RESEARCH ARTICLE

AMERICAN JOURNAL OF BOTANY

Phylogeography of the wild and cultivated stimulant plant qat (*Catha edulis*, Celastraceae) in areas of historical cultivation[1]

Luke R. Tembrock[2], Mark P. Simmons[2,13], Christopher M. Richards[3], Patrick A. Reeves[3], Ann Reilley[3], Manuel A. Curto[4,5], Harald Meimberg[5], Grace Ngugi[6], Sebsebe Demissew[7,8], Abdul Wali Al Khulaidi[9], Mansoor Al-Thobhani[10], Sheron Simpson[11], and Daniel M. Varisco[12]

PREMISE OF THE STUDY: Qat (*Catha edulis*, Celastraceae) is a woody plant species cultivated for its stimulant alkaloids. Qat is important to the economy and culture in large regions of Ethiopia, Kenya, and Yemen. Despite the importance of this species, the wild origins and dispersal of cultivars have only been described in often contradictory historical documents. We examined the wild origins, human-mediated dispersal, and genetic divergence of cultivated qat compared to wild qat.

METHODS: We sampled 17 SSR markers and 1561 wild and cultivated individuals across the historical areas of qat cultivation.

KEY RESULTS: On the basis of genetic structure inferred using Bayesian and nonparametric methods, two centers of origin in Kenya and one in Ethiopia were found for cultivated qat. The centers of origin in Ethiopia and northeast of Mt. Kenya are the primary sources of cultivated qat genotypes. Qat cultivated in Yemen is derived from Ethiopian genotypes rather than Yemeni wild populations. Cultivated qat with a wild Kenyan origin has not spread to Ethiopia or Yemen, whereas a small minority of qat cultivated in Kenya originated in Ethiopia. Hybrid genotypes with both Ethiopian and Kenyan parentage are present in northern Kenya.

CONCLUSIONS: Ethiopian cultivars have diverged from their wild relatives, whereas Kenyan qat has diverged less. This pattern of divergence could be caused by the extinction of the wild-source qat populations in Ethiopia due to deforestation, undersampling, and/or artificial selection for agronomically important traits.

KEY WORDS *Catha edulis*; centers of origin; clonal cultivation; East Africa; microsatellite markers; plant domestication

Over time, artificial selection produces plant cultivars that diverge morphologically and genetically from their wild progenitors (Harlan, 1992). Eventually, this process of divergence will produce cultivars that cannot survive without human intervention. This continuum of divergence is known as domestication, and those cultivars that cannot survive without human intervention are considered fully domesticated (Harlan, 1992; Miller and Gross, 2011). While the processes that produce this divergence are well studied among annual grain crops, they are far less well characterized among long-lived perennial crops (Hillman and Davies, 1990; Allaby et al., 2008; McKey et al., 2010; Miller and Gross, 2011).

There are two opposing theories regarding divergence of cultivars in long-lived perennial species. First, selection of individuals from the wild occurs once or a few times. These individuals possess desired traits that occurred from rare mutations in the wild and are subsequently maintained in cultivation via clonal propagation (e.g., Zohary and Spiegelroy, 1975; Harlan, 1992; Zohary, 2004;

---

[1] Manuscript received 16 December 2016; revision accepted 2 March 2017.

[2] Department of Biology, Colorado State University, Fort Collins, Colorado 80523, USA;
[3] United States Department of Agriculture, Agricultural Research Service, National Center for Genetic Resources Preservation, 1111 South Mason Street, Fort Collins, Colorado 80521, USA;
[4] CIBIO, Research Center in Biodiversity and Genetic Resources / InBio Associated Laboratory, University of Porto, Campus Vairão, Vairão 4485-661, Portugal;
[5] Institute for Integrative Nature Conservation Research, University of Natural Resources and Life Sciences, A-1180 Vienna, Austria;
[6] East African Herbarium, National Museums of Kenya, P.O. Box 40658-00100 Nairobi, Kenya;
[7] Department of Plant Biology and Biodiversity Management, College of Natural Sciences, Addis Ababa University, P.O. Box 3434, Addis Ababa, Ethiopia;
[8] Gullele Botanic Garden, P.O. Box 153/1029, Addis Ababa, Ethiopia;
[9] Department of Biology, College of Science and Art, Al Baha University, Baljurashi, Saudi Arabia;
[10] Department of Horticulture and Forestry, College of Agriculture, Sana'a University, Yemen;
[11] U.S. Department of Agriculture, Agricultural Research Service, Genomics and Bioinformatics Research Unit, 141 Experiment Station Road, Stoneville, Mississippi 38776, USA; and
[12] Center for Humanities and Social Sciences, Qatar University, Doha, Qatar
[13] Author for correspondence (e-mail: psimmons@rams.colostate.edu)

doi:10.3732/ajb.1600437

Kislev et al., 2006). This process is often referred to as "instant domestication" because of the limited time needed to produce a domesticated cultivar (Harlan, 1992). Alternatively, numerous individuals from several wild populations are selected, followed by hybridization both among cultivars and between cultivars and wild populations. In this scenario, both natural and artificial selection act upon a large number of genotypes at the early stages of domestication (e.g., Coart et al., 2006; Miller and Schaal, 2006; Achtak et al., 2010; Perrier et al., 2011; Bouby et al., 2013). This process has been documented in cassava, wherein seedling recruitment, interspecific gene flow, disease, and traditional cropping systems have all contributed to patterns of genetic diversity (e.g., Elias et al., 2001; Montero-Rojas et al., 2011; Bredeson et al., 2016).

Two processes of divergence between wild and cultivated lineages correspond to these alternative theories of long-lived perennial domestication. First, divergence may accumulate via natural selection and genetic drift that affects the wild populations only, while the clonal cultivars remain essentially unchanged. Second, divergence may occur by selection and drift acting on both wild and cultivated lineages. In clonally propagated crops, the change from sexual to asexual mating system may be an important means by which divergence is reinforced between wild and cultivated populations (McKey et al., 2010). It is thought that highly heterozygous genotypes are inherently selected for founding clonal lines because these genotypes show hybrid vigor (Balloux et al., 2003; Glémin et al., 2006).

Qat (Celastraceae: *Catha edulis* [Vahl] Forssk. ex Endl.) is a woody evergreen crop of major economic and cultural importance in East Africa and southwest Arabia that is grown for its stimulant alkaloids. An estimated 20 million people consume qat daily in eastern Africa (Al-Motarreb et al., 2002). Qat is second only to coffee as a traded commodity in Ethiopia (Bhalla, 2002; Gebissa, 2004; Anderson et al., 2007). Six percent of the gross domestic product in Yemen is generated from qat cultivation and sales (World Bank, 2007). Qat was estimated to be worth $100 million to the Kenyan economy (Chemitei, 2015), with >70,000 ha of cultivation in the Igembe region of Meru County alone (Baariu and Mulaku, 2015). In addition to local and regional consumption of qat, international export has expanded since the 1950s (e.g., Goldsmith, 1999; Gebissa, 2004; Carrier and Gezon, 2009). Despite the cultural and economic importance of qat, historical and oral accounts of its origin are often contradictory. The origin of cultivated qat has not been addressed using molecular markers.

In wild forest environments (Fig. 1A), qat grows as a woody shrub or tree (Fig. 1B) 2–25 m tall (Robson et al., 1994). Qat grows on rocky outcrops and at the margins of evergreen submontane forests (1100–2400 m; Robson et al., 1994). Qat ranges from Ethiopia to South Africa, and west to Angola (Palgrave, 1983; Robson et al., 1994). Flowers of qat are bisexual (Fig. 1E). Clonal growth is common in the wild but is limited to the area immediately adjacent to the parental plant. Wind-dispersed seeds largely account for dispersal of qat.

Cultivated qat is primarily propagated by rooted cuttings (e.g., Kennedy, 1987; Gebissa, 2004; Carrier, 2007) and trained into numerous forms (Fig. 1C; e.g., Kennedy, 1987; Carrier, 2007) for the purpose of encouraging growth and simplifying harvest (Fig. 1D). Qat cultivation takes place at scales ranging from backyard gardens to monocropped orchards (particularly in Yemen), but it is most typically grown in Africa on small-scale farms where it is often intercropped with maize, potatoes, beans, ensete, and other food crops (Fig. 1C).

Qat is among a small number of plant species that are widely cultivated for their stimulant properties, including coca (*Erythroxylum coca* and *E. novagranatense*), coffee (*Coffea arabica* and *C. canephora*), and tea (*Camellia sinensis*). It is the only such species in Celastraceae that is cultivated on a large scale for its stimulant properties. Qat is cultivated and/or collected from the wild and consumed in Ethiopia, Israel, Kenya, Madagascar, Rwanda, Somalia, Tanzania, Uganda, and Yemen (e.g., Kennedy, 1987; Gebissa, 2004; Anderson et al., 2007; Carrier, 2007; Carrier and Gezon, 2009). It is generally consumed by chewing fresh young leaves and shoots that contain the stimulant alkaloids: (S)–cathinone (United Nations Narcotics Laboratory, 1975; Schorno and Steinegger, 1979; Szendrei, 1980), (1R,2S)–norephedrine, and (1S,2S)–norpseudoephedrine, which is commonly referred to as cathine (Wolfes, 1930). Because of the rapid conversion of cathinone into the less potent cathine after harvest, qat is treated as a perishable commodity and is usually consumed within two to three days after harvest, with the highest value being given to the freshest leaves (e.g., Altabachew et al., 2013). The medicinal properties of qat were discussed in Arabic texts relating to Yemen from the early 16th century (Varisco, 2012).

