, an inundation during the Quaternary pluvial, followed

541 by the formation of a Mega-Lake was suggested to have isolated baobab populations of West

542 Africa from those of East Africa (Pock Tsy *et al*., 2009). The historical Mega-Lake zone also

543 coincides with the haplotype disjunctions observed in this study between Central -West and

544 West African *A. senegal* populations. The two populations that mark this contact zone are

545 Tourba, Chad (east of Lake Chad) and Maroua, Cameroon (south of Lake Chad). However,

546 the Maroua population (fixed *trn*H-*psb*A haplotype 6) would have also been isolated or

547 restricted either on the fringes of tropical rainforest or riparian conditions. Further West, dry

548 riverine woodlands and gallery forests along the West African extensive river systems

549 (Senegal, Niger, Volta and Gambia) may have also acted as refugia during periods of extreme

550 aridity that also promoted southward expansion of the Sahara and regression of tropical. The

551 possible refugial zones for *A. senegal* in this region would therefore include the westernmost

552 part of West Africa (Western Senegal) and Central Africa (north of Cameroon). Although

553 poorly represented in pollen cores, *Acacia*-type pollen records have also been reported for the

554 Sudano-Sahelian region during the Holocene humid period (e.g. Jikariye Lake, Nigeria, ~11

555 000 years BP, Waller *et al*., 2007; Lake Yoa, Chad, ~ 6 000 years BP, Lézine *et al*., 2009).

556 However, lack of pollen data predating the Holocene and LGM may not preclude the likely

557 occurrence of ancient *Acacia* refugia in West Africa. In other parts of the range beyond

558 Africa, *Acacia*-type pollen have also been recorded from the semi-desert around Kwar al

559 Jaramah (Oman, Arabian Peninsula) and Makran on the Pakistani coast and the Indian

560 subcontinent (~ 6 000 years BP, Lézine, 2009).

561

562 Conclusions

563 The combined nrDNA and cpDNA analysis of *A. senegal* have shown high haplotypic

564 variation with both regional and rangewide phylogeographic patterns suggesting a long

565 history of colonisation and expansion events characterised by extensive, recurrent gene flow

566 among populations and regions. The phylogenetic analysis revealed multiple evolutionary

567 divergence events that also separated var. *leiorhachis* from the other three (*senegal*, *kerensis*

568 and *rostrata*). The phylogeographic structure separating East and Southern Africa from West

569 African populations reflects a pattern reported for other drylands species straddling the

570 Sudano-Sahelian region (e.g. baobab, Pock Tsy *et al*., 2009; shea, Allal *et al*., 2011).

571 However, contrary to our prediction and in contrast to the other species, within-region

572 differentiation of *A. senegal* was greater in West Africa than East and Southern Africa, clearly

573 showing species differences at the regional scale. Spatial studies at the population and

574 landscape scales will be necessary to elucidate factors that have influenced the observed

575 regional differences. Cytological studies within the species complex would provide evidence

576 for introgression or hybridization events and show whether certain morphological variants

577 (intraspecies) have been more successful than others in colonisation and range expansion in

578 the more arid conditions. Further studies that take a comparative approach, such as a

579 combined analysis of key species within the wider *A. senegal* complex will be required to

580 understand how arid-adapted plant species have dispersed and occupied various dryland

581 habitats and vegetation since diversification.

582

591

592 **Author contributions:** DWO led the draft of the manuscript. DWO, SC, AT and JW

593 designed the work, acquired and analysed data; AG led field design, sample and data

594 collection in Senegal. All authors contributed ideas, comments and revised the manuscript.

595

Conflict of interest

597 The authors declare no conflict of interest.

598

599 **Supplementary information is available at Heredity's website**

600

Data archiving

Sequence data have been submitted to the GenBank: accession numbers HQ605042-

HQ605077.

References

604

605 Allal F, Sanou H, Millet L, Vaillant A, Camus-Kulandaivelu L, Logossa ZA *et al*. (2011).

606   Past climate changes explain the phylogeography of *Vitellaria paradoxa* over Africa.

607   *Heredity* **107:** 174-186.

608 Álvarez I, Wendel JF (2003). Ribosomal ITS sequences and plant phylogenetic inference.

609   *Mol Phylogenet Evol* **29:** 417-434.

610 Ariati SR, Murphy DJ, Udovicic F, Ladiges PY (2006). Molecular phylogeny of three groups

611   of acacias (*Acacia* subgenus *Phyllodineae*) in arid Australia based on the internal and

612   external transcribed spacer regions of nrDNA. *Syst Biodivers* **4:** 417-426.

613 Ayele TB, Gailing O, Umer M, Finkeldey R (2009). Chloroplast DNA haplotype diversity

614   and postglacial recolonization of *Hagenia abyssinica* (Bruce) J.F. Gmel. in Ethiopia.

615   *Plant Syst Evol* **280:** 175-185.

616 Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ

617   (1995). The its region of nuclear ribosomal region: a valuable source of evidence on

618   angiosperm phylogeny. *Ann Mo Bot Gard*, **82**: 242-277.

619 Bouchenak-Khelladi Y, Maurin O, Hurter J *et al*. (2010). The evolutionary history and

620   biogeography of Mimosoideae (Leguminosae): An emphasis on African acacias. *Mol*

621   *Phylogenet Evol* **57:** 495-508.

622 Brown GK, Murphy DJ, Miller JT, Ladiges PY (2008). *Acacia s.s.* and its relationship among

623   tropical legumes, Tribe Ingeae (Leguminosae: Mimosoideae). *Syst Bot* **33:** 739-751.

624 Bukhari YM (1997) Cytoevolution of taxa in *Acacia* and *Prosopis* (Mimosaceae). *Aust J Bot*
625   **45:** 879-891.

626 Byrne M, Macdonald B, Coates D (2002). Phylogeographical patterns in chloroplast DNA
627   variation within the *Acacia acuminata* (Leguminosae: Mimosoideae) complex in
628   Western Australia. *J Evol Biol* **15:** 576-587.

629 Caetano C, Prado D, Pennington RT, Beck S, Olivera-Filho A, Spichiger R, Naciri Y (2008).
630   The history of seasonally dry forests in eastern South America: inferences from the
631   genetic structure of the tree *Astronium urundeuva* (Anacardiaceae*). Mol Ecol* **17:**
632   3147-3159.

633 Cavers S, Navarro C, Lowe AJ (2003). Chloroplast DNA phylogeography reveals
634   colonization history of *Cedrela odorata* L., in Mesoamerica. *Mol Ecol* **12:** 1451-1460.

635 Collevatti RG, Grattapaglia D, Hay JD (2003). Evidences for multiple maternal lineages of
636   *Caryocar brasiliense* populations in the Brazilian Cerrado based on the analysis of
637   chloroplast DNA sequences and microsatellite haplotype variation. *Mol Ecol* **12:** 105-
638   115.

639 Cowling SA, Cox PM, Jones CD, Maslin MA, Peros M, Spall SA (2008). Simulated glacial
640   and interglacial vegetation across Africa: implications for species phylogenies and
641   trans-African migration of plants and animals. *Glob Change Biol* **14:** 827-840.

642 Daïnou K, Bizoux J-P, Doucet J-L, Mahy G, Hardy OJ, Heuertz M (2010). Forest refugia
643   revisited: nSSRs and cpDNA sequences support historical isolation in a wide-spread
644   African tree with high colonization capacity, *Milicia excelsa* (Moraceae). *Mol Ecol*
645   **19:** 4462-4477.

646 Dupanloup I, Schneider S, Excoffier L (2002). A simulated annealing approach to define the
647   genetic structure of populations. *Mol Ecol* **11:** 2571-2581.

648 El Amin HM (1976). *Acacia laeta* (R. Br.) ex. Benth., considered as a species of hybrid
649   origin. *Sudan Silva* **3:** 14-23.

650 Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using
651   the software STRUCTURE: a simulation study. *Mol Ecol* **14:** 2611-2620.

652 Excoffier L, Laval G, Schneider S (2005). Arlequin (version 3.0): an integrated software
653   package for population genetics data analysis. *Evol Bioinform* **1:** 47-50.

654 Fagg CW, Allison GE (2004). *Acacia senegal and the gum arabic trade. Monograph and*
655   *annotated bibliography*. Tropical Forestry Papers No. 42, Oxford Forestry Institute,
656   Oxford.

