SCIENTIFIC REPORTS



SUBJECT AREAS: BIOINFORMATICS EARTH AND ENVIRONMENTAL SCIENCES PLANT SCIENCES BIODIVERSITY

> Received 25 May 2011

> > Accepted 6 July 2011

Published 21 July 2011

Correspondence and requests for materials should be addressed to M.Y.S. (mark.stoeckle@rockefeller.edu)

Commercial Teas Highlight Plant DNA Barcode Identification Successes and Obstacles

Mark Y. Stoeckle¹, Catherine C. Gamble², Rohan Kirpekar², Grace Young², Selena Ahmed³ & Damon P. Little⁴

¹Program for the Human Environment, The Rockefeller University, New York, NY 10065, USA, ²Trinity School, New York, NY 10024, USA, ³Department of Biology, Tufts University, Medford, MA 02115, USA, ⁴Lewis B. and Dorothy Cullman Program for Molecular Systematics, The New York Botanical Garden, Bronx, New York, NY 10458, USA.

Appearance does not easily identify the dried plant fragments used to prepare teas to species. Here we test recovery of standard DNA barcodes for land plants from a large array of commercial tea products and analyze their performance in identifying tea constituents using existing databases. Most (90%) of 146 tea products yielded *rbcL* or *matK* barcodes using a standard protocol. Matching DNA identifications to listed ingredients was limited by incomplete databases for the two markers, shared or nearly identical barcodes among some species, and lack of standard common names for plant species. About 1/3 of herbal teas generated DNA identifications not found on labels. Broad scale adoption of plant DNA barcoding may require algorithms that place search results in context of standard plant names and character-based keys for distinguishing closely-related species. Demonstrating the importance of accessible plant barcoding, our findings indicate unlisted ingredients are common in herbal teas.

queous infusions prepared from dried plants, broadly known as teas, are popular beverages with desirable physiologic activities and potential health benefits. Accurate labeling is important for consumers, marketers, and regulators, as tea constituents cannot be easily identified to species by visual appearance. Their taxonomic diversity and fragmentary nature present a ready and demanding test of DNA-based identification. Here we report the successes with and obstacles to identifying tea ingredients using a short DNA sequence from a uniform locality within the genome, DNA barcoding¹.

Tea properly refers to infusions prepared from leaves of the tea plant, *Camellia sinensis* (L.) Kuntze, an evergreen flowering tree in the family Theaceae, native to the mountainous regions of southwestern China and neighboring countries²⁻⁴. The two main commercial varieties are small-leafed *C. sinensis* var. *sinensis*, adapted to cool weather and high altitude, and large-leafed *C. sinensis* var. *assamica* (J. W. Mast.) Kitam., which grows well in tropical and sub-tropical environments. Tea plant leaves contain a high concentration of phytochemicals including polyphenolic catechins and the methylxanthine caffeine⁵⁻¹¹. Tea drinking originated in southern China at least 2000 years ago, and today tea is the most widely consumed beverage in the world^{12,13}. Different processing methods, ranging from drying and baking to months of microbial fermentation, produce the variety of tea types—white, green, black, oolong, and pu-erh—which differ in catechin content and antioxidant activity^{14,15}.

In addition to *C. sinensis*, infusions are prepared from a diversity of other plants and plant parts—beverages also commonly referred to as tea. In the following we use "CS" to indicate *C. sinensis* and "herbal" for other plants. Some herbal teas have pharmacologically active compounds and may have therapeutic or toxic effects. Fatalities and serious illnesses have occurred after drinking herbal teas, caused by overdose, mislabeled products, or allergic reactions¹⁶⁻¹⁸.

In 2009, the Plant Working Group of the Consortium for the Barcode of Life (CBOL) endorsed a proposal to use defined portions of the plastid genes *rbcL* (~550 bp segment) and *matK* (~790 bp segment) as standard barcodes for land plants¹⁹. These and other candidate markers have been tested in various floristic and taxonomic settings^{20–24}. As compared to animals, plants generally have less barcode variation both within and among species. A relatively large proportion of plants (~15%–30%) share barcodes among multiple species. Plant barcodes generally do not exhibit the strong clustering pattern observed in most animal species (intraspecific variation). These observations apply even when longer sequences or additional markers are sampled,

which may reflect fundamental differences in plant and animal biology and evolution²³. Notwithstanding these limitations, standard plant barcodes are efficacious in a number of scientific and applied settings and have enormous potential for wider use²⁵.

In this study we explored a practical application of plant barcoding: matching commercial tea ingredients to product labels. We searched a public reference database for the closest match to each barcode sequence and compared the result to the listed ingredients. Because the tea specimens are morphologically unrecognizable, we cannot know with certainty if the source plants are represented in the reference database, a realistic and difficult test of barcode identification.

Results

Barcode recovery, haplotypes, matches. Using single sets of primers for each locus, readable *rbcL* or *matK* barcodes were recovered from 131 (90%) of 146 tea products, including 96% of CS and 84% of herbal teas. *rbcL* was recovered from 113/146 (77%), *matK* from 108/136 (79%), and both from 90/136 (66%). A total of 253 readable sequences were obtained, comprising 48 *rbcL* and 40 *matK* haplotypes (Figs. 1,2; additional details in Supplementary Tables S1,S2 online). There were no insertions or deletions in *rbcL* sequences; the *matK* alignment contained 14 different types of insertions or deletions. For each haplotype, BLAST searches of

GenBank and Barcode of Life databases were performed. The closest match in each database was recorded. As compared to results with GenBank, BOLD matches were on average lower identity and fewer were label ingredients, indicating that at the time of the study BOLD was less well populated with barcodes of plants used in commercial tea products. As a result, subsequent analyses were performed using GenBank. The *rbcL* haplotypes matched 42 species in 24 families; the *matK* haplotypes matched 25 species in 16 families (Figs. 1,2).

