Identification of *Erythroxylum* taxa by AFLP DNA analysis

Emanuel L. Johnson*, James A. Saunders, Sue Mischke, Charles S. Helling,
Stephen D. Emche

*Alternate Crops and Systems Laboratory, Plant Sciences Institute, Beltsville Agricultural Research Center, 10300 Baltimore Avenue,
Beltsville, MD 20705-2350, USA*

Dedicated to the memory of Professor Jeffrey B. Harborne

Abstract

Erythroxylum coca, indigenous to the Andean region of South America, is grown historically as a source of homeopathic medicine. However, in the last century, cultivation of *E. coca* and several closely-related species for the production of illicit cocaine has become a major global problem. Two subspecies, *E. coca* var. *coca* and *E. coca* var. *ipadu*, are almost indistinguishable phenotypically; a related cocaine-bearing species also has two subspecies (*E. novogranatense* var. *novogranatense* and *E. novogranatense* var. *truxillense*) that are phenotypically similar, but morphologically distinguishable. The purpose of this research was to discover unique AFLP DNA patterns ("genetic fingerprinting") that characterize the four taxa and then, if successful, to evaluate this approach for positive identification of the various species of coca. Of seven different AFLP primer pairs tested, a combination of five proved optimal in differentiating the four taxa as well as a non-cocaine-bearing species, *E. aerolatum*. This method of DNA fragment separation was selective, and faster, for coca identification, compared with analyses based on flavonoid chemotaxonomy. Using the 5-primer AFLP approach, 132 known and unknown coca leaf accessions were evaluated. Of these, 38 were collected in 1997–2001 from illicit coca fields in Colombia, and all were genetically differentiated from coca originating in Peru and Bolivia. Based on the DNA profiling, we believe that the Colombian coca now represents a hybridization of *E. coca* var. *ipadu*. Geographical profiling within Colombia also seems feasible as new coca production areas are developed or new types of coca are introduced within traditional growing areas.

1. Introduction

Only when coca leaf became an important pharmaceutical product in the late 19th century did its origin and varietal relationships become the object of scientific inquiry. However, it was not until the 1970's that the botanical origins, archeological records, and chemistry of coca were studied intensively (Plowman, 1984). Indeed, many investigators were unaware of the existence of distinct coca varieties even though geographical, ecological and morphological differences in varieties were detailed as early as the 16th century (Plowman, 1984). By the 1970's, cultivated coca was acknowledged to be the derivative of two discrete species of the genus *Erythroxylum*; *E. coca* var. *coca* Lam. and *E. novogranatense* var. *novogranatense* (Morris) Hieron. According to Plowman and Rivier (1983), this classification was supported by multidisciplinary research investigations on leaf and stem anatomy (Rury, 1981); leaf hybridization and flavonoid experiments (Bohm et al., 1982); breeding systems (Ganders, 1979); alkaloids (Holmstedt et al., 1977, 1979; Rivier, 1981); and traditional studies of the morphology, ecology, and geographical distribution of coca (Plowman, 1979, 1980, 1981, 1982). Furthermore, this body of work led to recognition that each cocaine-producing cultivated species of coca includes two varieties. One, *Erythroxylum coca* var. *coca* [(*E. c.* var. *coca*); 'Bolivian' or 'Huánuco

* Corresponding author. Tel.: +1-301-504-5323; fax: +1-301-504-5823.

E-mail address: johnsone@ba.ars.usda.gov (E.L. Johnson).

coca'] has been considered historically as the source of most of the world's supply of cocaine. It is found on the eastern slopes of the Andes in areas of humid montane forests. *E. coca* var. *ipadu* Plowman [(*E. c.* var. *ipadu*); 'Amazonian coca'] was considered to be cultivated only sparingly in the western part of the Amazon Basin. *E. novogranatense* var. *novogranatense* (Morris) Hieron [(*E. novo.* var. *novogranatense*); 'Colombian coca'] existed as a plantation crop in Colombia and as an ornamental throughout the tropics. Finally, *E. novogranatense* var. *truxillense* (Rusby) Plowman [(*E. novo.* var. *truxillense*); 'Trujillo coca'], mostly confined to the desert areas of northern Peru, was cultivated for 'coca-chewing' and in the 19th and early 20th centuries, for flavoring of an internationally popular, cola-based soft drink (Plowman and Rivier, 1983).

The Agricultural Research Service, US Department of Agriculture (ARS-USDA), at Beltsville, Maryland, has conducted basic and applied research on coca and other illicit drug crops since the 1970s. The extensive USDA living collection of Erythroxylaceae specimens includes all four cultivated taxa of *Erythroxylum* (*E. c.* var. *coca*, *E. c.* var. *ipadu*, *E. novo.* var. *novogranatense*, *E. novo.* var. *truxillense*), all of which were authenticated twice (Dr. Timothy Plowman, personal commun., 1988; Dr. Philip M. Rury, personal commun., 1993). Although the levels of cocaine and other tropane alkaloids differ among these taxa (Holmstedt et al., 1977; Rivier, 1981; Plowman and Rivier, 1983; Johnson and Emche, 1994; Johnson, 1996; Johnson and Foy, 1996), visual inspection of leaf morphology for at least two of these lines, *E. c.* var. *coca* and *E. c.* var. *ipadu*, is not very useful since they are usually phenotypically and morphologically indistinguishable. Similarly, leaf morphology for taxa *E. novo.* var. *novogranatense*, and *E. novo.* var. *truxillense* is phenotypically similar, but somewhat more distinguishable. Positive identification of the taxa for illicit coca is desirable, since potential yield (coca leaf, cocaine) may be affected by the type of coca being grown; furthermore, propagation or other aspects of plant production are influenced by the type of coca planted. Finally, accurate identification of the coca taxa may provide information that is useful in tracking the sources of coca germplasm, as new production areas are developed or old ones modified in response to coca eradication pressures.

Besides classical botanical approaches to coca identification, some recent progress has been made on using chemotaxonomy—the profiling of a plant through assessment of its chemical constituents—for identifying species/varieties of coca. Coca flavonoid composition, for example, has been associated with specific types of cultivated coca (Johnson et al., 1997, 2002; Johnson and Schmidt, 1999). Coca varieties may also be distinguished through the analyses of some 40–100 minor alkaloids in leaf tissue; most useful indicators among these seem to be truxillense and trimethoxycocaine (Casale and Moore, 1994, 1996; Moore et al., 1994). More recently, the isotope ratio (^{13}C:^{12}C and ^{15}N:^{14}N) combined with trace alkaloid composition in coca leaves or extracted cocaine has been associated with geographic origins for coca production in several major growing regions (Ehleringer et al., 2000). The combination of plant morphology and flavonoid analysis was used to confirm that illicit coca discovered in Panama in 1993 was *E. novo.* var. *novogranatense*, almost certainly introduced from nearby Colombia (Helling, 1993). Chemotaxonomy, through the use of HPLC, UV, GC/ LC/MS and NMR, however exact, is both time- and cost-consuming, and not practical for large population analysis. Recent DNA fingerprinting methodologies such as AFLP, RFLP, RAPDs and SNPs, represent improvements over chemotaxonomic analysis in terms of time and cost, and have enabled researchers to analyze genetic diversity within and/or among plant species and populations in a relatively short period of time (Welsh and McClelland, 1990; Williams et al., 1990; Amane et al., 1999; Bhattramakki and Rafalski, 2001; Saunders et al., 2001b; Li et al., 2002; Soleimani et al., 2002). This "genetic fingerprinting" tool was then applied to a bio-geographic survey of coca growing illegally in Colombia. Colombia was selected because of its now-dominant role as the source for production of both coca leaf and cocaine.