657 Hamilton AC (1982). *Environmental history of East Africa. A study of the Quaternary*.
658   Academic Press: London.

659 Huelsenbeck J P, Ronquist.F (2001). MRBAYES: Bayesian inference of phylogeny.
660   *Bioinformatics* **17**:754-755.

661 Jacobs BF (2004). Palaeobotanical studies from tropical Africa: relevance to evolution of
662   forest, woodland and savannah biomes. *Philos T Roy Soc B* **359:** 1573-1583.

663 Lavin M, Herendeen PS, Wojciechowski MF (2005). Evolutionary rates analysis of
664   Leguminosae implicates a rapid diversification of lineages during the Tertiary. *Syst*
665   *Bot*, **54:** 575-594.

666   Lézine A-.M (2009). Timing of vegetation changes at the end of the Holocene humid period

667   in desert areas at the northern edge of the Atlantic and Indian monsoon systems. *CR*

668   *Geoscience*, **341:** 750-759.

669   Lock JM (2006). The seasonally dry vegetation of Africa: parallels and comparisons with the

670   neotropics. In: Pennington RT, Lewis GP, Ratter JA (eds) *Neotropical savannas and*

671   *dry forests: diversity, biogeography, and conservation,* Taylor and Francis Ltd:

672   London. pp 449-467.

673   Lorenzen ED, Arctander P, Siegismund HR (2008). High variation and very low

674   differentiation in wide ranging plains zebra (*Equus quagga*): insights from mtDNA

675   and microsatellites. *Mol Ecol* **17:** 2812-2824.

676   Lorenzen ED, Masembe C, Arctander P, Siegismund HR (2010). A long-standing Pleistocene

677   refugium in Southern Africa and a mosaic of refugia in East Africa: insights from

678   mtDNA and the common eland antelope. *J Biogeogr* **37:** 571-581.

679   Lowe AJ, Harris D, Dormont E, Dawson IK (2010). Testing putative tropical forest refugia

680   using chloroplast and nuclear phylogeography. *Trop Plant Biol* **3:** 50-58.

681   Maley J (2001). The impact of arid phases on the African rain forest through geological

682   history. In: Weber W, White L, Vedder A, Naughton-Treves L (eds) *African rain*

683   *forest ecology and conservation,* Yale University Press: New Haven, Connecticut. pp.

684   68-87.

685   Meirmans PG, Hedrick PW (2010). Assessing population structure: F_{ST} and related measures.

686   *Mol Ecol Resour* **11**: 5-18.

687   deMenocal PB (1995). Plio-Pleistocene African climate. *Science* **270:** 53-59.

688   Muluvi GM, Sprent JI, Soranzo N, Provan J, Odee D, Folkard G and Powell W (1999).

689   Amplified fragment length polymorphism (AFLP) analysis of genetic variation in

690   *Moringa oleifera* Lam. *Mol Ecol* **8:** 463-470.

691   Omondi SF, Kireger E, Dangasuk OG, Chikamai B, Odee DW, Cavers S, Khasa DP (2010).

692   Genetic diversity and population structure of *Acacia senegal* (L.) Willd. in Kenya.

693   *Trop Plant Biol* **3:** 59-70.

694   Pennington RT, Lavin M, Oliveira-Filho A (2009). Woody plant diversity, evolution and

695   ecology in the tropics: perspectives from seasonally dry tropical forests. *Annu Rev*

696   *Ecol Evol Syst* **40:** 437-457.

697   Pennington RT, Lavin M, Prado DE, Pendry CA, Pell SK, Butterworth CA (2004). Historical

698   climate change and speciation: neotropical seasonally dry forest plants show patterns

699   of both Tertiary and Quaternary diversification. *Philos T Roy Soc B* **359:** 515-537.

700   Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005). Comparative

701   organization of chloroplast, mitochondrial and nuclear diversity in plant populations.

702   *Mol Ecol* **14:** 689-701.

703   Plana V (2004). Mechanisms and tempo of evolution in the African Guineo-Congolian

704   rainforest. *Philos T Roy Soc B* **359:** 1585-1594.

705   Pock Tsy J.-M L, Lumaret R, Mayne D, Vall AOM, Abutaba YIM, Sagna M *et al*. (2009).

706   Chloroplast DNA phylogeography suggests a West African centre of origin for the

707   baobab, *Adansonia digitata* L. (Bombacoideae, Malvaceae). *Mol Ecol* **18:** 1707-1715.

708   Pons O, Petit RJ (1996). Measuring and testing genetic differentiation with ordered versus

709   unordered alleles. *Genetics* **144:** 1237-1245.

710   Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using

711   multilocus genotype data. *Genetics* **155**: 945-959.

712   Ramos ACM, Lemos-Filho JP, Ribeiro RC, Santos FC, Lovato MD (2007). Phylogeography

713   of the Tree *Hymenaea stigonocarpa* (Fabaceae: Caesalpinioideae) and Influence of

714   Quaternary climate changes in the Brazilian Cerrado. *Ann Bot* **100**: 1219-1228.

715   Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM (2001). Rapid

716   diversification of a species-rich genus of neotropical rain forest trees. *Science* **293:**

717   2242-2245.

718   Ronquist F. Huelsenbeck JP (2003). MRBAYES 3: Bayesian phylogenetic inference under

719   mixed models. *Bioinformatics* **19**: 1572-1574.

720   Ross JH (1979). A conspectus of the African *Acacia* species. *Mem Bot Surv S Afr* **44:** 1-155.

721   Ross JH (1981). An analysis of the African *Acacia* species: their distribution, possible origins

722   and relationships. *Bothalia* **13:** 398-413.

723   Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J *et al*. (2005). The tortoise and the

724   hare: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic

725   analysis. *Am J Bot* **92:** 142-166.

726   Sprent JI, Odee DW, Dakora FD (2010). African legumes: a vital but under-utilized resource.

727   *J Exp Bot* **61:** 1257-1265.

728 Sun Y, Skinner DZ, Liang GH, Hulbert SH (1994). Phylogenetic analysis of *Sorghum* and

729 related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theor Appl*

730 *Genet* **89:** 26-32.

731 Swofford, D L (2003). *PAUP*. Phylogenetic Analysis Using Parsimony (*and other*

732 *methods).* Version 4.0 b10, version 4.0 beta 10. Sunderland: Sinauer Associates.

733 Vincens A, Tiercelin J-.J, Buchet G (2006). New Oligocene-early Miocene microflora from

734 the southwestern Turkana Basin. Palaeoenvironmental implications in the northern

735 Kenya Rift. *Palaeogeogr Palaeocl* **239:** 470-486.

736 Waller MP, Street-Perrot FA, Wang H (2007). Holocene vegetation history of the Sahel:

737 pollen, sedimentological and geochemical data from Jikariya Lake, north-eastern

738 Nigeria. *J Biogeogr* **34:** 1575-1590.

739 Wardill TJ, Graham GC, Zalucki N, Palmer WA, Playford J, Scott KD (2005). The

740 importance of species identity in the biocontrol process: identifying the subspecies of

741 *Acacia nilotica* (Leguminosae: Mimosoideae) by genetic distance and the implications

742 for biological control. *J Biogeogr* **32:** 2145-2159.

743 Weising K, Gardner RC (1999). A set of conserved PCR primers for the analysis of simple

744 sequence repeat polymorphisms in chloroplast genomes of dicotyledonous

745 angiosperms. *Genome* **42:** 9-19.

746

Titles and legends to figures

748

749 **Figure 1**

750 Bayesian 50 % majority-rule consensus phylogram of internal transcribed spacer (ITS)

751 sequences from *Acacia senegal* individuals sampled across its distribution range. Branches

752 are labelled with ≥ 70 % bootstrap support (below) and posterior probability (above) values.

753 Arrow indicates presence of a single clade in the maximum parsimony strict consensus tree.

754 The tree was rooted with *Acacia* [*Vachellia*] *farnesiana* and *Acacia* [*Vachellia*] *collinsii* (see

755 inset tree, double line denotes the point at which the break is in the main tree) obtained via

756 GenBank (accession numbers EF638219 and EF638216, respectively). *Acacia* [*Senegalia*]

757 *senegal* sequence is accession number EF638213 from Zimbabwe. Each haplotype is code-

758 labelled HX where X refers to a number 1-36 as described in Table S2, Supporting

759 information. Hypothesised colonisation or range expansion events are indicated. Clades are

760 identified by the different colours on the vertical bar (black, basal group; red, var. *leiorhachis*;

761 dark blue, var. *rostrata*; blue - light blue, vars. *kerensis-senegal*; dark green-light green, var.

762 *senegal*); clades that are phylogenetically related are shown with different shades of one

763 colour. The same colour coding is used for Figure 2a. Scale bar signifies 0.1 substitutions per

764 nucleotide site.