The oldest claimed, but doubtful, historical references to qat are from the year 1237, suggesting that qat originated from Kilwa (Tanzania) and Yemen (El Mahi, 1962). The earliest credible reference to qat cultivation in Ethiopia is in the chronicles of King ʿAmda Syon I, who ruled Ethiopia from 1314 to 1344 A.D., wherein Sultan Sabra al-Din proclaimed that he would plant qat in the Ethiopian city of Mar'ade (Huntingford, 1965). The first known historical claim for qat cultivation in Yemen is from the early 14th century, in which the Rasulid Sultan al-Malik al-Dāwud is said to have praised qat cultivation. But the first unambiguous reference in Yemen to the term *qat* is over a century later (Varisco, 2007). On the basis of these historical records, qat has been cultivated for >600 yr in Ethiopia and Yemen. In addition to the Ethiopian and/or Yemeni origin(s) of qat use, Goldsmith (1988, 1994, 1999) and Carrier (2007) suggested an independent origin of qat (known as *miraa* in Kenya) cultivation in central Kenya, with old trees found in cultivation by late-19th-century European explorers (Carrier and Gezon, 2009). This relatively recent initial and ongoing cultivation of qat makes it an ideal study species for understanding the early stages of domestication for long-lived perennials, a stage no longer observable for many long-lived perennial crop species (e.g., Ladizinsky, 1999; Denham et al., 2003).

Historical references suggest three hypotheses for the origin of cultivated qat. First, qat cultivation originated in Yemen and from there moved to Ethiopia and Kenya (El Mahi, 1962; Revri, 1983). Second, qat cultivation originated in Ethiopia and then spread to both Yemen and Kenya (Huntingford, 1965). Third, qat cultivation had an independent origin in Kenya distinct from that in Ethiopia and Yemen (Goldsmith, 1994; Carrier, 2007). Thus, the objectives of the present study were to determine the wild origin(s) of cultivated qat within the principal historical regions of cultivation, quantify dispersal of cultivated genotypes, and deduce whether divergent cultivars have evolved since the initial cultivation(s) of qat.

MATERIALS AND METHODS

Plant collections—A total of 659 silica-gel-preserved leaf specimens were collected from Ethiopia and 822 from Kenya. Of these 1481 specimens, 567 cultivated specimens were sampled from 256

FIGURE 1 Wild and cultivated qat. (A) Ethiopian forest habitat from which wild qat was collected. (B) Wild qat tree habit (see arrow) in southern Kenya. (C) Typical Ethiopian qat farm, intercropped with maize. (D) Qat bundle of mixed varieties for sale in an Ethiopian market. (E) Qat flowers.

Ethiopian farms and 525 cultivated specimens were collected from 217 Kenyan farms (including those in which the farmers reported cultivating wild qat that was either in situ or had been translocated from the wild). Four wild populations (29–37 specimens per population) were sampled from Ethiopia (only three had been previously reported for the entire country), and 20 wild populations (5–30 specimens per population) were sampled from Kenya. Additional specimens from Ethiopia came from a feral population (18 specimens), a farm in which wild plants were cultivated (three specimens), and three markets (26 specimens).

A total of 55 cultivated qat specimens were collected from major qat-production areas in Yemen; we do not know of any evidence for wild qat in Yemen. Additional specimens included six cultivated specimens from northern Madagascar, and wild collections from Tanzania (16 specimens), Malawi (1), Zimbabwe (1), and South Africa (1). The complete set of 1561 specimens (Appendix S1; see Supplemental Data with this article) encompasses the regions in which qat is most intensively cultivated as well as all of the areas mentioned in historical literature as the places of origin for cultivated qat. Specimens with associated voucher herbarium specimens are indicated in Appendix 1. Qat is illegal to cultivate in Tanzania (Carrier, 2007), and therefore Tanzanian cultivars could not be collected. Mapping of individuals and populations was conducted using Google Earth version 5.0 (Google, Mountain View, California, USA).

DNA extraction, polymerase chain reaction (PCR), and microsatellite genotyping—DNA was extracted from herbarium specimens or leaf tissue dried in silica gel. We applied a customized DNA extraction process (Appendix S2) that combined methods from protocols in De la Cruz et al. (1995), Alexander et al. (2007), and Lemke et al. (2011). This protocol reduced tannins and PCR-inhibiting metabolites present in qat. Prior to PCR, DNA extractions were quantified using Quant-iT Pico Green dsDNA assay (Life Technologies, California, USA) and equilibrated to 100–150 ng/μL when initial concentrations allowed.

The 17 microsatellites used for genotyping were selected from Curto et al. (2013) on the basis of consistent amplification, allelic richness, and their ability to be multiplexed. PCR bi/triplexes were designed using Multiplex Manager version 1.2 (Holleley and Geerts, 2009; Appendix S3). Two plexed reactions were combined before fragment sizing. Forward primers were directly labeled with fluorophores while reverse primers were synthesized with PIG-tails to reduce stutter artifacts (Brownstein et al., 1996). PCR was conducted in 5 μL reactions with 2.5 μL Qiagen Multiplex Master Mix (Qiagen, Venlo, The Netherlands), 0.5 μL multiplex primer set, with each primer at 1 μM, 1.45 μL molecular grade H_2O, and 1 μL DNA template. Thermocycler conditions were 15 min at 95°C followed by 40 cycles of 30 s at 94°C, 90 s at 55°C, and 60 s at 72°C, with a final extension of 30 min at 60°C.

PCR products were diluted 20-fold and run with the GeneScan 500LIZ Size Standard (Applied Biosystems, Foster City, California, USA) using an ABI 3130 or 3730XL DNA Analyzer. GeneMapper version 5.0 (Applied Biosystems) was used for scoring. Bin fidelity and mutant alleles were checked using Flexibin version 2.0 (Amos et al., 2007).

The total error rate across all loci was estimated using a single genotype (replicated four times on each 384-well plate, totaling 16 replicates) and dividing the number of incorrect calls for the replicate genotype by the number of total calls across all replicates and loci. The error rate across all genotypes was calculated by including the primers CE3 and CE42 in two different PCR multiplexes, comparing the allele calls across all genotypes, and dividing the total number of different calls by the total number of calls. Any individual with ≥17% missing data was removed from the final data set. The 17% cutoff was determined experimentally from our dataset to be the point at which ambiguous individual assignments could occur. The total percentage of missing data was kept below 5%, based on simulations demonstrating that levels of missing data above 5% can negatively influence genetic-clustering methods (Reeves et al., 2016).

Genetic analyses—Clones were identified using the P_{sex} calculation with 10,000 simulations in MLGsim version 2.0 (Stenberg et al., 2003). The P_{sex} calculation is a test statistic to determine whether similar genotypes are the result of chance sexual reproduction or are truly clonal, based on simulated recombination from the observed allele frequencies. In addition, genotypes that could not be distinguished because of missing data (a maximum of 17%) or that differed by a single allele were considered clones. Including genotypes differing at a single allele accounted for the error rate calculated in our dataset. Clonal groups were then given unique numeric designations as well as the codes "M" if they included genotypes with missing data and "E" if they included genotypes differing by a single allele (Appendix S1). A single representative individual was selected from each clonal group for further analyses. Data quality of the final set of wild genotypes was assessed using both global linkage disequilibrium tests and Hardy-Weinberg equilibrium as implemented in FSTAT version 2.9.3.2 (Goudet, 1995) and GenAlEx version 6.5 (Peakall and Smouse, 2012).

Population subdivision was assessed using STRUCTURE version 2.3.4 (Pritchard et al., 2000; Falush et al., 2003, 2007; Hubisz et al., 2009). STRUCTURE is a model-based approach that employs a Bayesian algorithm to assign genotypes to clusters that maximize Hardy-Weinberg equilibrium and linkage between loci within clusters (K).

A range of K values from 1 to 24 were assessed from the 24 Ethiopian and Kenyan wild collection sites, conducting 40 simulations per K value with 30,000 iterations per simulation (the first 10,000 as burn-in) using an admixture model. The optimal value for K was determined using the ΔK method (Evanno et al., 2005) as implemented in Structure Harvester version 0.6.94 (Earl and vonHoldt, 2012). Optimal values of K from the ΔK method were further assessed through examination of simulation congruence in CLUMPAK version 1.1 (Kopelman et al., 2015). Substructuring was examined by running STRUCTURE again for each cluster identified (Rosenberg et al., 2001). STRUCTURE parameters were identical across all runs with the exception that a smaller range of K values was used to identify substructures.

Cultivated genotypes were assigned to the genetic clusters among the wild populations using GeneClass version 2.0 (Piry et al., 2004; Baudouin and Lebrun, 2001). Using the wild genetic clusters (best-supported partition for wild genotypes) as a reference set, the origins of cultivated genotypes were inferred on the basis of which genetic cluster a given genotype was assigned to at ≥0.95 posterior probability (PP). Assignments of cultivated genotypes were also conducted within each wild genetic cluster using the wild subclusters as reference sets to infer more specific origins for cultivated genotypes. The within-cluster assignments used only the cultivated genotypes assigned to the initial wild cluster. The assignment results were also used to infer translocations. That is, cultivated genotypes collected geographically distant from the wild cluster to which they were assigned were inferred to have been translocated.