In the current study, AFLP was used to analyze 132 *Erythroxylum* samples in order to characterize and positively identify these varieties of coca and in addition, to develop AFLP DNA primers that would differentiate between the cultivated taxa of *Erythroxylum* that produce cocaine.

2. Results and discussion

2.1. Leaf profile

E. coca var. *coca*, *E. c.* var. *ipadu*, *E. novo.* var. *novogranatense*, and *E. novo.* var. *truxillense* are all present in the Andean regions of South America as prominent perennial shrubs, with waxy leaves and a conspicuous red berry (i.e., a drupe). Fig. 1 shows both the abaxial (left) and adaxial (right) surfaces of leaves harvested from the four cultivated taxa of *Erythroxylum*, as well as a new taxon (**c**) under widespread production within Colombia [**a**: *E. c.* var. *coca*; **b**: *E. c.* var. *ipadu*; **c**: *E.* sp. (hybrid); **d**: *E. novo.* var. *novogranatense*; and **e**: *E. novo.* var. *truxillense*]. The difficulty in attempting to identify the taxa by leaf shape and/or size is apparent from this Figure. Note the similarity between the leaf pairs '*a*,' '*b*,' and '*c*'. Leaf **c** represents the current line of coca under cultivation in at least part of Colombia; these specimens were collected from an illicit farm in 2001. The ovate
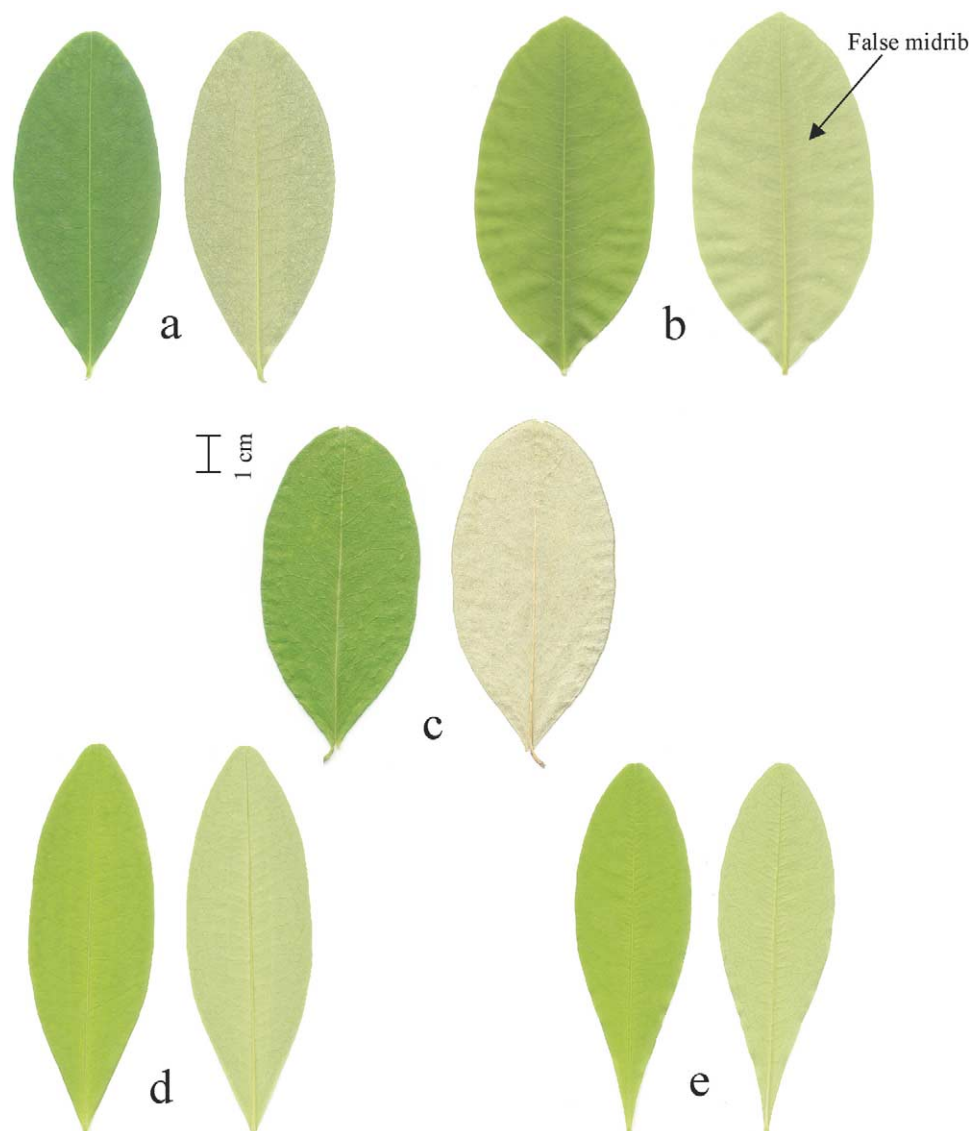
Fig. 1. Leaf morphology of *Erythroxylum*: (a) *E. coca* var. *coca*, (b) *E. coca* var. *ipadu* (Colombia Ca. 1972), (c) *E.* sp. (Colombia, 1997–2001), (d) *E. novogranatense* var. *novogranatense*, (e) *E. novogranatense* var. *truxillense*. Each leaf is represented abaxially (left) and adaxially (right). Leaf size and false midrib are as shown.

leaves have a similar general appearance. Minor differences shared by leaves from the '*a*', and '*b*' lines include a false midrib seen on either side of the true midrib which is not present on sample '*c*' (see arrow, Fig. 1). Leaf pairs '*d*' and '*e*' appear similar, both bearing barely discernible false midribs on their adaxial surfaces near the true midrib, but the proximal area adjoining the petiole in '*d*' differs from that of '*e*' (Fig. 1). Rury (1981) demarcated these differences and strongly suggested that leaf anatomy was not a good taxonomic marker and should not be used to identify these taxa. Further reinforcing and complicating the unreliability of coca leaf morphology alone as an identification tool, is the fact that leaf shape can vary even on individual plants.