765

766 **Figure 2**

767 Rangewide distribution of nuclear (**a**) and chloroplast (**b**) haplotypes in *Acacia senegal*. Data

768 are: (**a**) haplotypes from internal transcribed spacer (ITS) of the nuclear ribosomal DNA and

769 (**b**) PCR-RFLP haplotypes of chloroplast *trn*H-*psb*A intergenic spacer of *Acacia senegal*;

770 each circle represents a single population with colour denoting haplotype variation. The

771 number of samples per population is presented in Supplementary Tables S2 and S3,

772 respectively. Inset: approximate distribution of *A. senegal* in Africa (hatch-shaded area).

773

774

775 **Supplementary Information**

776 **Table S2** Details of sequenced samples.

777 **Table S3** Description of *trn*H-*psb*A haplotypes and their distribution within populations of

778 *Acacia senegal*. Polymorphic fragments are presented as presence (1) or absence (2). RBGE

779 represents sources from Royal Botanic Garden, Edinburgh, UK.

780 **Table S4** Haplotype diversity (within-population diversity h_S, v_S and total diversity h_T, v_T)

781 and differentiation (G_{ST}, N_{ST} and R_{ST}) for *Acacia senegal* populations geographic divisions

782 within Africa (East, including Southern Africa, and West, including Central Africa). Number

783 of populations, *n*; standard errors (SE) in parentheses.

784 **Table S5** Distribution of cpSSR haplotypes within populations of *Acacia senegal.*

785 **Table S6** cpSSR SAMOVA delineated groups and affiliations to geographic region, putative

786 variety and *trn*H-*psb*A haplotypes from 40 populations of *Acacia senegal.*

787 **Table S7** Hierarchical analysis of genetic differentiation for rangewide and pairwise

788 comparisons between groups using RFLP and cpSSR markers. SAMOVA delineated Groups

789   1, 2 and 3, namely; East Africa, mainly West Africa, and East and Southern Africa,

790   respectively. Number of populations, *n*.

791   **Figure S1**

792   Geographic distribution of groups of *Acacia senegal* populations delineated by chloroplast

793   microsatellite (cpDNA) SAMOVA; circle sizes are proportional to the number of individuals

794   (*cf*. STRUCTURE clusters, Figures S2 a and b; see also Table S5).

795

796   **Figure S2**

797   Genetic structure of cpSSR of 290 individuals from 39 populations. (**a**) Bar plot showing

798   clustering of individuals by STRUCTURE with $K = 2$ (Pritchard *et al*., 2000). Colour

799   represents proportion of ancestry derived from each cluster; red = cluster 1 ($F_{ST} = 0.188$,

800   mainly East Africa), green = cluster 2 ($F_{ST} = 0.283$, mainly West Africa). Gray lines represent

801   populations listed as follows: 1, Ngane; 2, Diamenar; 3, Daiba; 4, Kidira; 5, Aite; 6, Kirane;

802   7, Djiguéni; 8, Somo; 9, Karofane; 10, Burkina Di; 11, Burkina Bissiga Fc; 12, Maroua; 13,

803   Tourba; 14, Sudan RBGE specimens; 15, Fallatu; 16, Kordofan; 17, Sodera; 18, Kaleing'; 19,

804   Kakuma; 20, Lokichar; 21, Ngurunit; 22, Merille; 23, Serolipi; 24, Rimoi; 35, Kulamawe; 26

805   Ngarendare; 27, Marigat; 28, Koriema; 29, Ntumburi; 30, Magadi; 31, Kajiado; 32, Kibwezi;

806   33, Kigwe; 34, Wangingombe; 35, South Africa RBGE specimens; 36, Oman (Dhofar) RBGE

807   specimens; 37, India (Jodhpur World Agroforestry Centre collection); 38, India Jodhpur

808   Inde50; 39, India Jodhpur Inde60. (**b**) Bar plot showing clustering in regions and subregions:

809   1, West Africa; 2, Central Africa (Cameroon and Chad); 3, Sudan; 4, East Africa; 5, South

810   Africa; 6, Arabian Peninsula (Dhofar, Oman); 7, India.

811

812　**Table 1** Growth and morphological characters for distinguishing *Acacia senegal* varieties
813　(Source: Ross, 1979; Fagg and Allison, 2004)
814

Key characters	*senegal*	*kerensis*	*rostrata*	*leiorhachis*
Inflorescence axis	Sparsely to densely pubescent	Sparsely to densely pubescent	Sparsely to densely pubescent	Glabrous or subglabrous, except sometimes for some basal hairs
Tree, shrub or bush	Tree with a distinct trunk	Shrub or bush without a distinct trunk	Tree with a distinct trunk	Slender spindly tree, whippy
Shape of pod	Rounded to acute (sharply pointed) at apex	Rounded to acute, seldom acuminate	Markedly acuminate or rostrate (beaked) at apex	Rounded to acute
Phenology	Flowers usually produced after the leaves	Not known	Not known	Flowers often produced before or with the young leaves

815
816

817

Table 2 Locations of all sampled populations of *Acacia senegal*

Region	Country	Population	Variety	n	Latitude	Longitude
East Africa	Ethiopia	Sodera	*senegal*	3	08°24'N	39°23'E
	Kenya	Archer's Post	*kerensis*	2	00°39'N	37°39'E
		Daaba	*kerensis*	2	00°32'N	37°46'E
		Kajiado	*senegal*	7	01°54'S	36°45'E
		Kakuma	*kerensis*	5	03°45'N	34°40'E
		Kaleing'	*kerensis*	6	04°24'N	35°31'E
		Kargi	*kerensis*	1	02°39'N	37°28'E
		Kibwezi	*kerensis*	5	02°13'S	38º04'E
		Kulamawe	*leiorhachis*	6	00°32'N	37°60'E
		Koriema	*senegal*	7	00º27'N	35°52'E
		Lokichar	*kerensis*	3	02°22'N	35°38'E
		[a]Machakos	[c]*kerensis*	1	01°31'S	37°16'E
		Magadi	*leiorhachis*	7	01°32'S	36°35'E
		Marigat	*kerensis*	6	00°28'N	35°54'E
		Merille	*kerensis*	3	01°32'N	37°73'E
		Ngarendare	*kerensis*	8	00°28'N	37°25'E
		Ngurunit	*leiorhachis*	6	01°43'N	37°17'E
		Ntumburi	*senegal*	6	00º12'N	37°31'E
		Rimoi	*senegal*	7	00º40'N	35°34'E
		Serolipi	*kerensis*	5	01°15'N	37°59'E
		Taita- Taveta	*senegal*	2	03°27'S	38°29'E
	Somalia	[a]Hargesia	*kerensis*	1	09°30'N	44°03'E
		[a]Afgoi	[c]*kerensis*	1	02°06'N	45°08'E
	Tanzania	Kigwe	*leiorhachis*	15	06°06'S	35°29'E
		Wangingombe	*senegal*	8	08°51'S	34°38'E
	Sudan	Fallatu Forest	*senegal*	12	13°06'N	30°08'E
		Kordofan	*senegal*	6	12°44'N	29°35'E
		[a]Khartoum	*senegal*	1	15°38'N	32°32'E
		[a]Kundoura, Nyala	*senegal*	1	12°52'N	30°13'E
		[a]Bora	*senegal*	1	08°49'N	26°11'E
		[a]Wad Medani,	*senegal*	1	14°24'N	33°31'E
		[a]El Haraza	*senegal*	1	11°18'N	24°11'E
		[a]Sobar	*senegal*	1	08°48'N	32°54'E
		[a]Rodom,	*senegal*	1	12°05'N	23°03'E
		[a]Kassala	*senegal*	1	15°27'N	36°24'E
		[a]El Felaya	*senegal*	1	09°36'N	28°26'E
Central Africa	Cameroon	Maroua	*senegal*	10	10º15'N	14º14'E
	Chad	Tourba	*senegal*	2	12°49'N	15°18'E
West Africa	Burkina Faso	Bissiga Fc	*senegal*	6	12°26'N	00°32'W
		Di (Sousou)	*senegal*	7	13°10'N	03°25'W
	Mali	Aïte	*senegal*	4	15º05'N	11º39'W
		Kirane	*senegal*	50	15°23'N	10°15'W
		Somo	*senegal*	5	13°17'N	04°54'W
	Mauritania	Djiguéni	*senegal*	6	15°44'N	08°40'W
		Kankossa	*senegal*	1	15°56'N	11°27'W
	Niger	Karofane	*senegal*	2	14°18'N	06°11'E