Hybrid cultivated genotypes were discovered by examining the STRUCTURE results. Those genotypes that had nearly equal assignment to two wild clusters or subclusters were treated as putative hybrids (Vähä and Primmer, 2006). The putative hybrids were then checked for the presence of diagnostic alleles that could elucidate their parentage. Genotypes geographically adjacent to putative hybrids were examined to check whether they possessed sets of alleles that could produce the genotype of those putative hybrids. Genetic diversity and partitioning results were also assessed and visualized

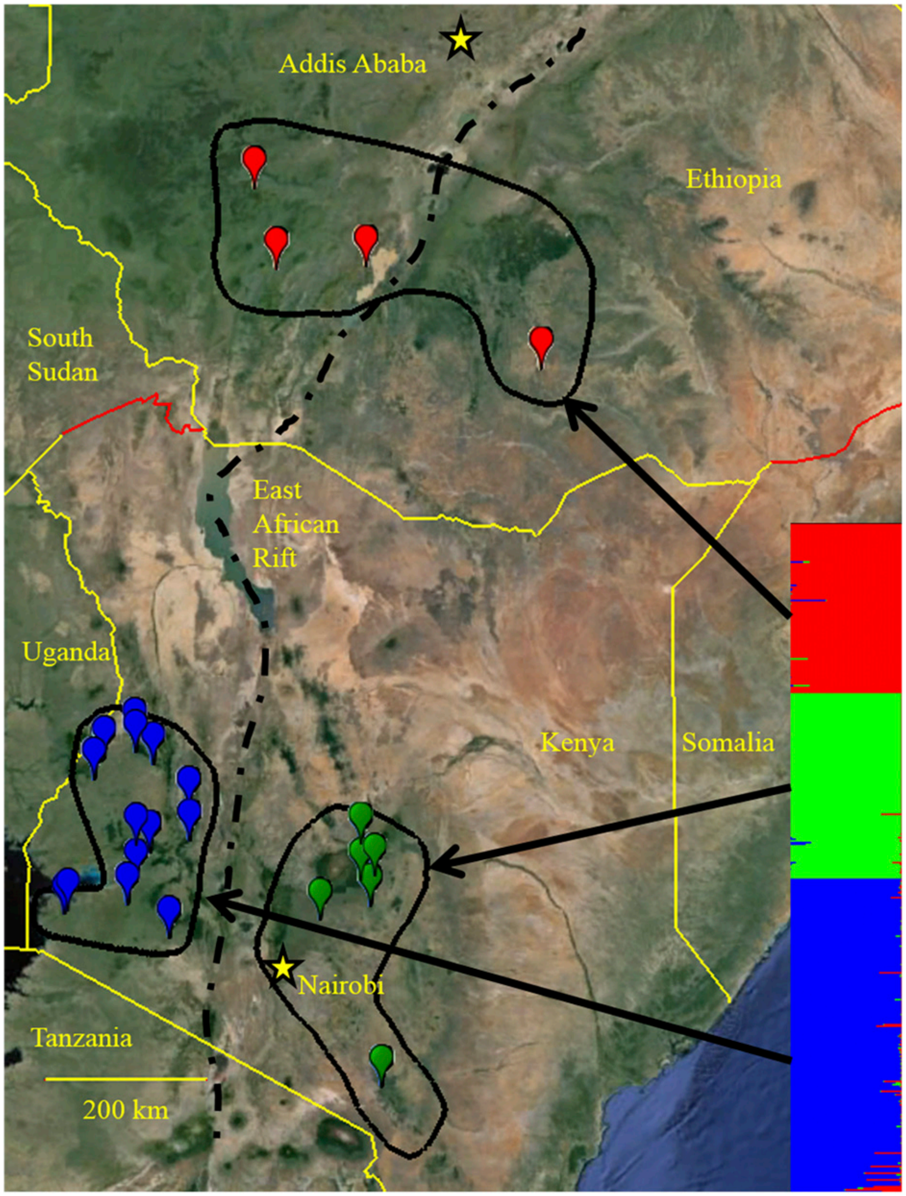

FIGURE 2 Wild qat collection sites in Ethiopia and Kenya with the highest supported genetic partitioning as determined through multiple STRUCTURE runs, as indicated by the bar plot at the lower right. The East African Rift is indicated by the dashed line.

using the nonparametric approach of PCoA (principal coordinate analysis) as implemented in GenAlEx.

RESULTS

Origins of cultivated qat—The final dataset included 1561 individuals, genotyped for ≤17 microsatellite loci, with 2.5% missing data and an estimated error rate of 2.4%. Clonal cultivars were identified in Ethiopia (567 cultivated specimens, 293 unique genotypes) Kenya (525, 305), Madagascar (6, 1), and Yemen (55, 30). Most cultivated genotypes were not found in the wild. The total number of cultivated genotypes found in each country followed by the number of cultivated genotypes that are also present in wild specimens is as follows: Ethiopia (293, 0), Kenya (305, 7), Madagascar (1, 0), and Yemen (30, 0).

The partition with the best likelihood from the set of replicate STRUCTURE runs was $K = 3$ for all wild qat genotypes (Fig. 2; Appendix S4). The three clusters are as follows: (1) genotypes from four wild collection sites in the highlands of southern Ethiopia, (2) genotypes from six wild collection sites around Meru/Mt. Kenya and adjacent to Mt. Kilimanjaro on the eastern side of the EAR (East African Rift), and (3) genotypes from the 12 wild collection sites from the eastern slope of Mt. Elgon in the north across the Endebess Plateau south to the Mau Forest on the western side of the EAR. The Malawian, South African, Tanzanian, and Zimbabwean genotypes were weakly supported as being clustered with the Mt. Elgon cluster. The weak support for these genotypes is based on the low percentage of individual assignment to any genetic cluster in the STRUCTURE analysis. Because of the small number and dispersed geographic collection of these samples, the power to group these genotypes into geographically meaningful clusters was limited.

In addition to the overall pattern of genetic structuring, we explored patterns of hierarchical substructuring within each of the three clusters. Each of the three clusters was found to have a pattern of substructuring generally consistent with geographic features and collection site. The Mt. Elgon cluster west of the EAR was split into two subclusters that are geographically separated by a lower-elevation area across the Endebess Plateau (Fig. 3). The Mt. Kenya/Meru cluster east of the EAR was split into two subclusters, each occurring on higher-elevation uplift features. The southern subcluster consisted of genotypes collected near Mt. Kilimanjaro, and the northern subcluster consisted of genotypes from five collection sites around the southern flanks of Mt. Kenya to the Meru area (Fig. 3). The Ethiopian cluster was split into two subclusters consisting of two collection sites on the west side of the EAR near Lake Abaya and two collection sites that span the EAR to the north (Fig. 3). The EAR does not appear to be an important geographic feature in separating the northern subcluster.

All but three cultivated genotypes could be assigned to one of the three genetic clusters at the 0.95 PP cutoff. Of the three genotypes that were not assigned, two were later inferred to be hybrids (see below). The remaining unassigned genotype was collected from the Nyambene area in Kenya. When the Tanzanian genotypes were included in the reference set, this genotype was assigned to the Tanzanian set with 0.99 PP. The number of unassigned genotypes increased when assignments were conducted at the subcluster level. This reduction in assignment percentage is expected, given that the range of allele frequencies is reduced because of the smaller size of the subpartitioned reference sets (e.g., Cain et al., 2000).

The three wild sources/genetic clusters contributed very different proportions of cultivated genotypes to the total number of

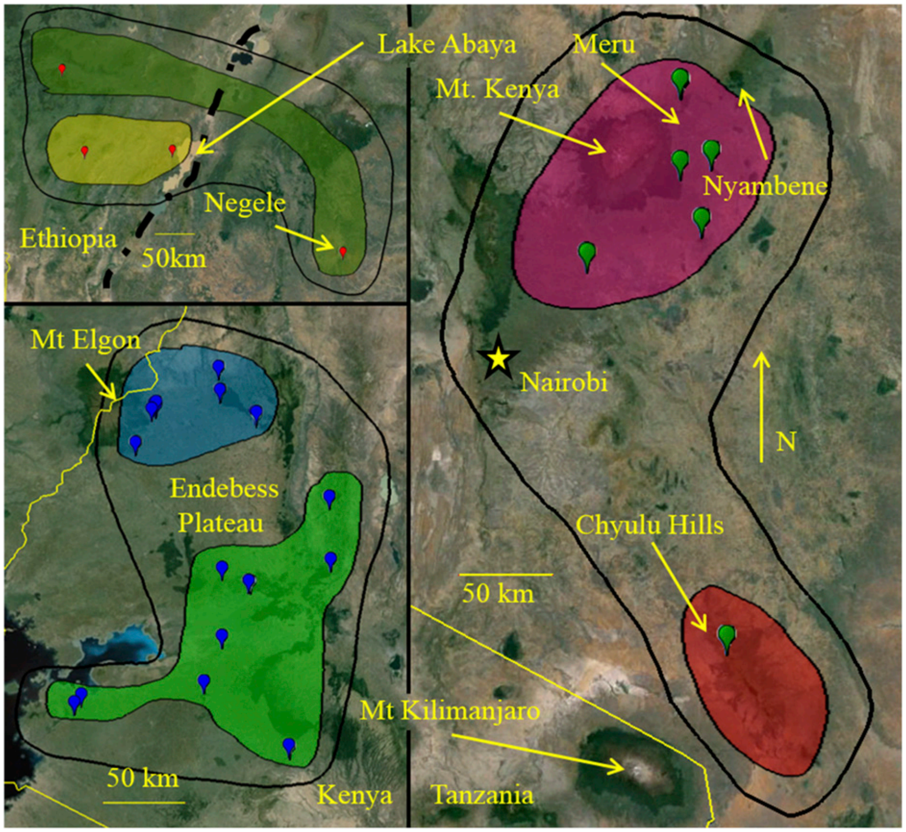

FIGURE 3 Genetic subclusters within each of the three highest-supported wild-qat partitions.

cultivated genotypes analyzed. A total of 101 cultivar genotypes were assigned to the cluster west of the EAR in Kenya, within which 34 were assigned to the northern subcluster and 58 to the southern subcluster, leaving nine cultivar genotypes that were not assigned to either subcluster (Fig. 4). A total of 204 cultivar genotypes was assigned to the cluster east of the EAR in Kenya, within which 185 were assigned to the northern subcluster around Mt. Kenya/Meru, and 18 to a southern subcluster near Mt. Kilimanjaro in the Chyulu Hills area, with one cultivar genotype unassigned to either subcluster (Fig. 4). A total of 293 cultivar genotypes were assigned to an Ethiopian origin; within this cluster, 54 were assigned to the northern subcluster, 198 to the southern subcluster west of Lake Abaya (Fig. 4), and 41 were not assigned to either subcluster.