2.2. Chromosome number and alkaloid content in leaf tissue

The four cultivated taxa of *Erythroxylum* are very similar morphologically and form a closely related group that differs from other species, both chemically and morphologically (Plowman, 1979, 1980, 1981; Youssefi et al., 1979). They have the same chromosome number ($n = 12$) and are completely allopatric (Plowman, 1979). Although each of the four cultivated *Erythroxylum* taxa produce cocaine, the alkaloid content in leaf tissue varies with leaf age and environmental conditions (Reens, 1919; Plowman and Rivier, 1983; Holmstedt et al., 1977; Johnson and Emche, 1994; Johnson, 1996; Johnson and Foy, 1996) and cannot be used as a specific marker for the taxa (Youssefi et al.,

1979, Rivier, 1981; Schultes, 1981; Turner et al., 1981). The average content for cocaine and cinnamoylcocaines in leaf tissue of the four cultivated taxa of *Erythroxylum* is shown in Table 1. Even though the two *E. novogranatense* taxa have the highest cocaine content, they are not preferred for illicit cocaine production. This is because their leaf tissue contains a higher content of the cinnamoylcocaines (e.g. *cis* and *trans*) as compared to that present in *E. c.* var. *coca* or *E. c.* var. *ipadu* (Table 1; also see Plowman and Rivier, 1983; Johnson and Foy, 1996). According to Plowman and Rivier (1983), the cinnamoylcocaines were referred to by early investigators as 'un-crystallizable cocaine', which reflected the turn-of-the-19th-century chemists' difficulty in extracting and purifying pharmaceutical cocaine from Java coca (*E. n.* var. *novogranatense*). Apparently, this problem in extracting cocaine from *E. novogranatense* is well known by the illicit cocaine producers.

2.3. Leaf flavonoid chemistry as related to AFLP DNA analysis

Bohm et al. (1982) first described the flavonoid chemistry of the four cultivated taxa of coca. In an earlier report (Johnson et al., 1997), we attempted to correlate unique flavonoids as chemotaxonomic markers for distinguishing coca varieties from samples acquired from Bolivia and Peru. In addition to HPLC separation with UV spectroscopic detection, LC/MS analysis was performed for determination of the flavonoids. These findings were confirmed by ^{1}H NMR spectroscopy to define each flavonoid for unambiguous identification of *Erythroxylum* varieties. Subsequently, the flavonoid profiles for the four cultivated varieties, as well as a feral taxon, were determined using key flavonoid(s) as markers for the each variety of coca (Johnson et al., 1998, 2002, 2003; Johnson and Schmidt, 1999).

Table 1 shows a composite chemical analysis of the key flavonoids present in leaf tissue of the four main cultivated varieties of *Erythroxylum*. These flavonoids are also present in leaves of the four cultivated varieties

of coca in the current study (Fig. 2). The ancestral taxon of the cultivated cocas is, according to Plowman (1979, 1980, 1984), *E. c.* var. *coca,* which he suggested contained a single subspecies *E. c.* var. *ipadu*. We found the major flavonoid for *E. c.* var. *coca* to be the flavanone eriodictyol (Johnson et al., 1997). Contrary to Plowman's (1979) suggestion, it seems apparent now that *E. coca* has two types of *E. c.* var. *ipadu* (Table 1). One type was observed earlier in the Amazon region of Colombia (Colombia 1972 in Table 1; see Johnson et al., 1998), and a second type that is currently under cultivation in much of Colombia, including parts of the Amazon basin (Colombia 1997–2001 in Table 1; see Johnson et al., 2002, 2003). Both Plowman (1979) and Rury (1981) described the *E. c.* var. *ipadu* as having two forms of leaf anatomy: one containing a false midrib adjacent to the true midrib, similar to that shown in Fig. 1b, and the other completely lacking a false midrib, similar to that found in Fig. 1c. We have observed morphological and chemical differences in leaf tissue samples identified as *E. c.* var. *ipadu* collected from Colombia in 1997 when compared with *E. c.* var. *ipadu* obtained from Colombian fields in 1972 (Johnson et al., 1998). These accessions, *E. c.* var. *ipadu* Colombia, 1972 and *E. c.* var. *ipadu* Colombia, 1997–2001 (Table 1), also differ in leaf flavonoid composition (Johnson et al., 2002, 2003). The sample *E. c.* var. *ipadu* 'Colombia, 1997–2001′ has flavonoids that are different from *E. c.* var. *coca* [from which it is said to be derived (Plowman, 1979, 1982, 1984)] and also different from those present in the *E. c.* var. *ipadu* obtained in 1972 (Table 1). In addition, the new Colombian ipadu also has a rare flavonol, kaempferol 4′-*O*-(rhamnosyl)glucoside, which was found in neither *E. c.* var. *coca* nor *E. c.* var. *ipadu* currently in our living collection (Johnson et al., 2003). Thus, based on the flavonoid chemistry profiles, *E. c.* var. *ipadu* currently under cultivation in Colombian fields appears to be a hybrid that resulted from a cross between *E. c.* var. *coca*×*E. n.* var. *truxillense*, which apparently occurred after 1972 (Johnson et al., 2002).

Table 1
Phytochemicals present in leaf extracts of four cultivated *Erythroxylum* taxa

Coca type	Selected phytochemical constitutents[a]							
	Eriodictyol	Taxifolin	Quercetin	Luteolin	Fisetin	Kaempferol	Cocaine[b] (%)	Cinnamoyl-cocaines, total (%)
E. c. var. *coca*	+	−	−	-	−	−	0.58	0.18
E. c. var. *ipadu* (Colombia, 1972)	−	+	+	−	−	−	0.41	0.006
E. c. var. *ipadu* (Colombia, 1997-2001)	+	+	+	−	−	+	0.41	0.007
E. n. var. *novogranatense*	−	−	−	+	−	+	0.77	0.46
E. n. var. *truxillense*	−	−	+	−	+	+	0.71	0.38