	Senegal	Daïba	*senegal*	6	15°22'N	13°08'W
		Diaménar	*senegal*	6	16°00'N	15°54'W
		Kidira	*senegal*	4	14°28'N	12°13'W
		Ngane	*senegal*	3	14°08'N	16°12'W
Southern Africa	Botswana	[a]Tsau	*rostrata*	1	21°47'S	24°04'E
		[a]Mbeleapudi	*rostrata*	1	20°09'S	22°19'E
	South Africa	[a]Jozinidam	*rostrata*	3	27°32'S	31°58'E
		[a]Kamlushwa	*rostrata*	1	25°42'S	31°45'E
		[a]Sokwe,	*rostrata*	1	26°52'S	32°13'E
		[a]Zululand	*rostrata*	1	27°08'S	31°59'E
		[a]Pongolo	*rostrata*	1	27°25'S	32°04'E
		[a]Hlabisa	*rostrata*	1	28°08'S	31°52'E
Arabian Peninsula, Pakistan and India	India	[b]Jodhpur	[c]*senegal*	15	26°43'N	73°09'E
	India	Jodhpur (Inde50)	*senegal*	4	26°19'N	79°31'E
		Jodhpur (Inde60)	*senegal*	3	26°19'N	79°31'E
		[a]Old Delhi Ridge	*senegal*	1	28°42'N	77°13'E
	Oman	[a]Dhofar	*senegal*	3	16°52'N	53°47'E
	Pakistan	[a]Sind	*senegal*	2	25°19'N	68°04'E

[a]Herbarium specimens held by Royal Botanic Garden, Edinburgh (RBGE); further details are presented in Supplementary Table S1.

[b]Collections from World Agroforestry Centre, formerly ICRAF.

[c]Samples with uncertain intraspecific affiliations.

n, number of samples.

.

829 **Table 3** Haplotype diversity (within-population diversity h_S, v_S and total diversity h_T, v_T) and differentiation (G_{ST}, N_{ST} and R_{ST}) for *Acacia*
830 *senegal* populations across its range and within SAMOVA delineated groups 1 (mainly East Africa), 2 (mainly West Africa) and 3 (East and
831 Southern Africa). Number of populations, n; standard errors (SE) in parentheses.

Marker	Region	n	h_S	h_T	G_{ST}	v_S	v_T	N_{ST}
a) *trn*H-*psb*A	Mainly East Africa	21	0.162 (0.0487)	0.545 (0.0838)	0.703 (0.0924)	0.159 (0.0530)	0.545 (0.0904)	0.709 (0.0987), < 0.736 (permuted)
	Mainly West Africa	16	0.138 (0.0610)	0.584 (0.0534)	0.764 (0.1100)	0.146 (0.0646)	0.584 (0.0534)	0.750 (0.1121), < 0.806 (permuted)
	East and Southern Africa	3	0.200 (0.2000)	0.667 (0.3289)	0.700 (NC)[a]	0.200 (0.2000)	0.667 (0.3289)	0.850 (NC)[a], < 0.850 (permuted)
	Rangewide	40	0.155 (0.0370)	0.711 (0.0263)	0.782 (0.0514)	0.162 (0.0413)	0.711 (0.0421)	0.772 (0.0555), < 0.799 (permuted)
	East Africa	22	0.138 (0.0467)	0.523 (0.0972)	0.737 (0.0835)	0.120 (0.0433)	0.524 (0.1077)	0.771 (0.0782), = 0.771 (permuted)
	West Africa	11	0.155 (0.0787)	0.613 (0.0782)	0.747 (0.1358)	0.169 (0.0860)	0.612 (0.0608)	0.723 (0.1393), < 0.805 (permuted)
	Africa-wide	33	0.143 (0.0400)	0.693 (0.0406)	0.793 (0.0562)	0.150 (0.0443)	0.693 (0.0606)	0.784 (0.0589), < 0.807 (permuted)

Marker	Region	n	h_S	h_T	G_{ST}	v_S	v_T	R_{ST}
b) cpSSR	Mainly East Africa	21	0.641 (0.0646)	0.903 (0.0255)	0.290 (0.0726)	0.779 (0.5359)	0.901 (0.3768)	0.136 (0.2649), < 0.437 (permuted)
	Mainly West Africa	15	0.501 (0.0869)	0.925 (0.0203)	0.458 (0.0927)	0.287 (0.1039)	0.937 (0.2443)	0.694 (0.0383) , > 0.641 (permuted) P = 0.05
	East and Southern Africa	3	0.500 (0.2646)	0.833 (0.0912)	0.400 (0.4243)	0.452 (0.2609)	0.853 (0.0639)	0.471 (0.4404), < 0.759 (permuted)
	Rangewide	39	0.576 (0.0516)	0.948 (0.0092)	0.392 (0.0537)	0.312 (0.1842)	0.954 (0.1773)	0.673 (0.1555), > 0.576 (permuted) P = 0.01

832 [a]NC, not computed due to small sample size.
833
834
835
836
837
838
839
840
841

Supplementary Information Tables S1 – S7

843

844 **Table S1** Geographical origin, variety, sample size and climatic conditions of *Acacia senegal* samples used in this study

845

Region and Country	[a]Population	Variety	n	Latitude	Longitude	Alt (masl)	Precipitation mm/year	Mean Temp (°C)	[b]Monthly rainfall pattern
East Africa									
Ethiopia	Sodera	*senegal*	3	08°24'00.0"N	39°23'00.0"E	1393	778	20.7	6/10/2
Kenya	Archer's Post	*kerensis*	2	00°38'48.6"N	37°39'01.2"E	876	364	23.7	10/11/0
Kenya	Daaba	*kerensis*	2	00°32'00.2"N	37°45'39.9"E	940	332	23.7	10/11/0
Kenya	Kajiado	*senegal*	7	01°52'53.5"S	36°45'29.0"E	1741	693	19.1	7/9/2
Kenya	Kakuma	*kerensis*	5	03°45'26.1"N	34°39'59.5"E	670	535	24.0	7/11/0
Kenya	Kaleing'	*kerensis*	6	04°24'28.0"N	35°31'03.1"E	614	361	26.9	11/12/0
Kenya	Kargi	*kerensis*	1	02°38'35.8"N	37°27'34.7"E	454	859	20.0	6/9/2
Kenya	Kibwezi	*kerensis*	5	02°12'49.1"S	38°04'22.4"E	713	674	24.0	8/9/3
Kenya	Kulamawe	*leiorhachis*	6	00°32'21.5"N	37°59'41.5"E	1021	373	23.7	10/11/0
Kenya	Koriema	*senegal*	7	00°26'55.8"N	35°52'11.9"E	1348	1112	19.3	2/8/3
Kenya	Lokichar	*kerensis*	3	02°22'13.3"N	35°38'21.0"E	794	198	26.9	11/12/0
Kenya	Machakos, Yatta plains, P Evans, RBGE #1081	[c]*kerensis*	1	01°30'51.1"S	37°15'45.7"E	1820	856	19.2	5/8/2
Kenya	Magadi	*leiorhachis*	7	01°31'52.8"S	36°34'31.5"E	1460	588	22.2	7/10/2
Kenya	Marigat	*kerensis*	6	00°27'49.6"N	35°53'30.1"E	1348	641	23.8	5/12/0
Kenya	Merille	*kerensis*	3	01°31'40.8"N	37°73'29.6"E	655	859	20.0	6/9/2
Kenya	Ngarendare	*kerensis*	8	00°28'04.3"N	37°25'02.4"E	1005	580	23.3	6/9/1
Kenya	Ngurunit	*leiorhachis*	6	01°43'17.0"N	37°17'24.3"E	723	693	19.1	7/9/2
Kenya	Ntumburi	*senegal*	6	00°11'54.7"N	37°31'00.2"E	1694	1414	18.5	4/6/2
Kenya	Rimoi	*senegal*	7	00°40'08.8"N	35°33'47.0"E	1190	1355	15.8	1/6/5
Kenya	Serolipi	*kerensis*	5	01°15'16.1"N	37° 59'63.9"E	750	832	20.6	6/8/2