Dispersal of cultivated qat—Ethiopia is the source of the greatest number of translocations to other sampled regions, most notably Yemen (Fig. 5). Thirty Yemeni cultivar genotypes were assigned to the Ethiopian cluster and all but one were assigned to the southern subcluster (the unassigned genotype also had the highest PP [0.63] of assignment to the southern subcluster). We applied four different methods to test for an independent origin of Yemeni cultivated qat from undiscovered or extinct wild populations in Yemen. First, as noted above, genetic clustering with assignment did not support a separate Yemeni origin. Second, we calculated Goldstein et al.'s (1995) linear genetic distance. The mean genetic distance between the Ethiopian wild source populations and Ethiopian cultivars is 6.1, whereas it is 5.0 between Ethiopian wild source populations and Yemeni cultivars. Thus, in the context of the genetic distances between groups, the Yemeni cultivars are not clearly distinguishable from Ethiopian cultivars. Third, to test the possibility that extant Yemeni cultivated qat had both a Yemeni origin and an Ethiopian origin, PCoA (Appendix S5) was applied to identify any outliers among the Yemeni genotypes. In the PCoA analysis, the Yemeni genotypes clustered with Ethiopian cultivars, providing additional support for an Ethiopian origin of Yemeni cultivated qat. Fourth, no private alleles were identified in any of the Yemeni cultivars. Translocations of cultivated qat from Ethiopia into Kenya were identified. Eleven cultivar genotypes found in Kenya were assigned to the southern Ethiopian subcluster. In Madagascar, all six individuals sampled were from a single clonal genotype that was assigned to the southern Ethiopian subcluster (Fig. 5).

We did not identify any locations that had cultivated genotypes translocated from Kenya to other countries that we sampled based on the assignment tests. But translocations across the EAR and within Kenya were common (Fig. 5; Appendix S6). From the southern subcluster of the Mt. Elgon cluster, two genotypes were found in the Nyambene area northeast of Meru, one genotype near Nairobi, one genotype in the city of Nakuru in the EAR, and 24 near the village of Poro in Samburu County. Two Poro farmers stated that cultivated individuals were translocated from unsampled wild populations nearby.

Hybridization among cultivated qat genotypes—Clear examples of hybridization were found among cultivated genotypes in the isolated highland areas of Marsabit and Kulal in northern Kenya, which are located between the southern edge of the Ethiopian highlands and the northern edge of the Mt. Kenya uplift area. Hybrids from Marsabit and Kulal were inferred on the basis of STRUCTURE results in which genotypes had ~50% membership coefficient to both an Ethiopian and Kenyan origin together with an intermediate placement between Ethiopian and Kenyan genotypes in a PCoA plot (Appendix S8). Furthermore, when considering only the genotypes found in Marsabit, 19 private alleles were found among the genotypes of Kenyan origin and 15 private alleles among the genotypes of an Ethiopian origin. These 34 private alleles are from the loci CE22, CE37, CE41, CE42, and CE43. All hybrids from Marsabit and Kulal were heterozygous at these five loci, possessing one Ethiopian and one Kenyan private allele (Appendix S9).

Cultivated qat in Marsabit was represented by seven Ethiopian genotypes, eight Mt. Kenya/Meru genotypes, and four hybrid genotypes. In Kulal the cultivated qat was represented by three Ethiopian genotypes and three hybrid genotypes (Appendix S8).

Divergence between wild and cultivated qat—To assess the divergence of cultivated populations from wild progenitors, we combined one of each of the cultivated genotypes with those from wild progenitors in an assessment of population structure, both across the entire sampling area and within each of three origins of qat

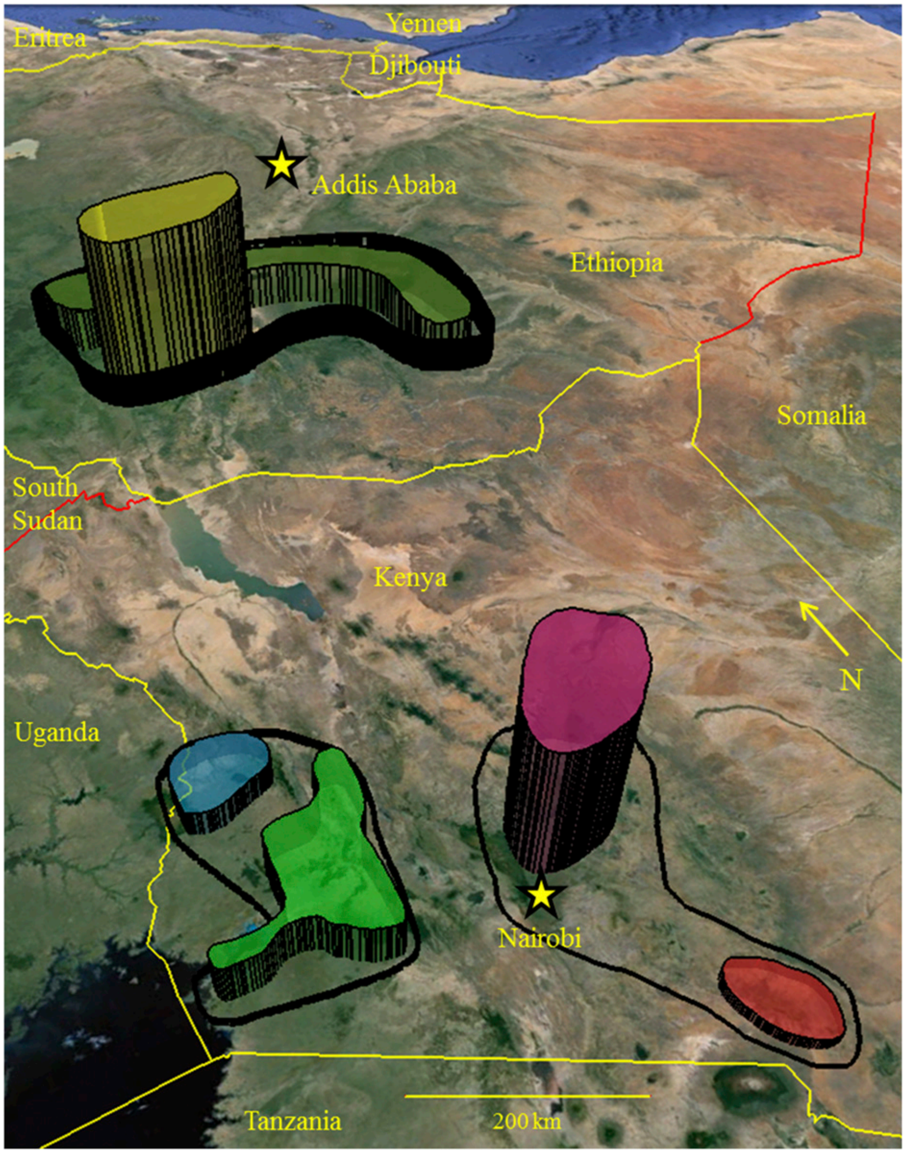

FIGURE 4 The contribution of each wild cluster and subcluster to the number of cultivated genotypes. Height of the subclusters is proportional to the number of cultivated genotypes assigned to each. Height of the outline around clusters is proportional to the number of genotypes assigned to the larger cluster but not assigned to any subcluster by ≥0.95 PP.

cultivation. When analyzing the cultivated and wild genotypes together across the entire sampling area, the most highly supported partitioning of the data from the STRUCTURE analysis is $K = 3$, which is the same number of clusters found for the wild-only analysis of population structure (Fig. 2).

When analyzing only the wild and cultivated genotypes assigned to the Ethiopian center of origin, a pattern of divergence between cultivated and wild populations is clear at the highest supported partitioning of $K = 2$ (Fig. 6A). At $K = 2$, 125 (at a 0.95 cutoff and averaged across all simulations) cultivated genotypes are separated from the remaining 272 wild and cultivated genotypes. From the STRUCTURE results, a weaker pattern of divergence is found with the cultivated and wild genotypes from the two Kenyan centers of origin. In these cases, the cultivated genotypes clustered with their corresponding wild genotypes at the highest supported K value of 2 (Fig. 6B, C). However, the divergence of cultivated from wild genotypes is indicated in the PCoA plot for the Mt. Kenya/Meru cluster (Fig. 6B).

In exploring a range of K values, 74 (at a 0.95 cutoff and averaged across all simulations) of the 205 cultivated genotypes from the Mt. Kenya/Meru origin form a separate cluster in all simulations at $K = 3$, whereas the separate clustering of some cultivated genotypes from wild genotypes with a Mt. Elgon origin did not occur until $K = 4$, and at that point only 23% of the 40 independent simulations indicated separation of some of the cultivated genotypes (Appendix S10). Taken together, these patterns show that the genetic signal of divergence between the wild and cultivated genotypes is strongest in Ethiopia, moderate in Mt. Kenya/Meru, and weakest in Mt. Elgon.