Presence of the aglycone in samples of leaf tissue is indicated by a positive sign (+) and absence is indicated by a negative sign (−). There was no disparity among the phytochemical profiles analyzed in replicate samples for each taxa.
[a] For a complete listing of the flavonoid glycosides and their structural configurations in leaf tissue of the four cultivated taxon of *Erythroxylum*, see Johnson et al. (1997, 1998, 2003).
[b] Average content (%) in leaf tissue, dry wt basis (Johnson, 1996; Johnson and Emche, 1994; Johnson and Foy, 1996; Rivier, 1981).
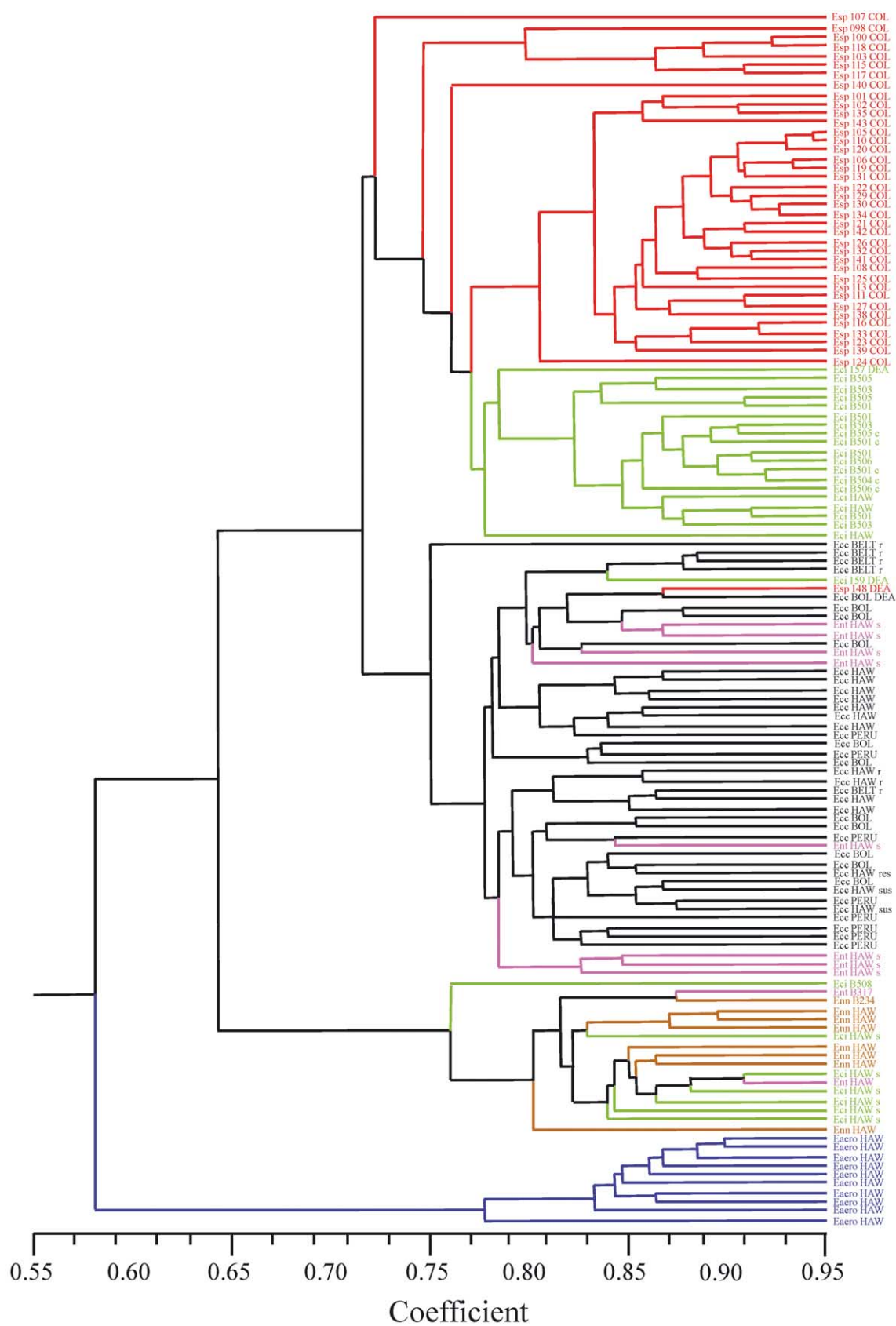
Fig. 2. Dendrogram of 132 *Erythroxylum* samples.

The concept that the predominant coca in production today in Colombia is a hybrid is further supported by AFLP DNA analysis, as shown in the dendrogram in Fig. 2. In the dendrogram, the accessions of *E. c.* var. *ipadu* more recently collected from Colombian fields (Fig. 2; red) and our 1972 (living collection; Fig. 2; green) appear as disjuncts. The dendrogram shows that these samples of *E. c.* var. *ipadu* (represented as Esp COL accessions 098-142) currently under cultivation in fields in Colombia are all contained within the top cluster. The accessions listed in the dendrogram as Eci B (Fig. 2; green), Eci HAW (Fig. 2; not suffixed with 's'; green), as well as Eci DEA (a collection from Colombia supplied by the US Drug Enforcement Administration; Fig. 2; green) form a distinct subcluster within the Esp COL group (Fig. 2; red). These results exhibit a close, but genetically distinct relationship between the Esp COL (Fig. 2; red) and Eci B, Eci HAW, Eci DEA (Fig. 2; all green) clusters. The only exceptions are two outliers, samples Eci B508 and Eci 159 DEA, possibly influenced by edaphic and/or climatic conditions. *E. c.* var. *coca* harvested from fields in Bolivia (Ecc BOL; Fig. 2; black), Peru (Ecc PERU; Fig. 2; black), or the USDA's research site in Hawaii (Ecc HAW; Fig. 2; black) all show a close relationship in forming individual clusters (Fig. 2), a relationship corroborated by the leaf flavonoid data (Table 1).

E. n. var. *novogranatense* and *E. n.* var. *truxillense* are also considered to be derivatives of the ancestral taxon *E. c.* var. *coca* through selective breeding (Bohm et al., 1982; Plowman, 1984). In terms of their leaf flavonoid chemistry, both share the same kaempferol flavonol (Table 1). The dendrogram also supports a close link between taxa which are denoted as Enn HAW (Fig. 2; brown) and Ent B (Fig. 2; pink) or HAW (not suffixed with 's'; Fig. 2; pink). It further shows that they are distantly related to *E. c.* var. *coca* (Fig. 2).

Samples suffixed with an 's' designation were seedlings harvested from the noted species. The seed from both Eci HAW (*E. c.* var. *ipadu*; Fig. 2; green) and Ent HAW (*E. n.* var. *truxillense;* Fig. 2; pink) were from field collections in Hawaii and analyzed for possible hybridization. Because *E. c.* var. *ipadu* is considered a true cultigen (Plowman, 1979), it was particularly interesting to see if the seed was viable and how the crossing had occurred. Our dendrogram (Fig. 2) clearly indicates that *E. c.* var. *ipadu* will readily cross with *E. n.* var. *novogranatense*, while *E. n.* var. *truxillense* hybridizes *E. c.* var. *coca*.

The use of leaf flavonoids as chemotaxonomic markers for the *Erythroxylum* taxa has been successful due to the unique leaf chemistry of each variety of cultivated coca. Unfortunately, this technique requires extended time as well as sophisticated instrumentation for the analyses. In addition, field-grown coca samples can exhibit significantly higher concentrations of flavonoids as compared to greenhouse-grown coca (Johnson, per-

sonal commun.). The use of AFLP DNA analyses reduces the time required for flavonoid analyses by 10-fold and the data acquired were consistent with those of leaf flavonoid chemistry (Table 1 and Fig. 2).