Kenya	Taita- Taveta	*senegal*	2	03°27'05.5"S	38°28'42.0"E	654	570	24.9	8/10/0
Somalia	Hargesia, JB Allen and AA Elmi, RBGE #264,	*kerensis*	1	09°30'00.0"N	44°03'00.0"E	1371	418	21.8	8/12/0
Somalia	Afgoi, JB Allen, RBGE #233	*[c]kerensis*	1	02°06'00.0"N	45°08'00.0"E	80	526	27.3	8/10/0
Tanzania	Kigwe	*leiorhachis*	15	06°06'00.0"S	35°29'00.0"E	980	566*	23.6	7/8/0
Tanzania	Wangingombe	*senegal*	8	08°51'00.0"S	34°38'00.0"E	1450	873*	22.2	8/8/3
Sudan	Fallatu Forest	*senegal*	12	13°06'00.0"N	30°08'24.0"E	570	365*	27.3	9/11/0
Sudan	Kordofan	*senegal*	6	12°44'00.0"N	29°35'00.0"E	510	365*	26.0	9/11/0
Sudan	Khartoum nursery, H Elamin, RBGE #1407	*senegal*	1	15°37'59.2"N	32°31'58.8"E	386	127	28.9	11/12/0
Sudan	Kundoura Forest Nyala, Southern Darfur, H Elamin, RBGE #1608	*senegal*	1	12°51'46.1"N	30°13'03.5"E	520	317	27.7	9/11/0
Sudan	Bora, H Elamin, RBGE #1605	*senegal*	1	08°49'01.0"N	26°10'48.0"E	540	1140	25.9	6/7/4
Sudan	Wad Medani, H Elamin, RBGE #30	*senegal*	1	14°23'58.8"N	33°31'12.9"E	400	343	28.6	9/10/0
Sudan	El Haraza, H Elamin, RBGE #1604	*senegal*	1	11°18'00.0"N	24°11'00.0"E	564	610	27.1	8/9/1
Sudan	Sobar, H Elamin, RBGE #1614	*senegal*	1	08°47'38.6"N	32°53'55.3"E	400	811	27.6	6/7/1
Sudan	Rodom, West Darfur, H Elamin, RBGE #1503	*senegal*	1	12°05'17.1"N	23°03'19.2"E	801	657	25.7	8/8/2
Sudan	Kassala nursery, H Elamin, RBGE #31	*senegal*	1	15°26'53.9"N	36°24'00.7"E	552	251	29.6	10/12/0
Sudan	El Felaya, South West Kordofan, H Elamin, RBGE #1553	*senegal*	1	09°35'41.6"N	28°26'20.4"E	403	841	27.3	7/8/3

Central Africa									
Cameroon	Villages of Mouda, Laf and Mousourtouk, near Maroua	*senegal*	10	10º15'00.0"N	14º14'00.0"E	448	790	27.9	7/8/3
Chad	Tourba	*senegal*	2	12°49'00.0"N	15°18'00.0"E	288	351	29.9	9/10/1
West Africa									
Burkina Faso	Bissiga Fc	*senegal*	6	12°26'00.0"N	00°32'00.0"W	308	748	28.5	7/9/2
Burkina Faso	Di (Sousou)	*senegal*	7	13°10'00.0"N	03°25'00.0"W	255	697	28.3	8/9/2
Mali	Aïte	*senegal*	4	15º05'00.0"N	11º39'00.0"W	71	547	30.3	8/9/1
Mali	Kirane	*senegal*	50	15º23'00.0"N	10°15'00.0"W	206	550*	28.9	9/11/0
Mali	Somo	*senegal*	5	13°17'00.0"N	04°54'00.0"W	284	700*	28.1	8/9/3
Mauritania	Djiguéni	*senegal*	6	15°44'00.0"N	08°40'00.0"W	200	200*	28.3	9/12/0
Mauritania	Kankossa	*senegal*	1	15°56'00.0"N	11°27'00.0"W	150	250*	29.5	9/11/0
Niger	Karofane	*senegal*	2	14°18'00.0"N	06°11'00.0"E	330	485	28.9	8/10/1
Senegal	Daïba	*senegal*	6	15°22'00. 0"N	13°08'00.0"W	48	458*	29.6	9/11/0
Senegal	Diaménar	*senegal*	6	16°00'00.0"N	15°54'00.0"W	12	284*	25.4	9/10/0
Senegal	Kidira	*senegal*	4	14°28'00.0"N	12°13'00.0"W	33	505*	28.9	8/9/2
Senegal	Ngane	*senegal*	3	14°08'00.0"N	16°12'00.0"W	8	712*	28.0	9/9/2
Southern Africa									
Botswana	Tsau, Ngamiland district, DG Long and DAH Rae, RBGE #246	*rostrata*	1	21°46'48.0"S	24°03'35.0"E	940	345	22.6	9/12/0

Botswana	Mbeleapudi Hills, Ngamiland district, DG Long and DAH Rae, RBGE #154	*rostrata*	1	20°08'42.0"S	22°19'19.0"E	940	445	22.6	8/11/0
South Africa	Turnoff to Jozinidam on Mkuze-Candover Rd, JH Ross, RBGE #1643	*rostrata*	1	27°31'34.0"S	31°58'26.2"E	165	500	22.6	6/12/0
South Africa	Jozinidam on Mkuze-Candover Rd, JH Ross, RBGE #1646	*rostrata*	1	27°31'16.9"S	31°58'40.6"E	171	500	22.6	6/12/0
South Africa	Kamlushwa district, near Mzinti, M Stalmans, RBGE #2175	*rostrata*	1	25°41'39.0"S	31°45'03.2"E	244	571	23.5	5/11/0
South Africa	Jozinidam, JH Ross, RBGE #1645	*rostrata*	1	27°31'16.9"S	31°58'40.6"E	131	500	22.6	6/12/0
South Africa	Sokwe, Ndumu Game Reserve, Ingwavuma district, ES Pooley RBGE #1297	*rostrata*	1	26°52'27.4"S	32°12'30.7"E	44	608	22.9	6/11/0
South Africa	Zululand, Ingwavuma district, ES Pooley, RBGE #1158	*rostrata*	1	27°07'55.9"S	31°59'39.2"E	44	608	22.9	6/11/0
South Africa	Pongolo, JH Ross, RBGE #1702	*rostrata*	1	27°25'05.5"S	32°04'17.6"E	340	500	22.6	7/12/0
South Africa	Hlabisa, CJ Ward, RBGE #5614	*rostrata*	1	28°08'22.3"S	31°51'57.6"E	300	1073	19.9	4/6/5
Arabian Peninsula, Pakistan									

and India									
India	Jodhpur ICRAF general collection	[c]*senegal*	15	26°42'54.4"N	73° 08'31.4"E	357	402	26.5	9/10/0
India	Jodhpur Inde50	*senegal*	4	26°19'00.0"N	79°31'00.0"E	210	300*	25.8	7/8/3
India	Jodhpur Inde60	*senegal*	3	26°19'00.0"N	79°31'00.0"E	210	300*	25.8	7/8/3
India	Old Delhi Ridge, S. Jalan and U Singh, RBGE, date of collection 5 October 1959	*senegal*	1	28°42'00.0"N	77°13'00.0"E	216	792	25.0	8/9/2
Oman	Dhofar, Mughsayl, AG Miller, RBGE #7674	*senegal*	1	16°52'22.9"N	53°47'13.2"E	128	130	25.9	12/12/0
Oman	Dhofar, Lejer Waterhole, AG Miller, RBGE #6232	*senegal*	1	17°10'59.9"N	54°55'59.9"E	598	109	25.6	12/12/0
Oman	Dhofar, Lejer Waterhole, RBGE, AG Miller, RBGE #7236	*senegal*	1	17°06'00.0"N	55°04'59.9"E	41	109	25.6	12/12/0
Pakistan	Sind, Off Thano-Bula-Kotri Highway, J Lamond, RBGE #811	*senegal*	1	25°19'28.9"N	68°03'58.9"E	68	175	27.8	10/12/0
Pakistan	Sind, Off Budhapur-Menjhand Rd, J Lamond, RBGE #826	*senegal*	1	25°41'23.4"N	68°19'01.5"E	31	175	27.8	10/10/0

846   [a]Populations collected during this study and/or herbarium specimens held by Royal Botanic Garden, Edinburgh (RBGE), UK with collectors' names and
847   voucher numbers where available.
848   [b]Monthly rainfall patterns: first number = no. of months with < 50 mm, second number = no. of months with < 100 mm, last number: no. of months
849   precipitation > potential evapo-transpiration; e.g. 9/10/0 indicates 9 months < 50mm, 10 months < 100mm and 0 = precipitation deficit throughout the year.
850   Precipitation and temperature: *****Data obtained from actual site; the rest are obtained with Local climate estimator (New-**Loc**Clim, FAO 2005).
851   [c]Samples with uncertain intraspecific affiliations.
852