The quantification of genetic divergence using G-statistics (Meirmans and Hedrick, 2011) and similar measures (Chakraborty and Jin, 1993) also indicates a similar trend in divergence with G_{ST} between Ethiopian cultivated and wild material the greatest at 0.07, whereas G_{ST} between the cultivated and wild material was less at 0.04 and 0.02 for the Mt. Kenya/Meru and Mt. Elgon centers, respectively (Table 1). No major loss of heterozygosity was found among all cultivated genotypes, with a mean proportion of heterozygous loci of 0.51, compared to the mean proportion of heterozygous loci at 0.55 for wild genotypes.

DISCUSSION

Multiple origins for cultivated qat—Qat is cultivated in many highland areas of East Africa, especially in Ethiopia and Kenya, where historians have proposed that qat cultivation began. For the first time, using genetic data, we have confirmed that qat cultivation has multiple origins from wild sources in Ethiopia and Kenya. In addition, we have shown that cultivated genotypes have evolved from their wild progenitors through genetic divergence and novel hybridization.

The phylogeographic patterns for wild qat populations are similar to the findings for other studies examining genetic patterns within the Afromontane flora (e.g., Kebede et al., 2007; Kadu et al., 2013; Mairal et al., 2015). For instance, Kebede et al.'s (2007) study of genetic structuring in *Lobelia giberroa* Hemsl. (Campanulaceae) and palynological work (e.g., Vincens et al., 2005) indicate that forest habitat "bridges" that connected high-altitude areas across the EAR in Kenya began to contract with interglacial drying ~5500 calibrated years before present. Similarly, this trend of glacial-period habitat-bridge recession was also inferred to affect plant population structuring between the Ethiopian highlands and the

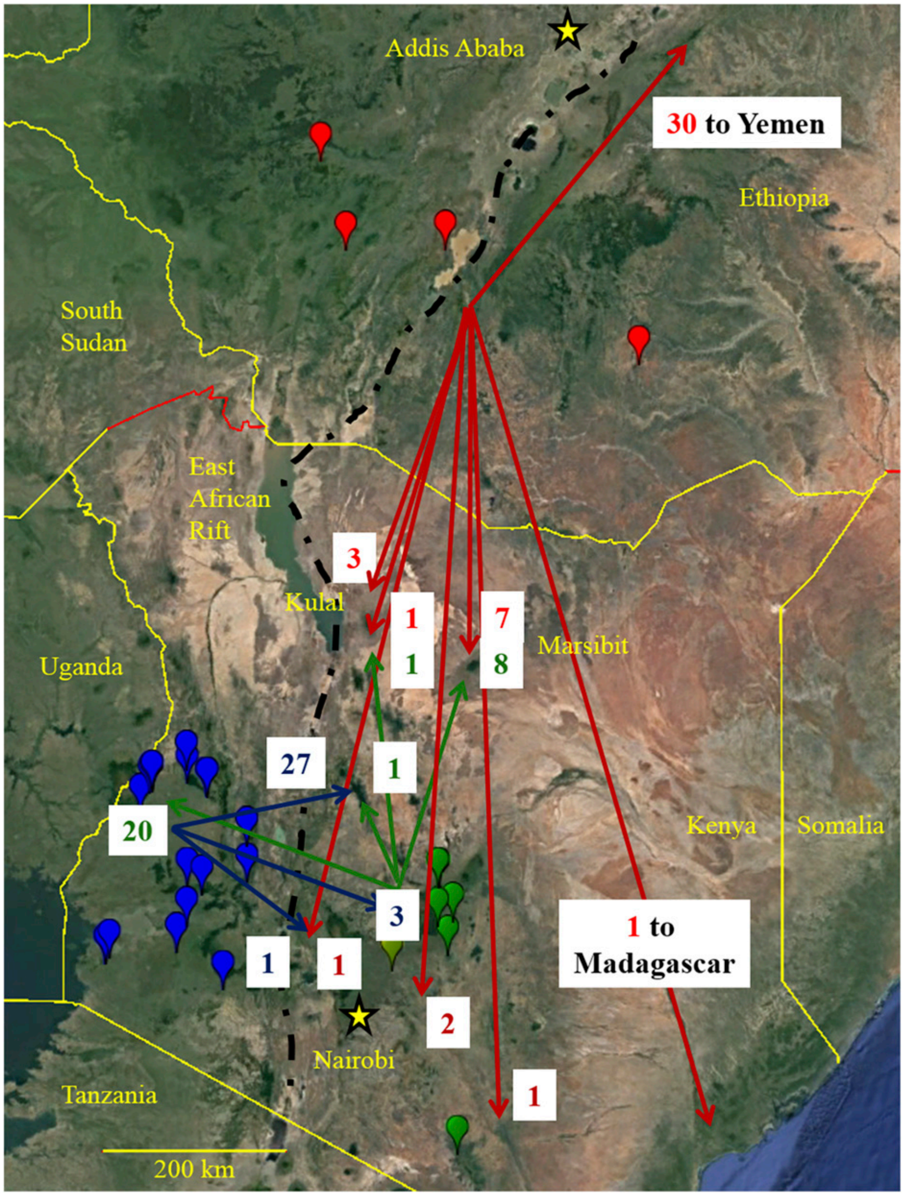

FIGURE 5 Translocation of genotypes from centers of origin, with arrows originating from the origins. Numbers adjacent to arrows indicate the number of translocated genotypes found in a given area. Arrows and numbers are color coded by wild-cluster place of origin.

mountainous areas of south central Kenya (e.g., Kebede et al., 2007). Thus, a similar biogeographic explanation probably applies to the broad-scale pattern in which wild qat populations are not connected by gene flow during warmer, dry periods.

The population structure found among wild qat genotypes indicates that the EAR and other geographic features are important barriers in limiting gene flow. The four Kenyan subclusters delimited by STRUCTURE each consist of genotypes associated with a given uplift feature, which indicates persistent isolation effects between these uplift areas. The exception to this biogeographic patterning among subclusters is two collection sites in Ethiopia that form a subcluster but are separated by the EAR (Fig. 3). This pattern might be attributable to more recent dispersal and gene flow across the EAR in Ethiopia and/or to fewer wild populations being available for collection. Whichever the case, the northern subcluster was not sampled as intensively in the founding of cultivated qat, with 68% of the applicable cultivar genotypes being assigned to the southern subcluster. Furthermore, when using individual wild collection sites as reference sets for assignments, the Negele site in Ethiopia to the east of the EAR was not inferred to be the source population for any Ethiopian cultivated genotypes. On the basis of these analyses, Ethiopian cultivated qat was mainly derived from wild populations to the west of the EAR.

Ethiopia and Mt. Kenya/Meru are the areas consistent with the primary sources for cultivated genotypes. Wild genetic diversity, long-term cultivation, and the system of farming (e.g., degree of trade between farmers and propagation methods) are thought to be key factors in producing genotypically diverse cultivars (Harlan, 1992). This dynamic between wild diversity and cultivation intensity is demonstrated by comparing the Mt. Elgon and Mt. Kenya/Meru centers of origin. Wild genotypic diversity is very high in the Mt. Elgon center of origin, yet twice as many cultivar genotypes (205 vs. 103) originated from the Mt. Kenya/Meru center of origin, which indicates longer and/or more intensive cultivation in that area. Our genetically based finding of multiple origins for qat in Kenya and Ethiopia corroborates earlier anthropologically based assertions (e.g., Huntingford, 1965; Carrier, 2007).

Wild qat genotypes that overlap with cultivated genotypes in PCoA plots indicate a relatively recent initial cultivation or ongoing introgression with wild genotypes. By contrast, "ancient" crops often cannot be clearly matched to extant wild populations because either the wild progenitor populations are extinct or the cultivated and wild populations have diverged to a degree that confident assignments are not possible (e.g., Ladizinsky, 1999; Denham et al., 2003). On the basis of this expected pattern, western Ethiopia and Mt. Kenya/Meru are the oldest areas of cultivation, in that they are the sources of the greatest number of cultivated genotypes and divergence was found between many of the cultivated and wild genotypes (Fig. 6A, B).