To confirm the efficiency of the AFLP DNA analyses in the current study, we extracted and analyzed DNA from a feral, non-related *Erythroxylum* taxon and used the same five-AFLP-primer set for the analyses. Ten accessions of *Erythroxylum aerolatum* L. were evaluated by AFLP DNA analysis and appear in the dendrogram as Eaero HAW (Fig. 2; blue). All are similarly clustered in the dendrogram with a coefficient (0.58) which is indicative of a high degree of differentiation from the remaining samples tested (Fig. 2).

2.4. Identification of Colombian coca

To further elucidate the genetic relationship of coca currently under cultivation in Colombia, a series of coca leaf collections were made by ARS in the regions of highest coca production, during 1997–2001. These were done as an ancillary task to the main coca verification missions, i.e., to assess periodically the effectiveness of the ongoing coca eradication program (through aerial spraying of a herbicide). As expected, such collection is a relatively high risk endeavor.

Field samples of the coca plants were collected and estimates of each field's original planting date were made. Since estimating coca age is difficult (there are no "tree rings" to count), the date provided is the latest year in which planting could have occurred for each accession. Fig. 3 depicts an expanded portion of the dendrogram of Fig. 2; in it, the clustered linkages of the suspected Colombian *E. c.* var. *ipadu* (Esp COL; Johnson et al., 2002, 2003) are more easily studied. Because the provenance of all samples (save one) was exactly known, several clusters identifiable on the Fig. 3 dendrogram can be precisely localized geographically within Colombia (Fig. 4). An example is the genetic linkage group labeled "Cluster A" in Fig. 3, comprised of accession numbers Esp 098, 100, 118, 103, 115, and 117. They were exclusively located within northern Guaviare Department [state] and along the border area of adjacent Meta Department, indicating that a regionalized population of *E. c.* var *ipadu* has been developed in that area within the last few decades (Figs. 3 and 4). Another small, genetically-linked cluster ('E') of accessions identified as Esp 116, 133, 123, and 139 were located in the southwestern Departments of Caqueta, Cauca, and Putumayo; these were scattered among another larger, but genetically distinct cluster ['C'; accessions Esp 105, 110, 120, 106, 119, 131, 122, 129, 130, 134, 121, 142, 126, 132, and 141]. With the single exception of sample Esp 110, from the northeastern Colombia state of Norte de Santander, all accessions from the population defined by clusters C and E were
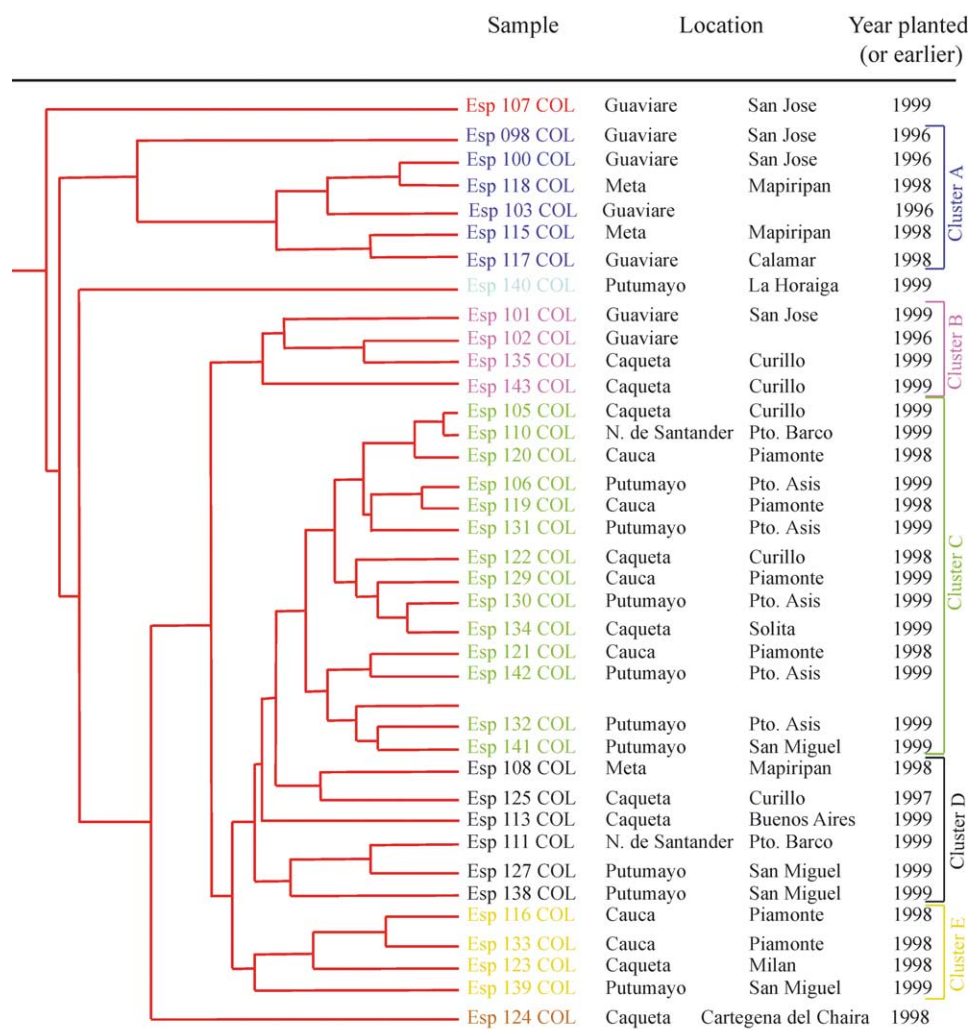
Fig. 3. Colombian sample geography and planting date.

exclusive to Putumayo, Cauca, and Caqueta. Two other genetic clusters ('B' and 'D', Fig. 3) appeared in the Putumayo/Caqueta region, but also were found in other areas of Colombia. Three other unrelated populations of *E. c.* var. *ipadu* (Esp 124, 107, and 140) which were not genetically linked to either the five main clusters or to each other, were found in both the Meta/Guaviare as well as the Putumayo/Caqueta coca-growing nuclei.

From the analysis of the genetic data, it appears that regionalized populations of *E. c.* var. *ipadu* are being developed rapidly in Colombia in a number of areas. The main center for these new populations is the Putu-mayo/Caqueta region of Colombia, although other parts of the country may have some form of independent coca varietal development programs. At least 5–8 new populations of *E. c.* var *ipadu* have been or are being developed and several of those populations have gained regional widespread use. It is also clear that additional new populations are being developed and field tested within several regions in Colombia, probably

relatively informally and independently once the germplasm (as cuttings or seeds) is brought into new regions. There is ample evidence to indicate that, although some regional populations of coca are being restricted to a local area, in time, popularized high-producing lines are being distributed to other regions of the country that are suitable for growing the illicit coca. Additional sampling and analyses will, in conjunction with other sources of information, further solidify the trends underway in Colombia. The same AFLP DNA fingerprinting approach should prove valuable for characterizing coca grown elsewhere as well.