853 **Table S2** Details of ITS sequenced samples.
854

No.	Country of origin	Variety	Provenance	Haplotype number	GenBank number
1	Tanzania	*leiorhachis*	Kigwe	1	HQ605042
2	Kenya	*kerensis*	Kaleing	2	HQ605043
3	Tanzania	*leiorhachis*	Kigwe	3	HQ605044
4	India	^a*senegal*	India (ICRAF)	4	HQ605045
5	Pakistan	*senegal*	Sind, RBGE #811	5	HQ605046
6	Pakistan	*senegal*	Sind, RBGE #826	5	HQ605046
7	Kenya	*kerensis*	Kakuma	6	HQ605047
8	Kenya	*kerensis*	Merille	7	HQ605048
9	Kenya	*senegal*	Taita Taveta	8	HQ605049
10	Mali	*senegal*	Kirane	9	HQ605050
11	Kenya	*senegal*	Kajiado	10	HQ605051
12	Mali	*senegal*	Kirane	11	HQ605052
13	Mauritania	*senegal*	Kankoussa	12	HQ605053
14	Mali	*senegal*	Kirane	13	HQ605054
15	Burkina Faso	*senegal*	Di (Sousou)	13	HQ605054
16	Burkina Faso	*senegal*	Di (Sousou)	13	HQ605054
17	Burkina Faso	*senegal*	Bissiga Fc	13	HQ605054
18	India	*senegal*	Jodhpur Inde50	13	HQ605054
19	Mali	*senegal*	Somo	13	HQ605054
20	Mauritania	*senegal*	Djiguéni	13	HQ605054
21	Mauritania	*senegal*	Djiguéni	13	HQ605054
22	Niger	*senegal*	Karofane	13	HQ605054
23	Senegal	*senegal*	Daiba	13	HQ605054
24	Senegal	*senegal*	Kidira	13	HQ605054
25	Senegal	*senegal*	Ngane	13	HQ605054
26	Chad	*senegal*	Tourba	13	HQ605054
27	Chad	*senegal*	Tourba	13	HQ605054
28	Sudan	*senegal*	RBGE #30	13	HQ605054
29	Kenya	*kerensis*	Kaleing	14	HQ605055
30	Sudan	*senegal*	RBGE #1407	15	HQ605056
31	Sudan	*senegal*	Kordofan	16	HQ605057
32	Sudan	*senegal*	RBGE #31	16	HQ605057
33	India	*senegal*	Jodhpur Inde50	17	HQ605058
34	Sudan	*senegal*	Kordofan	17	HQ605058
35	Sudan	*senegal*	RBGE #1553	17	HQ605058
36	Sudan	*senegal*	Fallatu Forest	17	HQ605058
37	Mali	*senegal*	Kirane	18	HQ605059
38	Cameroon	*senegal*	Maroua	19	HQ605060
39	Cameroon	*senegal*	Maroua	19	HQ605060
40	Senegal	*senegal*	Diamenar	20	HQ605061
41	Kenya	*senegal*	Kibwezi	21	HQ605062
42	Mali	*senegal*	Aite	22	HQ605063
43	Mali	*senegal*	Aite	22	HQ605063
44	Senegal	*senegal*	Daiba	22	HQ605063
45	Senegal	*senegal*	Diamenar	22	HQ605063
46	Senegal	*senegal*	Kidira	22	HQ605063
47	Senegal	*senegal*	Ngane	22	HQ605063

48	India	*senegal*	Jodhpur Inde60	23	HQ605064
49	India	*senegal*	Jodhpur Inde60	23	HQ605064
50	Mauritania	*senegal*	Kankoussa	24	HQ605065
51	Kenya	*kerensis*	Kakuma	25	HQ605066
52	Kenya	*kerensis*	Serolipi	25	HQ605066
53	Kenya	*kerensis*	Serolipi	26	HQ605067
54	Kenya	*senegal*	Kibwezi	26	HQ605067
55	South Africa	*rostrata*	RBGE #1158	27	HQ605068
56	South Africa	*rostrata*	RBGE #1643	28	HQ605069
57	South Africa	*rostrata*	RBGE #2175	29	HQ605070
58	Kenya	*kerensis*	Ngurunit	30	HQ605071
59	Kenya	*kerensis*	Ngurunit	31	HQ605072
60	Kenya	*senegal*	Ntumburi	32	HQ605073
61	Kenya	*senegal*	Ntumburi	32	HQ605073
62	Tanzania	*senegal*	Wangingombe	33	HQ605074
63	Niger	*senegal*	Karofane	34	HQ605075
64	Kenya	*leiorhachis*	Magadi	35	HQ605076
65	Botswana	*rostrata*	RBGE #246	36	HQ605077

[a]Samples with unverified intraspecific affiliations.

Table S3 Description of *trn*H-*psb*A haplotypes and their distribution within populations of *Acacia senegal*. Polymorphic fragments are presented as presence (1) or absence (2). *trn*H-*psb*A was digested restriction enzyme *Dra*I. RBGE represents sources from Royal Botanic Garden, Edinburgh, UK.

Haplotype description

Haplotype	*Polymorphic fragments*								
1	1	1	2	2	2	2	2	2	2
2	2	2	1	1	2	2	2	2	2
3	2	1	2	2	1	2	2	2	2
4	2	2	2	2	2	1	1	2	2
5	2	2	2	2	2	2	1	1	2
6	2	1	2	2	2	2	2	2	1

Haplotype distribution

Region and Country		*Haplotypes*						
East Africa	*[a]Population*	1	2	3	4	5	6	Total
Ethiopia	Sodera			3				3
Kenya	Archer's Post	2						2
Kenya	Daaba	2						2
Kenya	Kajiado	7						7
Kenya	Kakuma	4	1					5
Kenya	Kaleing	6						6
Kenya	Kargi	1						1
Kenya	Kibwezi	1		4				5
Kenya	Koriema	6		1				7
Kenya	Kulamawe	6						6
Kenya	Lokichar	3						3
Kenya	Magadi	7						7

49

Country	Location					
Kenya	Marigat	5		1		6
Kenya	Merille	3				3
Kenya	Ngarendare	8				8
Kenya	Ngurunit	6				6
Kenya	Ntumburi	6				6
Kenya	Machakos, RBGE #1801	1				1
Kenya	Rimoi	7				7
Kenya	Serolipi	5				5
Kenya	Taita Taveta	2				2
Somalia	Hargesia, RBGE #264		1			1
Somalia	Afgoi, RBGE #233			1		1
Tanzania	Kigwe	11	4			15
Tanzania	Wangingombe		8			8
Sudan	Fallatu Forest		12			12
Sudan	Kordofan		6			6
Sudan	Bora, RBGE #1605	1				1
Sudan	Kundoura Forest, RBGE #1608		1			1
Sudan	El Haraza, RBGE #1604	1				1
Sudan	Sobar, RBGE #1614		1			1
Sudan	Wad Medani, RBGE #30		1			1
Sudan	Kassala nursery, RBGE #31		1			1
Sudan	Rodom, West Darfur RBGE #1503		1			1
Cameroon	Maroua				10	10
Chad	Tourba		2			2
Burkina Faso	Bissiga Fc			6		6
Burkina Faso	Di (Sousou)			7		7
Mauritania	Aite		3	1		4

Country	Location					Total
Mauritania	Djiguéni		6			6
Mauritania	Kankossa		1			1
Mauritania	Kirane	1	49			50
Mauritania	Somo		5			5
Niger	Karofane		2			2
Senegal	Daiba	6				6
Senegal	Diamenar	6				6
Senegal	Kidira	3	1			4
Senegal	Ngane	2	1			3
Botswana	Tsau, Ngamiland district, RBGE #246		1			1
Botswana	Mbeleapudi Hills, Ngamiland district, RBGE #154	1				1
South Africa	Jozidam on Mkuze-Candover Rd, RBGE #1646			1		1
South Africa	Kamlushwa district, near Mzinti, RBGE #2175			1		1
South Africa	Jozinidam, RBGE #1645			1		1
South Africa	Ndumu Game Reserve, Ingwavuma district, RBGE #1297			1		1
South Africa	Zululand, Ingwavuma district, RBGE #1158			1		1
South Africa	Pongolo, RBGE #1702			1		1
South Africa	Hlabisa, RBGE #5614	1				1
India	India Jodhpur ICRAF general collection	12	1		2	15
India	Jodhpur Inde50	3	1			4
India	Jodhpur Inde60	3				3

Country	Location							
India	Old Dehli Ridge, RBGE, date of collection 5 October 1959	1						1
Oman	Dhofar, Mughsayl, RBGE #7674					1		1
Oman	Dhofar, Lejer Waterhole, RBGE #6232	1						1
Oman	Dhofar, Lejer Waterhole, RBGE #7236	1						1
Pakistan	Sind, Off Thano-Bula-Kotri Highway, RBGE #811	1						1
Pakistan	Sind, Off Budhapur-Menjhand Rd, RBGE #826	1						1
		106	80	92	3	4	12	297

862 [a]Population: RBGE (Royal Botanic Garden, Edinburgh, UK); represents individual herbarium specimens.