Differential patterns of translocation from wild centers of origin—Extant wild populations of qat are known in Ethiopia and Kenya (Palgrave, 1983; Robson et al., 1994) but have not been described from Yemen (Gabali, 1995; Wood, 1997). There are suitable habitats in Yemen for wild qat, but field studies in these areas have not reported wild qat there (Hall et al., 2008). We found no results that indicated an independent Yemeni origin of cultivated qat derived from Yemeni wild populations. Rather, the results from our STRUCTURE analyses, assignment tests, genetic distance, and PCoA analyses support an Ethiopian wild origin of Yemeni cultivated qat. By way of comparison, our methods of genetic-cluster

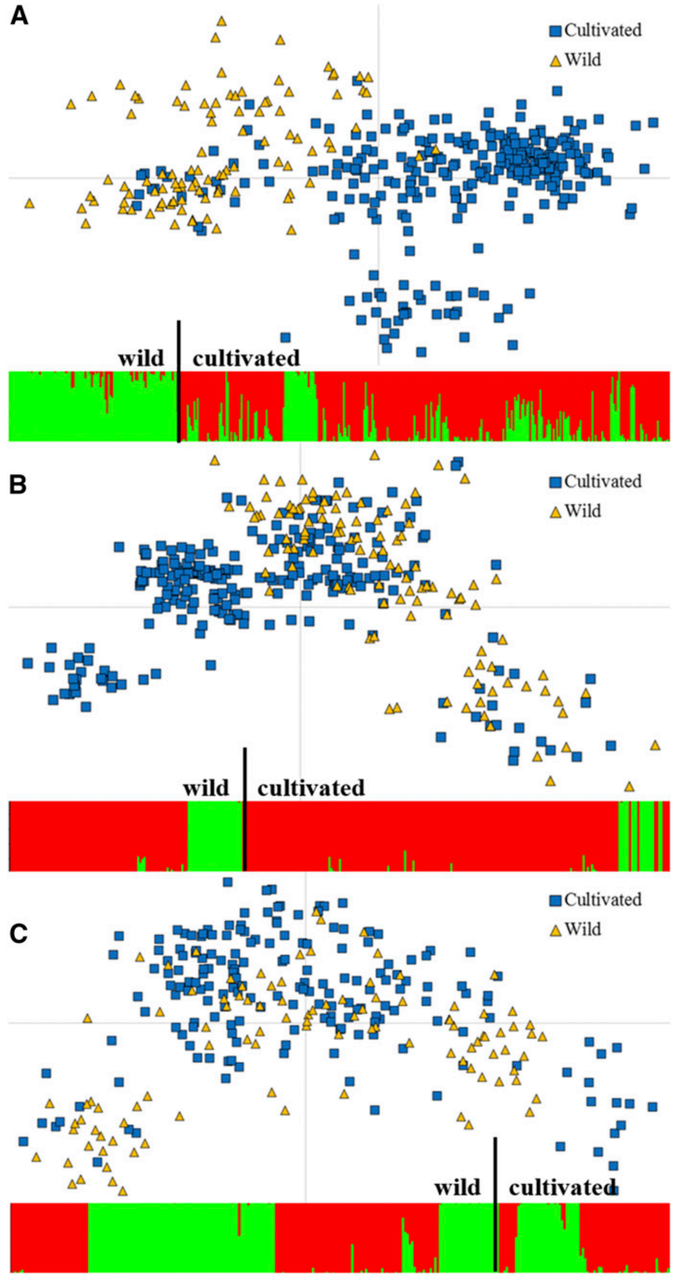

FIGURE 6 PCoA and STRUCTURE bar plots for each of the three centers of origin, including both wild and cultivated genotypes that were assigned to that center of origin. PCoA plots are color coded by cultivated (blue squares) and wild (yellow triangles) status, and STRUCTURE bar plots are color coded by cluster membership (red = cluster 1, green = cluster 2). (A) Ethiopian center of origin, with both methods indicating divergence between wild and cultivated genotypes. (B) Mt. Kenya/Meru center of origin, with PCoA indicating divergence. (C) Mt. Elgon center of origin, with neither method indicating a clear pattern of divergence.

TABLE 1. Divergence as quantified by G_{ST} and related measures between cultivated and wild genotypes in each of the three centers of origin

Centers of origin	G_{ST}	G'_{ST} (Nei)	G'_{ST} (Hedrick)	G''_{ST}	D_{est}
Mt. Elgon	0.018	0.035	0.11	0.126	0.094
Mt. Kenya/Meru	0.039	0.076	0.148	0.181	0.113
Ethiopia	0.07	0.13	0.229	0.279	0.171

identification separated genetic groups on either side of the EAR in Kenya that may have been connected by gene flow as recently as 5000 yr ago (e.g., Kebede et al., 2007). The 6 million yr separation between Africa and the Arabian Peninsula (Fernandes et al., 2006) would have resulted in numerous allelic differences between Ethiopian and any wild Yemeni qat populations, such that individual genotypes from a possible wild Yemeni origin would be distinct from genotypes of African origin. Yet no consistent differences were found among the cultivated Yemeni genotypes (Appendices S5 and S7). Additionally, if Yemeni cultivated material was sourced from native Yemeni populations, then we expect the mean genetic distance to be distinctly higher between the wild Ethiopian and cultivated Yemeni genotypes, which was not the case. However, in Yemen the majority of cultivated genotypes were not collected in Ethiopia or Kenya, and thus Yemen is an important source of cultivated-qat genotypic diversity. A similar situation applies to Ethiopian and Yemeni coffee genotypes (e.g., Silvestrini et al., 2007).

The source of Malagasy qat is thought to be a translocation shortly after 1920 from Yemen (Carrier and Gezon, 2009). The six individuals obtained from Madagascar are all from a single clonal genotype that was the most broadly distributed and commonly encountered clonal group in our dataset. Our findings are consistent with a Yemeni or Ethiopian origin. This pattern of reduced genetic diversity associated with distant translocation from wild centers of origin in qat is very similar to the translocation pattern in coffee. Coffee was native to the highlands of Ethiopia, numerous cultivar genotypes were brought to Yemen, and very few were taken from Yemen to found coffee production in other parts of the world (Simpson and Ogorzaly, 2000).

Distant translocation from either of the Kenyan centers of origin was not found in our dataset, though translocation within Kenya was common. The Meru area is the primary source from which cultivated qat is translocated. Translocations to Mt. Marsabit led to the formation of novel hybrid genotypes from genetically distant parents (Appendix S8). This pattern was not documented in other areas founded by translocated genotypes. Unlike the Ethiopian and Mt. Kenya/Meru centers of origin, few translocated qat genotypes were from the Mt. Elgon center of origin, which indicates recent and/or less intensive cultivation from this region.

Hybridization between translocated genotypes—The pattern of hybridization among qat cultivars (Appendix S8) is similar to the pattern reported for other long-lived perennial crops, such as apple (e.g., Cornille et al., 2015; Volk et al., 2015) and almond (e.g., Delplancke et al., 2012), wherein novel hybridization occurred after genotypes from disparate centers of origin were brought together. However, translocation followed by hybridization of distantly related genotypes appears to be atypical when considering the progressively reduced genotypic and genetic diversity among qat translocated from Ethiopia to Yemen and to Madagascar. Furthermore, in Wote and Wundanyi in southern Kenya, where both Ethiopian and Kenyan genotypes are grown together, hybrid genotypes were not found.

Given that only Ethiopian genotypes and hybrids were found in Kulal, it seems unlikely that the hybridizations occurred in Kulal. Three farmers growing hybrids in Kulal mentioned that these plants were more bitter and potent than other genotypes grown in the area. All six interviewed farmers in Kulal named Marsabit, Meru, and/or neighboring farms as their sources of cuttings to establish their orchards.

Different levels of divergence between cultivated and wild qat based on geographic origin—Separation between cultivated and wild genotypes is not detectable at $K = 2$ when analyzing the dataset encompassing all wild and cultivated collections from Ethiopia and Kenya (Appendix S10). Thus, broad-scale biogeographic processes at this level appear to override the signal of divergence created through artificial selection. However, when examining higher levels of K across the entire dataset or examining divergence within each center of origin, the next partitions to emerge are the cultivated genotypes from Ethiopia (Appendix S10). Thus, the processes of artificial selection and cultivation in the areas of historical qat cultivation may have resulted in similar levels of divergence to the natural processes that created patterns of subclustering within each of the three wild centers of origin.

Two sampling issues are important to take into consideration when explaining the patterns of divergence seen in the different centers of origin. First, and unrelated to artificial selection, is the issue of sampling wild populations while conducting fieldwork. The wild populations from which a cultivated lineage was sourced may not have been sampled during fieldwork, which has been shown to influence individual assignments to populations (e.g., Cain et al., 2000). Second, the wild population(s) from which a cultivar was sourced could be extinct. This latter scenario may explain the pattern found among Ethiopian wild and cultivated samples, given that the vast majority of primary Ethiopian forests have been cleared and/or converted into agricultural uses (Dessie and Kleman, 2007; Dessie and Kinlund, 2008; Tadesse et al., 2014; Guillozet et al., 2015).

If extinction of the wild-source populations is causing the observed patterns of divergence for the Ethiopian center of origin, then farming methods are an important factor in the preservation not only of genetic diversity in general, but also of genetic diversity from wild source populations. The observed trend among long-lived perennials is for relatively high genetic diversity to be retained among the cultivated genotypes in relation to wild source populations (Miller and Gross, 2011). However, the degree to which this relative measure of diversity has been influenced by wild-population extinction vs. on-farm preservation of cultivars has not been broadly assessed in the literature.

In Ethiopia, cultivated genotypes are more divergent from wild genotypes than in the Kenyan centers of origin (Fig. 6; Table 1; Appendix S10). Different levels of divergence may be the result of different times since initial cultivation. Historical documentation indicates that the Ethiopian center of origin is the oldest and Mt. Elgon the youngest (e.g., Huntingford, 1965; Carrier, 2007), which is supported by our data.

Which theory of domestication best describes cultivated qat? The instant domestication theory (Zohary and Spiegelroy, 1975; Harlan, 1992; Zohary, 2004; Kislev et al., 2006) is not supported for qat, given the number of cultivated genotypes within multiple areas of cultivation. However, only three clonal genotypes made up nearly half of all clonal genotypes, and thus it appears that while a large number of genotypes are cultivated, only a small number of genotypes became widespread in cultivation. From our data, hybridization between distantly related genotypes does not appear to be a major factor in the early stages of qat domestication, but is important in later stages of domestication outside centers of origin, as seen in Marsabit and Kulal.