3. Experimental

3.1. Plant material

3.1.1. Isolation of DNA from leaf tissue
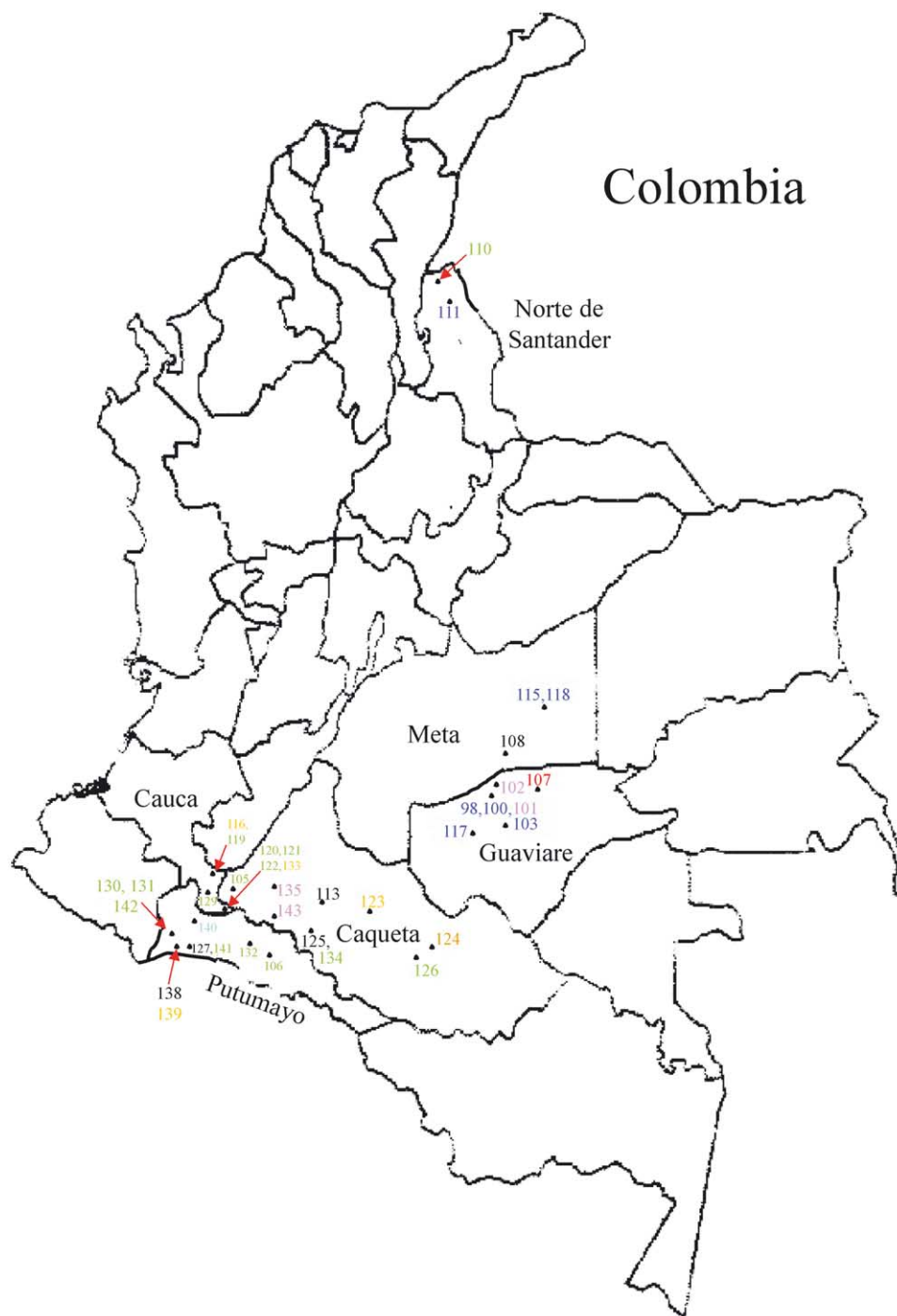Genomic DNA (i.e., total DNA) was extracted using a modification of the Qiagen Dneasy Plant Mini Kit™

Fig. 4. Colombian map of collection sites of unidentified *Erythroxylum* spp. (1997–201). Numbers directly correlate to dendrogram samples.

protocol (Qiagen Inc., Valencia, CA, USA). An equivalent of 50 mg dry weight of tissue was placed into a 2-ml lysing matrix cylinder that contained two 0.53-cm ceramic spheres (QBiogene, Inc., Carlsbad, CA, USA). Buffer API (650 μl) and RNase stock (2 μl) solution were added and mixed thoroughly. Tissue was then homogenized using a FastPrep FP120 Instrument (Savant Instruments, Holbrook, NY, USA) at a speed setting of 6.5 for 45 s. Each sample

was incubated for 10 min at 65 °C in an Eppendorf Thermomixer (Eppendorf, Hamburg, Germany). During incubation, samples were mixed several times by inverting the cylinders. AP2 buffer (160 μl) was added to the lysate and incubated for 5 min at 0 °C (to precipitate the detergent, proteins, and polysaccharides). The samples were then centrifuged at 16,100 g (23 °C) for 5 min, in an Eppendorf Centrifuge 5415D. The supernatant was transferred to a QIAshredder spin column situated

in a 2-ml collecting tube and centrifuged as above for 2 min. The eluant was transferred to a 2-ml microcentrifuge tube and 1.5 volumes of buffer AP3/E were added and mixed thoroughly. Buffered eluant (650 µl) was loaded onto a DNeasy mini spin column and centrifuged for 1 min at 6000 g. The eluant was discarded and the step repeated several times (total buffered eluant volume ≈2 ml). The DNeasy column was placed into a 2-ml collection tube and 500 µl of buffer AW were added and centrifuged for 1 min as above. The eluant was discarded and the wash repeated twice. The DNeasy column was transferred to a 1.5-ml microcentrifuge tube and 100 µl of preheated (65 °C) buffer AE was added. The sample was incubated for 5 min at room temperature and centrifuged for 1 min at 6000 g. This step was repeated, resulting in a final volume of 200 µl of extracted DNA.

3.1.2. Harvest and storage of leaf material

Leaf tissue from the various locations was transported to the USDA-ARS laboratory at the Beltsville Agricultural Research Center, Beltsville, MD. Accessions from Beltsville (B or BELT suffix; Fig. 2) were isolated from plants authenticated by Drs. T. Plowman (1988) and P. M. Rury (1993). Other accessions listed in Fig. 2 with the following suffixes were obtained as follows: HAW = Hawaii, PERU = Peru, BOL = Bolivia, Esp = Colombia-suspected *E. c.* var. *ipadu*, DEA = DEA-provided plant. Lower case suffixes are indicated as follows: r = replicate samples, s = seedlings, c = cuttings. Replicate samples (r) were complete replicates of AFLP and fragment analysis procedures from the same sample isolate. Seedlings (s) are presumed hybrids but are designated with the species from which the seed was collected. Because *E. c.* var. *ipadu* is a true cultigen, production of new plants can theoretically only be made through the use of cuttings (c).