52

863

864 **Table** S4 Haplotype diversity (within-population diversity h_S, v_S and total diversity h_T, v_T) and differentiation (G_{ST}, N_{ST} and R_{ST}) for *Acacia*

865 *senegal* populations in Africa (East, including Southern Africa, and West, including Central Africa). Number of populations, n; standard

866 errors (SE) in parentheses.

Marker	Region	n	h_S	h_T	G_{ST}	v_S	v_T	N_{ST}
a) *trn*H-*psb*A	East Africa, including Southern Africa	22	0.138 (0.0467)	0.523 (0.0972)	0.737 (0.0835)	0.120 (0.0433)	0.524 (0.1077)	0.771 (0.0782), = 0.771 (permuted)
	West Africa, including Central Africa	11	0.155 (0.0787)	0.613 (0.0782)	0.747 (0.1358)	0.169 (0.0860)	0.612 (0.0608)	0.723 (0.1393), < 0.805 (permuted)
	Africa-wide	33	0.143 (0.0400)	0.693 (0.0406)	0.793 (0.0562)	0.150 (0.0443)	0.693 (0.0606)	0.784 (0.0589), < 0.807 (permuted)
								R_{ST}
b) cpSSR	East Africa, including Southern Africa	22	0.657 (0.0583)	0.914 (0.0245)	0.281 (0.0607)	0.216 (0.0530)	0.931 (0.3165)	0.768 (0.0836), > 0.486 (permuted) (p=0.01)
	West Africa, including Central Africa	11	0.509 (0.0937)	0.918 (0.0264)	0.445 (0.1046)	0.306 (0.1365)	0.935 (0.3111)	0.673 (0.0317), > 0.656 (permuted)t (p=0.05)
	Africa-wide	33	0.607 (0.0505)	0.945 (0.0109)	0.357 (0.0532)	0.153 (0.0328)	0.957 (0.1637)	0.840 (0.0370), > 0.531 (permuted) (p=0.01)

867 [a]NC, not computed due to small sample size.

868

869

870

Table S5 Distribution of cpSSR haplotypes within populations of *Acacia senegal*.

l

[a]Region and Population	Haplotypes																																				Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	
East Africa																																					
Sodera														2					1																		3
Archer's Post														1					1																		2
Daaba										2																											2
Kajiado														1				3	3																		7
Kakuma														3					1					1													5
Kaleing										2				2	1				1																		6
Kargi														1																							1
Kibwezi															4						1																5
Koriema					1	2			1	2				1																							7
Kulamawe						6																															6
Lokichar						1														2																	3
Magadi													3	1				2	1																		7
Marigat						3				3																											6
Merille							1			1	1																										3
Ngarendare										6	1			1																							8
Ngurunit										2					3									1													6
Ntumburi			3	3																																	6
Machakos, [b]RBGE #1081					1																																1
Rimoi										7																											7
Serolipi										4					1																						5
Taita Taveta													1						1																		2
Hargesia RBGE #264														1																							1
Kigwe										3		1	2	5	4																						15
Wangingombe										1				2	3		2																				8
Fallatu Forest																												2	2	1	2	2	1	2			12
Kordofan																									2		4										6
Bora, RBGE #1605																							1														1
Kundoura Forest Nyala, RBGE #1608						1																															1
El Haraza, RBGE #1604							1																														1
Sobar, RBGE #1614																				1																	1
Kassala nusery, RBGE #31																						1															1
Rodom, West Darfur RBGE #1503																						1															1

Region / Location	Values	Total
Central Africa		
Maroua	1 9	10
Tourba	1 1	2
West Africa		
Bissiga Fc	3 ... 1 2	6
Di (Sousou)	7	7
Aite	3 1	4
Somo	1 ... 2 1 ... 1	5
Kirane	1 1 ... 4 1 21 6 ... 7 9	50
Djigueni	6	6
Kankossa	1	1
Karofane	2	2
Daiba	3 ... 3	6
Diamenar	2 3 ... 1	6
Kidira	1 3	4
Ngane	1 ... 2	3
Southern Africa		
Jozinidam on Mkuze-Candover Rd, RBGE #1646	1	1
Kamlushwa district, near Mzinti, RBGE #2175	1	1
Jozinidam, RBGE #1645	1	1
Ndumu Game Reserve, Ingwavuma district,, RBGE #1297	1	1
Zululand, Ingwavuma district, RBGE #1158,	1	1
Arabian Peninsula, Pakistan and India		
India ICRAF general collection	2 7 ... 1 1 ... 3 1	15
Jodhpur Inde50	4	4
Jodhpur Inde60	3	3
Mughsayl, RBGE #7674	1	1
Lejer Waterhole, RBGE #2462	1	1
Lejer Waterhole, RBGE #6232	1	1
Sind, Off Thano-Bula-Kotri Highway, RBGE # 811	1	1
Budhapur-Menjhand Rd, RBGE #826	1	1

	1	1	5	4	2	14	3	2	1	33	2	1	7	32	13	1	3	5	17	7	16	11	1	22	11	25	17	2	2	1	11	12	1	2	1	1	290

882 **Table S6** cpSSR SAMOVA delineated groups and affiliations to geographic region, putative variety and *trn*H-*psb*A haplotypes from 40
883 populations of *Acacia senegal.*
884

SAMOVA group and geographic region	Population	Variety	Country	*trn*H-*psb*A haplotypes	ITS lineages
1. Mainly East Africa	Sodera	*senegal*	Ethiopia	3	nd[4]
	[a]Sudan	*senegal*	Sudan	1, 2	IV-3
	Kajiado	*senegal*	Kenya	1	IV-3
	Kakuma	*kerensis*	Kenya	1, 2	IV-3
	Kaleing'	*kerensis*	Kenya	1	IV-3
	Kibwezi	*kerensis*	Kenya	1, 3	IV-3
	Koriema	*senegal*	Kenya	1, 3	nd[4]
	Lokichar	*kerensis*	Kenya	1	nd[4]
	Magadi	*leiorhachis*	Kenya	1	III
	Marigat	*kerensis*	Kenya	1, 3	nd[3]
	Merille	*kerensis*	Kenya	1	IV-3
	Ngarendare	*kerensis*	Kenya	1	nd[3]
	Ngurunit	*leiorhachis*	Kenya	1	II
	Rimoi	*senegal*	Kenya	1	
	Serolipi	*kerensis*	Kenya	1	IV-3
	Kigwe	*leiorhachis*	Tanzania	1, 2	III
	Wangingombe	*senegal*	Tanzania	2	IV-3
	Diamenar	*senegal*	Senegal	2	IV-3
	Tourba	*senegal*	Chad	2	IV-3
	Jodhpur (ICRAF)	[b]*senegal*	India	2, 3, 6	III
	[a]Oman	*senegal*	Oman	1, 5	nd[4]
	[a]Sudan	*senegal*	Sudan	1, 2	

2. Mainly West Africa	Bissiga	*senegal*	Burkina Faso	3	IV-3
	Di (Sousou)	*senegal*	Burkina Faso	3	IV-3
	Aite	*senegal*	Mauritania	2, 3	IV-3
	Djigueni	*senegal*	Mauritania	3	IV-3
	Kirane	*senegal*	Mauritania	2, 3	IV-3
	Somo	*senegal*	Mauritania	3	IV-3
	Karofane	*senegal*	Niger[3]	3	IV-3
	Daiba	*senegal*	Senegal	2	IV-3
	Kidira	*senegal*	Senegal	2, 3	IV-3
	Ngane	*senegal*	Senegal	2, 3	IV-3
	Maroua	*senegal*	Cameroon	6	IV-3
	Kordofan	*senegal*	Sudan	2	IV-3
	Fallatu	*senegal*	Sudan	2	IV-3
	Jodhpur50	*senegal*	India	2, 3	IV-3
	Jodhpur60	*senegal*	India	2	IV-3
3. East and Southern Africa	Kulamawe	*leoirhachis*	Kenya	1	[d]ND
	Ntumburi	*senegal*	Kenya	1	IV-3
	[a]Southern Africa	*rostrata*	[c]Botswana	2, 4, 5	IV-2

885

886   [a]Quasi-populations from country or regionally located herbarium specimens as listed in Table 1; [b]Samples with unverified intraspecific affiliations; [c]Contains
887   individuals forming the basal group (see Figure 1); [d]ND, not determined/not sequenced.