Divergence between cultivated and wild genotypes in the historical areas of cultivation in Ethiopia and Mt. Kenya/Meru indicates that a dynamic process of selection and extinction has taken place among cultivated and wild qat populations and/or indicates potentially unsampled wild populations. Quantifying divergence, assigning cultivars to wild centers of origin, and identifying hybrid cultivars further clarifies the history of qat cultivation and provides a set of research questions for further study. Two such questions are as follows. Do wild populations persist in northern Ethiopia or Nyambene, Kenya, that are more similar to the wild progenitors of cultivated qat in those areas? Has hybridization between distantly related genotypes produced novel and desirable chemical traits? These questions can be addressed through further fieldwork and growth-chamber experiments, respectively.

ACKNOWLEDGEMENTS

The authors thank two anonymous reviewers, A. Angert, J. Cappa, T. Gilligan, D. Sloan, S. Ward, and J. Warren for their constructive comments and valuable suggestions on numerous aspects of the manuscript; R. Bonnart and A. Henk for assisting with technical aspects of the experiments; the individuals involved with collecting material for this study, including W. Abebe, R. Archer, T. Awas, A. Hailu, E. Kelbessa, C. Lukhoba, G. Mwachala, P. Puppo, H. Razamatsoa, J. Solomon, M. Wondafrash; as well as the Ethiopian, Kenyan, Malagasy, and Yemeni farmers who allowed us to collect their crops. Funding was provided by the American Institute for Yemeni Studies (to L.R.T.), the Harold David Harrington Fellowship (to L.R.T.), the John Simon Guggenheim Foundation (to M.P.S.), and the National Geographic Society (to M.P.S.).

AUTHOR CONTRIBUTIONS

L.R.T., M.P.S., C.M.R., P.A.R., A.R., M.A.C., H.M., and S.S. conceived the experiments. L.R.T., P.A.R., A.R., M.A.C., and S.S. conducted the experiments. M.P.S., G.N., S.D., A.W.A.K., and M. A. conducted fieldwork (including farmer interviews) and contributed samples. L.R.T., M.P.S., C.M.R., and P.A.R. analyzed the data. L.R.T., M.P.S., C.M.R., and P.A.R. drafted the manuscript, and all authors edited the manuscript.

DATA ACCESSIBILITY

The DNA sequence data used to develop the SSR markers for this study can be retrieved at GenBank from accessions JX406758–JX406811. A matrix including SSR fragment lengths for each locus, collection numbers, clone designations, and latitude and longitude used in the analyses for this study is in Appendix S1 (see Supplemental Data with this article; https://figshare.com/s/d284c056a61ee7ecc289).

LITERATURE CITED

Achtak, H., M. Ater, A. Oukabli, S. Santoni, F. Kjellberg, and B. Khadari. 2010. Traditional agroecosystems as conservatories and incubators of cultivated plant varietal diversity: The case of fig (*Ficus carica* L.) in Morocco. *BMC Plant Biology* 10: 28.

Alexander, P. J., G. Rajanikanth, C. D. Bacon, and C. D. Bailey. 2007. Recovery of plant DNA using a reciprocating saw and silica-based columns. *Molecular Ecology Notes* 7: 5–9.

Allaby, R. G., D. Q. Fuller, and T. A. Brown. 2008. The genetic expectations of a protracted model for the origins of domesticated crops. *Proceedings of the National Academy of Sciences, USA* 105: 13982–13986.

Al-Motarreb, A., K. Baker, and K. J. Broadly. 2002. Khat: Pharmacological and medical aspects and its social use in Yemen. *Phytotherapy Research* 16: 403–413.

Altabachew, M., B. S. Chandravanshi, M. Redi–Abshiro, N. Torto, S. Chigome, and B. O. Pule. 2013. Evaluation of the effects of various drying techniques on the composition of the psychoactive phenylpropylamino alkaloids of khat (*Catha edulis* Forsk) chewing leaves. *Bulletin of the Chemical Society of Ethiopia* 27: 347–358.

Amos, W., J. I. Hoffman, A. Frodsham, L. Zhang, S. Best, and A. V. S. Hill. 2007. Automated binning of microsatellite alleles: Problems and solutions. *Molecular Ecology Notes* 7: 10–14.

Anderson, D., S. Beckerleg, D. Hailu, and A. Klein. 2007. The khat controversy: Stimulating the debate on drugs. Berg, Oxford, UK.

Baariu, S. N., and G. C. Mulaku. 2015. Mapping khat (miraa) by remote sensing in Meru County, Kenya. *International Journal of Remote Sensing Applications* 5: 54–61.

Balloux, F., L. Lehmann, and T. de Meeüs. 2003. The population genetics of clonal and partially clonal diploids. *Genetics* 164: 1635–1644.

Baudouin, L., and P. Lebrun. 2001. An operational Bayesian approach for the identification of sexually reproduced cross fertilized populations using molecular markers. *Acta Horticulturae* (546): 81–93.

Bhalla, N. 2002. Ethiopia's khat dilemma. BBC News.

Bouby, L., I. Figueiral, A. Bouchette, N. Rovira, S. Ivorra, T. Lacombe, T. Pastor, S. Picq, P. Marinval, and J.-F. Terral. 2013. Bioarchaeological insights into the process of domestication of grapevine (*Vitis vinifera* L.) during roman time in southern France. *PLoS ONE* 8: e63195.

Bredeson, J. V., J. B. Lyons, S. E. Prochnik, G. A. Wu, C. M. Ha, E. Edsinger-Gonzales, J. Grimwood, et al. 2016. Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nature Biotechnology* 34: 562–570.

Brownstein, M. J., J. D. Carpten, and J. R. Smith. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques* 20: 1004–1010.

Cain, M. L., B. G. Milligan, and A. E. Strand. 2000. Long-distance seed dispersal in plant populations. *American Journal of Botany* 87: 1217–1227.

Carrier, N. C. M., Ed. 2007. Kenyan khat: The social like of a stimulant. Brill, Leiden, The Netherlands.

Carrier, N. C. M., and L. Gezon. 2009. Khat in the western Indian Ocean. *Plantes et Sociétés, Études Océan Indien* 43: 271–297.

Chakraborty, R., and L. Jin. 1993. A unified approach to study hypervariable polymorphisms: Statistical considerations of determining relatedness and population distances. *EXS* 67: 153–175.

Chemitei, J. 2015. One year on, Kenyan farmers affected by UK khat ban. On Our Radar: Kenyan Insecurity, 18 June 2015. http://www.onourradar.org/kenya/2015/06/18/one-year-on-kenyan-farmers-affected-by-uk-khat-ban/.

Coart, E. S., M. VanGlabeke, A. S. DeLoose, I. Larsen, and A. Roldán–Ruiz. 2006. Chloroplast diversity in the genus *Malus*: New insights into the relationship between European wild apple (*Malus sylvestris* (L.) Mill.) and the domesticated apple (*Malus domestica* Borkh.). *Molecular Ecology* 15: 2171–2182.

Cornille, A., A. Feurtey, U. Gélin, J. Ropars, K. Misvanderbrugge, P. Gladieux, and T. Giraud. 2015. Anthropogenic and natural drivers of gene flow in a temperate wild fruit tree: A basis for conservation and breeding programs in apples. *Evolutionary Applications* 8: 373–384.

Curto, M. A., L. R. Tembrock, P. Puppo, M. Nogueira, M. P. Simmons, and H. Meimberg. 2013. Evaluation of microsatellites of *Catha edulis* (qat: Celastraceae) identified using pyrosequencing. *Biochemical Systematics and Ecology* 49: 1–9.

De la Cruz, R., R. Whitkus, and L. Mota-Bravo. 1995. Tropical tree DNA isolation and amplification. *Molecular Ecology* 4: 787–790.

Delplancke, M., N. Alvarez, A. Espíndola, H. Joly, E. Brouck, and N. Arrigo. 2012. Gene flow among wild and domesticated almond species: Insights from chloroplast and nuclear markers. *Evolutionary Applications* 5: 317–329.

Denham, T. P., S. G. Haberle, C. Lentfer, R. Fullagar, J. Field, M. Therin, N. Porch, and B. Winsborough. 2003. Origins of agriculture at Kuk swamp in the highlands of New Guinea. *Science* 301: 189–193.

Dessie, G., and P. Kinlund. 2008. Khat expansion and forest decline in Wondo Genet Ethiopia. *Geografiska Annaler. Series B, Human Geography* 90: 187–203.

Dessie, G., and J. Kleman. 2007. Pattern and magnitude of deforestation in the south central rift valley region of Ethiopia. *Mountain Research and Development* 27: 162–168.

Earl, D. A., and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

Elias, M., L. Penet, P. Vindry, D. McKey, O. Panaud, and T. Robert. 2001. Unmanaged sexual reproduction and the dynamics of genetic diversity of a vegetatively propagated crop plant, cassava (*Manihot esculenta* Crantz), in a traditional farming system. *Molecular Ecology* 10: 1895–1907.

El Mahi, T. 1962. A preliminary study on khat (*Catha edulis* Forskal) together with the institutional history of coffee as a beverage in relation to khat, World Health Organization, Regional Office for the Eastern Mediterranean, WHO document EM/RC 11/10.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14: 2611–2620.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.

Falush, D., M. Stephens, and J. K. Pritchard. 2007. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes* 7: 574–578.

Fernandes, C. A., J. E. Rohling, and M. Siddall. 2006. Absence of post-miocene red sea land bridge: Biogeographic implications. *Journal of Biogeography* 33: 961–966.

Gabali, S. A. 1995. Plant life in Yemen: A general survey and preliminary check-list of the flowering plant species. Aden University Press, Aden, Yemen.