3.1.3. DNA quantification

DNA was quantified by UV (260/280 nm ratios) and fluorimetry in all samples to reduce variability in analyses. All UV measurements for DNA were determined using the 260/280 ratio on 20 µl of the DNA extract and 680 µl of the eluting buffer with a Beckman DU® 530 Series UV/VIS Spectrophotometer (Beckman Instruments, Fullerton, CA, USA). All sample measurements were replicated three times to assure precision. For fluorimetric DNA quantitation, samples were diluted 1:20 (with TE) and analyzed using the PicoGreen dsDNA quantitation kit (Molecular Probes, Eugene, OR, USA) in a Fluoroskan Ascent microplate reader equipped with 485/538 excitation/emission filter settings (LabSystems, Helsinki, Finland). In a 96-well platform plate (Greiner Bio-one, Longwood, FL, USA), 50 µl of working PicoGreen solution (1:200 dilution), 2.5 µl of dilute DNA (1:20), and 47.5 µl of de-ionized water were added to each well. A standard curve was generated from DNA standards (PicoGreen kit) ranging from 10 to 500 ng/ml on the same plate. The final sample dilution was 1:800 and all measurements were repeated three times. The fluorimetric DNA quantitation technique was routinely checked against UV measurements for quality control.

3.2. Amplified fragment length polymorphisms (AFLP)

DNA fragments were amplified using the procedure by Vos et al. (1995) modified as follows. Template DNA (500 ng) was digested by *Eco*RI and *Mse*I (New England BioLabs, Beverly, MA) and ligated to commercial *Eco*RI and *Mse*I oligonucleotide adapters (Applied Biosystems, Foster City, CA) in a single step, with incubation overnight at room temperature. Solutions were as described previously, except as noted below (Saunders et al., 2001a,b).

The first (preselective) amplification of the restricted and ligated fragments utilized commercial *Eco*RI and *Mse*I AFLP preselective primers and AFLP core mix (both from Applied Biosystems). The thermocycling program for this amplification was: 94 °C for 3 min, followed by 20 cycles of the following profile: 94 °C for 20 s, 56 °C for 30 s, and 72 °C for 2 min with a final hold of 60 °C for 30 min (GeneAmp® 9700 PCR system, Applied Biosystems). The products from the preselective amplification were diluted as described previously (Saunders et al., 2001a,b) and used as templates for the selective amplification. For the second (selective) amplification, custom primers (Table 2) were used, with a WellRED™ active ester dye (D1, D2 or

Table 2
Sequences of primers used for selective amplification. Unique sequences and dyes are shown in bold type

Primer pair	Forward (*Eco*RI) primer (5′→3′)	Reverse (*Mse*I) Primer (5′→3′)	# Polymorphic peaks
1	**D4**–GACTGCGTACCAATTC**ACT**	GATGAGTCCTGAGTAA**CAT**	67
3	**D4**-GACTGCGTACCAATTC**ACA**	GATGAGTCCTGAGTAA**CAA**	41
4	**D2**-GACTGCGTACCAATTC**AAC**	GATGAGTCCTGAGTAA**CAC**	76
6	**D3**-GACTGCGTACCAATTC**AAG**	GATGAGTCCTGAGTAA**CTA**	94
7	**D4**-GACTGCGTACCAATTC**ACC**	GATGAGTCCTGAGTAA**CTA**	76

D4) added to the 5′ end of each *Eco*RI primer (Research Genetics Inc., Huntsville, AL). No modifications were made to the *Mse*I primers. The labeled *Eco*RI primer was included at 0.05 μM and the unlabeled *Mse*I primer was included at 0.10 μM. The thermocycling profile was: 94 °C for 2 min, followed by 10 cycles of 94 °C for 20 s, 1 °C per cycle stepdown of annealing temperature from 66 °C, held for 30 s, and 72 °C for 2 min. This was followed by 25 cycles of 94 °C for 20 s, 56 °C for 30 s, and 72 °C for 2 min, with a final hold of 60 °C for 30 min.

A number of primer pairs were evaluated over a 4-year period and while all of these were successful, only primers with optimal resolution were selected for this study. Our previous studies with *Erythroxylum* and other crops (Saunders et al., 2001b) indicated that three primer pairs were sufficient to determine relationships in a population of this size (132 samples). To be conservative, we included five different primer pairs; our inclusion of two additional primers did not change the structure of the dendrogram.

Samples were prepared for analysis by diluting the final amplified product 1:30 (v/v) in commercial Sample Loading Solution (Beckman-Coulter, Fullerton, CA), which included 1% (v/v) Beckman-Coulter 400 BP DNA size standards. Fragment separation and detection was performed by a genetic analysis system (CEQ 8000™, Beckman-Coulter, Fullerton, CA); this used capillary electrophoresis on each sample, beginning with a 30-s electrophoretic injection at 2.0 kV, and 35 min separation at 50 °C and 6.0 kV. For quality assurance, both samples and AFLPs were periodically replicated.

The fragment data was analyzed using CEQ 8000™ software, with analysis parameters that recognized peaks with slope of 5 and 15% of the height of the second-largest peak, except for primer 3 (Table 2), which recognized peaks that were 10% of the second largest peak height. The size standard fit coefficient was 0.38 and it used a cubic model. For AFLP analysis, the maximum bin width was 1.00 nt, and no further Y threshold was applied. Each sample was scored for each bin as "1" if a fragment of that size was present, and "0" if not. Fully populated bins, sparsely populated bins containing 1, 2 or 3 members, and unpopulated bins were excluded. In rare cases where two fragments were present in one bin, the bin was scored as "1" for that sample. The number of polymorphic fragments produced by each primer set are listed in Table 2.

Data were imported into the multivariate data analysis program, NTSYSpc, (version 2.1, September 2000; Exeter Software, Setauket, NY). The module SimQual was used to generate a similarity matrix based on simple matching coefficients, and the module SAHN performed sequential, agglomerative, hierarchical and non-overlapping clustering by unweighted pair-group method, arithmetic average (UPGMA). The module Tree Plot was used to convert the data generated by SAHN into a dendrogram.

References

Amane, M., Lumaret, R., Hany, V., Ouazzani, N., Debain, C., Vivier, G., Deguilloux, M.F., 1999. Chloroplast-DNA variation in cultivated and wild olive (*Olea europaea* L.). Theor. Appl. Genet. 99, 133–139.

Bohm, B.A., Ganders, F.R., Plowman, T., 1982. Biosystematics and evolution of cultivated coca (Erythroxylaceae). Syst. Bot. 7, 121–133.