888

Table S7 Hierarchical analysis of genetic differentiation for rangewide and pairwise comparisons between groups using RFLP and cpSSR markers. SAMOVA delineated Groups 1, 2 and 3, namely; East Africa, mainly West Africa, and East and Southern Africa, respectively. Number of populations, *n*.

Region/group	*n*	RFLP (*trn*H-*psb*A)			cpSSR		
		F_{ST}	F_{SC}	F_{CT}	F_{ST}	F_{SC}	F_{CT}
Rangewide	39	0.792	0.754	0.156	0.871	0.389	0.789
Group 1 vs 2	21 vs 15	0.815	0.772	0.187	0.855	0.385	0.765
Group 1 vs 3	21 vs 3	0.603	0.620	-0.044	0.662	0.218	0.568
Group 2 vs 3	15 vs 3	0.850	0.828	0.127	0.964	0.603	0.909

Figure legend

Figure 1 (a) Bayesian 50 % majority-rule consensus phylogram of ITS sequences from *Acacia senegal* individuals sampled across its distribution range. Branches are labelled with ≥ 70 % bootstrap support (above) and posterior probability (below) values. Arrow indicates presence of a single clade in the maximum parsimony strict consensus tree. The tree was rooted with *Vachellia farnesiana* and *Vachellia collinsii* (see inset tree, double line denotes the point at which the break is in the main tree) obtained via GenBank (accession numbers EF638219 and EF638216, respectively). *Acacia* [*Senegalia*] *senegal* sequence is accession number EF638213 from Zimbabwe. Each haplotype is code-labelled HX where X refers to a number 1-36 as described in Table S1. Hypothesised colonisation or range expansion events are indicated. Scale bar signifies 0.1 substitutions per nucleotide site.

Figure 2 Rangewide distribution of nuclear (**a**) and chloroplast (**b**) haplotypes in *Acacia senegal*. Data are: (**a**) haplotypes from internal transcribed spacer (ITS) of the nuclear ribosomal DNA and (**b**) PCR-RFLP haplotypes of chloroplast *trn*H-*psb*A intergenic spacer of *Acacia senegal*; each circle represents a single population with colour denoting haplotype variation. The number of samples per population is presented in Supplementary Tables S2 and S3, respectively. Inset: approximate distribution of *A. senegal* in Africa (hatch-shaded area).

Figure 2
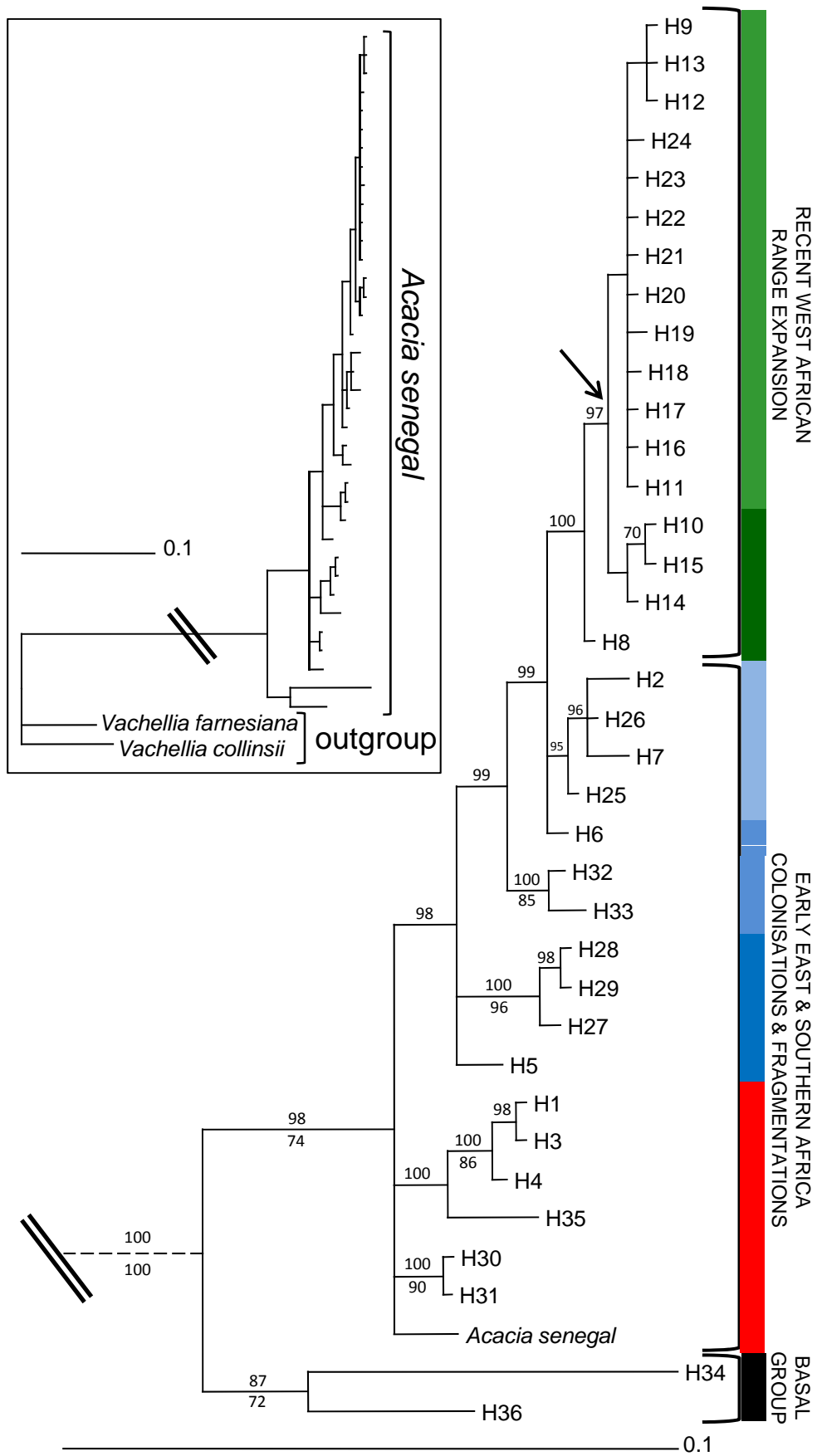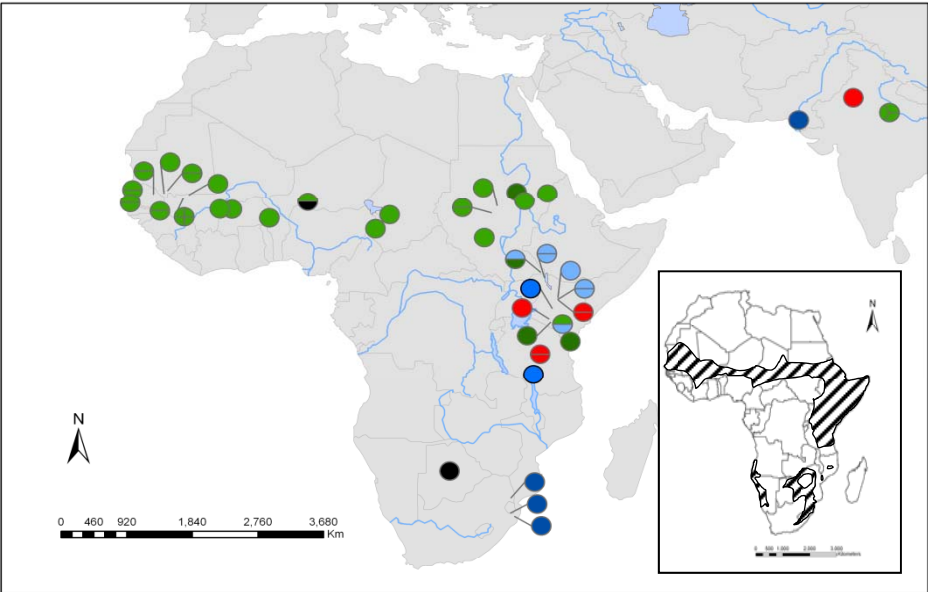
(a)

(b)