Gebissa, E. 2004. Leaf of Allah: Khat and agricultural transformation in Harerge, Ethiopia. James Currey, Oxford, UK.

Glémin, S., E. Bazin, and D. Charlesworth. 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proceedings of the Royal Society B* 273: 3011–3019.

Goldsmith, P. 1988. The production and marketing of miraa in Kenya. *In* R. Cohen [ed.], Satisfying Africa's food needs: Food production and commercialization in African agriculture. Lynne Rienner, Boulder, Colorado, USA.

Goldsmith, P. 1994. Symbiosis and transformation in Kenya's Meru District. Ph.D. dissertation, University of Florida, Gainesville, Florida, USA.

Goldsmith, P. 1999. The political economy of miraa. *East African Alternatives*, March/April, 15–19.

Goldstein, D. B., A. R. Linares, L. L. Cavalli-Sforza, and M. W. Feldman. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139: 463–471.

Goudet, J. 1995. FSTAT (version 1.2): A computer program to calculate *F*-statistics. *Journal of Heredity* 86: 485–486.

Guillozet, K., J. C. Bliss, and T. S. Kelecha. 2015. Degradation in an Afromontane forest in highland Ethiopia, 1969–2010. *Small-scale Forestry* 14: 121–137.

Hall, M., A. W. Al-Khulaidi, A. G. Miller, P. Scholte, and A. H. Al-Qadasi. 2008. Arabia's last forests under threat: Plant biodiversity and conservation in the valley forest of Jabal Bura (Yemen). *Edinburgh Journal of Botany* 65: 113–135.

Harlan, J. R. 1992. Crops and man. American Society of Agronomy, Madison, Wisconsin, USA.

Hillman, G. C., and M. S. Davies. 1990. Domestication rates in wild-type wheats and barley under primitive cultivation. *Biological Journal of the Linnean Society* 39: 39–78.

Holleley, C. E., and R. P. Geerts. 2009. Multiplex Manager 1.0: A cross–platform computer program that plans and optimizes multiplex PCR. *BioTechniques* 46: 511–517.

Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9: 1322–1332.

Huntingford, G. W. B. 1965. The glorious victories of 'Amda Seyon, king of Ethiopia. Clarendon Press, Oxford, UK.

Kadu, C. A. C., H. Konrad, S. Schueler, G. M. Muluvi, O. Eyog–Matig, A. Muchugi, V. L. Williams, et al. 2013. Divergent patterns of nuclear genetic diversity across the range of the Afromontane *Prunus africana* mirrors variable climate of African highlands. *Annals of Botany* 111: 47–60.

Kebede, M., D. Ehrich, P. Taberlet, S. Nemomissa, and C. Brochmann. 2007. Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopia and tropical east African mountains. *Molecular Ecology* 16: 1233–1243.

Kennedy, J. G. 1987. The flower of paradise: The institutional use of the drug qat in North Yemen. D. Reidel, Dordrecht, The Netherlands.

Kislev, M. E., A. Hartmann, and O. Bar–Yosef. 2006. Early domesticated fig in the Jordan Valley. *Science* 312: 1372–1374.

Kopelman, N. M., J. Mayzel, M. Jakobsson, and N. A. Rosenberg. 2015. CLUMPAK: A program for identifying clustering modes and packaging population structure inferences across *K. Molecular Ecology Resources* 15: 1179–1191.

Ladizinsky, G. 1999. On the origin of almond. *Genetic Resources and Crop Evolution* 46: 143–147.

Lemke, L., M. Rex, E. Zyprian, and R. Töpfer. 2011. A simple, inexpensive and environmentally friendly method for high throughput DNA extraction from grapevine (*Vitis* spp.). *Vitis* 50: 7–10.

Mairal, M., L. Pokorny, J. J. Aldasoro, M. Alarcón, and I. Sanmartín. 2015. Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: The case of the Rand Flora genus *Canarina* (Campanulaceae). *Molecular Ecology* 24: 1335–1354.

McKey, D., M. E. Benoit Punjol, and A. Duputie. 2010. The evolutionary ecology of clonally propagated domesticated plants. *New Phytologist* 186: 318–332.

Meirmans, P. G., and P. W. Hedrick. 2011. Assessing population structure: FST and related measures. *Molecular Ecology Resources* 11: 5–18.

Miller, A. J., and B. L. Gross. 2011. From forest to field: Perennial fruit crop domestication. *American Journal of Botany* 98: 1389–1414.

Miller, A. J., and B. A. Schaal. 2006. Domestication and the distribution of genetic variation in wild and cultivated populations of the Mesoamerican fruit tree *Spondias purpurea* L. (Anacardiaceae). *Molecular Ecology* 15: 1467–1480.

Montero-Rojas, M., A. M. Correa, and D. Siritunga. 2011. Molecular differentiation and diversity of cassava (*Manihot esculenta*) taken from 162 locations across Puerto Rico and assessed with microsatellite markers. *AoB Plants* 2011: plr010.

Palgrave, K. C. 1983. *Trees of Southern Africa*, 2nd ed. Struik, Cape Town, South Africa.

Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel, population genetic software for teaching and research—an update. *Bioinformatics* 28: 2537–2539.

Perrier, X., E. De Langhe, M. Donohue, C. Lentfer, L. Vrydaghs, F. Bakry, F. Carreel, et al. 2011. Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proceedings of the National Academy of Sciences, USA* 108: 11311–11318.

Piry, S., A. Alapetite, J.-M. Cornuet, D. Paetkau, L. Baudouin, and A. Estoup. 2004. GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* 95: 536–539.

Pritchard, J. C., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Reeves, P. A., C. L. Bowker, C. E. Fettig, L. R. Tembrock, and C. M. Richards. 2016. Effect of error and missing data on population structure inference using microsatellite data. *bioRxiv* 10.1101/080630.

Revri, R. 1983. *Catha edulis* Forsk.: Geographical dispersal, botanical, ecological and agronomical aspects with special reference to Yemen Arab Republic. University of Göttingen, Göttingen, Germany.

Robson, N. K. B., N. Hallé, N. B. Mathew, and R. Blakelock. 1994. Celastraceae. *In* R. M. Polhill [ed.], Flora of tropical East Africa, vol. 108. A.A. Balkema, Rotterdam, The Netherlands.

Rosenberg, N. A., T. Burke, K. Elo, M. W. Feldman, P. J. Freidlin, M. A. M. Groenen, J. Hillel, et al. 2001. Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics* 159: 699–713.

Schorno, X., and E. Steinegger. 1979. CNS-active phenylpropylamines of *Catha edulis* Forsk. (Celastraceae) of Kenyan origin. *Experientia* 35: 572–574.

Silvestrini, M., M. G. Junqueira, A. C. Favarin, O. Guerreiro-Filho, M. P. Maluf, M. B. Silvarolla, and C. A. Colombo. 2007. Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. *Genetic Resources and Crop Evolution* 54: 1367–1379.

Simpson, B., and M. Ogorzaly. 2000. Economic botany: Plants in our world. McGraw Hill, New York, New York, USA.

Stenberg, P., M. Lundmark, and A. Saura. 2003. MLGsim: A program for detecting clones using a simulation approach. *Molecular Ecology Notes* 3: 329–331.

Szendrei, K. 1980. The chemistry of khat. *Bulletin on Narcotics* 32: 5–35.

Tadesse, G., E. Zavaleta, C. Shennan, and M. FitzSimmons. 2014. Policy and demography factors shape deforestation patterns and socio-ecological processes in southwest Ethiopian coffee agroecosystems. *Applied Geography (Sevenoaks, England)* 54: 149–159.

United Nations Narcotics Laboratory. 1975. Studies on the chemical composition of khat: III, Investigations on the phenylalkylamine fraction. United Nations Narcotics Laboratory, Geneva, Switzerland.

Vähä, J.-P., and C. R. Primmer. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* 15: 63–72.

Varisco, D. M. 2007. Turning over a new leaf: The impact of qat (*Catha edulis*) in Yemeni horticulture. *In* M. Conan and W. J. Kress [eds.], Botanical progress, horticultural innovation and cultural change. Dumbarton Oaks Research Library and Collection, Washington, D.C., USA.

Varisco, D. M. 2012. Qat and traditional healing in Yemen. *In* I. Hehmeyer and H. Schönig [eds.], Herbal medicine in Yemen: Traditional knowledge and practice, and their value for today's world. Brill, Leiden, The Netherlands.

Vincens, A., G. Buchet, D. Williamson, and M. Taieb. 2005. A 23,000 yr pollen record from Lake Rukwa (8°S, SW Tanzania): New data on vegetation dynamics and climate in Central Eastern Africa. *Review of Palaeobotany and Palynology* 137: 147–162.

Volk, G. M., A. D. Henk, A. Baldo, G. Fazio, C. T. Chao, and C. M. Richards. 2015. Chloroplast heterogeneity and historical admixture of the genus *Malus. American Journal of Botany* 102: 1198–1208.

Wolfes, O. 1930. Über das Vorkommen von *d*-nor-iso-Ephedrin in *Catha edulis. Archiv der Pharmazie* 268: 81–83.

Wood, J. R. I. 1997. A handbook of the Yemen flora. Royal Botanic Gardens, Kew, UK.

World Bank. 2007. Yemen: Towards qat demand reduction. World Bank document report 39738-YE. http://documents.worldbank.org/curated/en/500871468183253500/pdf/397380YE.pdf.

Zohary, D. 2004. Unconscious selection and the evolution of domesticated plants. *Economic Botany* 58: 5–10.

Zohary, D., and P. Spiegel-Roy. 1975. Beginnings of fruit growing in the Old World. *Science* 187: 319–327.