Bhattramakki, D., Rafalski, A., 2001. Discovery and application of single nucleotide polymorphism markers in plants. In: Henry, R.J. (Ed.), Plant Genotyping; The DNA Fingerprinting of Plants. CIBA Publishing, New York, pp. 179–192 (Chapter 12).

Casale, J.F., Moore, J.M., 1994. 3′,4′,5′-Trimethoxy-substituted analogs of cocaine, *cis/trans*-cinnamoylcocaine and tropacocaine: characterization and quantitation of new alkaloids in coca leaf, coca paste and refined illicit cocaine. J. Foren. Sci. 39, 462–472.

Casale, J.F., Moore, J.M., 1996. Lesser alkaloids of cocaine-bearing plants. III. 2-Carbomethoxy-3-oxo substituted tropane esters: detection and gas chromatographic-mass spectrometric characterization of new minor alkaloids found in South American *Erythroxylum coca* var. *coca*. J. Chromatogr. A 756, 185–192.

Ehleringer, J.R., Casale, J.F., Lott, M.J., Ford, V.L., 2000. Tracing the geographical origin of cocaine. Nature 408, 311–312.

Ganders, F.R., 1979. Heterostyly in *Erythroxylum coca*, (Erythroxylaceae), Bot. J. Linn. Soc. 78, 11–20.

Helling, C.S., 1993. Coca Eradication in Panama: Final Techn. Rep. Agric. Res. Service, US Dept. Agric, Beltsville, MD, USA.

Holmstedt, B., Jäätmaa, E., Leander, K., Plowman, T., 1977. Determination of cocaine in some South American species of *Erythroxylum* using mass fragmentography. Phytochemistry 16, 1753–1755.

Holmstedt, B., Lindgren, J.E., Rivier, L., Plowman, T., 1979. Cocaine in blood of coca chewers. J. Ethnopharmacol 1, 69–78.

Johnson, E.L., 1996. Alkaloid content in *Erythroxylum coca* tissue during reproductive development. Phytochemistry 35, 35–38.

Johnson, E.L., Emche, S.D., 1994. Variation of alkaloid content in Erythroxylum coca leaves from leaf bud to leaf drop. Ann. Bot. 73, 645–650.

Johnson, E.L., Foy, C.D., 1996. Biomass accumulation and alkaloid content in leaves of *Erythroxylum coca* and *Erythroxylum novogranatense* var. *novogranatense* grown in soil with varying pH. J. Plant Phys. 149, 444–450.

Johnson, E.L., Schmidt, W.F., 1999. Flavonoids as chemotaxonomic markers for *Erythroxylum ulei*. Z. Naturforsch. 54c, 881–888.

Johnson, E.L., Schmidt, W.F., Norman, H.A., 1997. Leaf flavonoids as chemotaxonomic markers for two *Erythroxylum taxa*. Z. Naturforsch. 52c, 577–585.

Johnson, E.L., Schmidt, W.F., Norman, H.A., 1998. Flavonoids as markers for *Erythroxylum* taxa: *E. coca* var. *ipadu* and *E. novogranatense* var. *truxillense*. Biochem. Sys. Ecol. 26, 743–759.

Johnson, E.L., Schmidt, W.F., Cooper, D., 2002. Flavonoids as chemotaxonomic markers for cultivated Amazonian coca. Plant Physiol. Biochem. 40, 89–95.

Johnson, E.L., Schmidt, W.F., Emche, S.D., Mossoba, M.M., Musser, S.M., 2003. Kaempferol (rhamnosyl) glucoside, a new flavonol from *Erythroxylum coca* var. *ipadu*. Biochem. Sys. Ecol. 31, 59–67.

Li, X., Ralphs, M.H., Gardner, D.R., Wang, R.R.-C., 2002. Genetic variation within and among 22 accessions of three tall larkspur species (*Delphinium* spp.) based on RAPD markers. Biochem. Sys. Ecol. 30, 91–102.

Moore, J.M., Casale, J.F., Klein, R.F.X., Cooper, D.A., Lydon, J., 1994. Determination and in-depth chromatographic analyses of

alkaloids in South American and greenhouse-cultivated coca leaves. J. Chromatogr. A 659, 163–175.

Plowman, T., 1979. The identity of Amazonian and Trujillo coca. Bot. Mus. Leaf. 27, 45–68.

Plowman, T., 1980. The identity of Amazonian and Trujillo coca. Bot. Mus. Leaf. 27, 253–261.

Plowman, T., 1981. Amazonian coca. J. of Ethnopharmacol 3, 195–225.

Plowman, T., 1982. The identification of coca (Erythroxylum species): 1860–1910. Bot. J. Linn. Soc. 84, 329–353.

Plowman, T., 1984. The ethnobotany of coca (*Erythroxylum* spp., Erythroxylaceae). Adv. Econ. Bot. 1, 62–111.

Plowman, T., Rivier, L., 1983. Cocaine and cinnamoylcocaine content of *Erythroxylum* species. Ann. Bot. 51, 641–659.

Reens, E., 1919. La coca de Java. Bull. Sci. Pharmacol. 12, 497–505.

Rivier, L., 1981. Analysis of alkaloid in leaves of cultivated *Erythroxylum* and characterization of alkaline substances used during coca chewing. J. Ethnopharmacol. 3, 313–335.

Rury, R.P., 1981. Systematic anatomy of *Erythroxylum* P. Browne: practical and evolutionary implications for the cultivated cocas. J. Ethnopharmacol. 3, 229–263.

Saunders, J.A., Mischke, S., Hemeida, A.A., 2001a. The use of AFLP techniques for DNA fingerprinting in plants. Beckman Coulter Application Notes A1910A, pp. 1–9.

Saunders, J.A., Pedroni, M.J., Penrose, L.D.J., Fist, A.J., 2001b. AFLP DNA analysis of opium poppy. Crop Sci. 41, 1596–1601.

Schultes, R.E., 1981. Coca in the Northwest Amazon. J. Ethnopharmacol. 3, 173–194.

Soleimani, V.D., Baum, B.R., Johnson, D.A., 2002. AFLP and pedigree-based genetic diversity estimates in modern cultivars of drum wheat [Triticum turgidum L. Subsp. drum (Desf.) Husn.]. Theo. Appl. Genet. 104, 350–357.

Turner, C.E., Ma, C.Y., ElSohly, M.A., 1981. Constituents of Erythroxylum coca. II. Gas chromatographic analysis of cocaine and other alkaloids in coca leaves. J. Ethnopharmacol. 3, 293–298.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kulper, M., Zabeau, M., 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23, 4407–4414.

Welsh, J., McClelland, M., 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 18, 7213–7218.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., Tingey, S.V., 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18, 6531–6535.

Youssefi, M., Cooks, R.G., McLaughlin, J.L., 1979. Mapping of cocaine and cinnamoylcocaine in whole coca plant tissue. J. Am. Chem. Soc. 101, 3400–3402.