Taking into account uncertainties arising from incomplete databases, shared barcodes, and ambiguous common names, of 48 *rbcL* haplotypes, 32 were assigned to species, 10 to genus, and 6 to family. Of 40 *matK* haplotypes, 27 were assigned to species, 8 to genus, and 5 to family (Figs. 1,2). In most cases (58%), barcodes recovered from commercial tea products matched listed ingredients. It should be noted that our study was designed to enable comparison between CS and herbal teas, and not among individual products or manufacturers. Given this and potential liability issues, we assigned arbitrary alphanumeric codes to each product to protect the manufacturer's identity. Most of the barcodes that did not match listed ingredients reflected an incomplete reference database, lacking either a record for the relevant species or a record of an intraspecific variant. For example, an herbal tea labeled "Marshmallow (*Althaea officinalis*)" produced an *rbcL* sequence closest to *Anisodontea*

NN dif

			-							
Polypodiaceae	TT204z-1	Terpischore sp. indet				and made		T THEFT		
Monimiaceae	MKa6-1	Peumus holdus	2272-23	na serie de la comp						
Zingiboracoao	D1547 1	Zingiboracoao		or and						
Zingiberaceae	TT002-1	Augustaceae						-		
Poaceae	11223-1	Avena sativa						-		
	K22A-2	Cymbopogon citratus								
	K24-1	Poa annua	5 - S	22.1.57		2 - 1 C - 1 C	0.000			
Ranunculaceae	TT218A-2	Actaea sp. indet.	a 1 - 1	600 (00) - Co				1115		-
Fabaceae	TT2027-2	Aspalathus linearis		(x,y) = (x,y) = (x,y)				1.110.001		
	R43-1	Glycymhiza glabra		111111		111111111	32254	1.000		
	K267-2	Medicado sativa						J		
	TT200 1	Serve en indet						-		
B	11209-1	Serina sp. indel.					-			
Rosaceae	K28-2	Rubus sp. Indet.						-		_
Urticaceae	11208-1	Urtica doica		10000				- 1997	1.1	
Myrtaceae	G13-1	Psidium guajava		1.545.57(1955)		10.0	1000	11101	111	
Rutaceae	K22B-1	Citrus sinensis				1000				
Malvaceae	TT219-1	Althaea officinalis		10000				1.2.2.4		
martabbab	K29-1	Hibiscus sp indet		00000				·		1.00
Polygonaceae	TT213-1	Rumey japonicus				100	in the	·		
Theorem	C16 20	Comollio sinonsis		200				-		
Theaceae	010-50					1.1		-		
	G20-2	Carriellia sinensis					_	-		
	K19-3	Camellia sinensis					_	-		-
	R41-1	Camellia sinensis								
	R9-32	Camellia sinensis		12.5			-	-		
Oleaceae	R4A-1	Jasminum sp. indet.		and the second second			1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	A	1.0.00	
Scrophulariaceae	TT203-1	Verbascum thapsis		2003						1.00
Lamiaceae	K27-6	Lamiaceae	52 E.	1933303033333				100		
	R10B-1	Lamiaceae	55.55	weedeed				11111		
	R8D-1	Lamiaceae		1723-184 (1956)				1000		
	TT205-1	Lavandula stoechas					- 1 C			
	TT225-1	Melissa officinalis		0000000						
	TT222-1	Neneta cataria								
	G6-1	Prinella vulgaris	o o nore	montern		Sec. Sec.	······	-		
	P23 1	Prinella vulgaris						-	and the	
Orehenskesses	TT207_1	Odontitos vomus						-		
Orobanchaceae	MCm17.1	Lontono on indet						-		
verbenaceae	NG1117-1	Lanana sp. Indel.								
Aquifoliaceae	AC3-1	liex paraguariensis			·			-	1.1	
	G31-1	llex sp. Indet.							1.000	
Asteraceae	TT215A-2	Asteraceae		2.C. 2			_	-		
	11211-1	Achillea millefolium	5 C (1	10000000000		18 JU 19	100	1000		1000
	TT202-1	Echinacea purpurea	00.00	22220000				111.01		1277
	K15-9	Matricaria recutita		A CONTRACTOR						
	MRa6-1	Stevia rebaudiana						1.000		1
Caprifoliaceae	G10-1	Lonicera sp. indet.		sected as		210 11 22	1.110	201220	311	
Pittosporaceae	K22v-1	P. pentandrum		the second second			37325			
Araliaceae	AC1-1	Panax ginseng	22722	nesseicus.		and show		d and a	and the	
Alalaceae	AC2-2	Panax quinquefolius	22				1000	S TRACT	1000	
Anionno	R45-1	Aniaceae						-		
Aplaceae	R8y-1	Aniaceae					· · · · ·			
	1.07-1	huddac						-		
			97 98	3 99 1	100					
			% ir	dentity						

Closest match NN same

Figure 1 | *rbcL* barcode identifications. For each haplotype, alphanumeric code, number of isolates, identification, and graphic representation of match results are shown. Color bars depict percent identity of closest match, nearest neighbor (NN) in the same genus, and NN in a different genus, with scale at bottom. Haplotypes for which the second closest match was in a different genus have a blank in "NN same genus" column. (Note: *P. pentandrum = Pittosporum pentandrum*).





Figure 2 | *matK* barcode identifications. For each haplotype, alphanumeric code, number of isolates, identification, and graphic representation of match are shown as described in Fig. 1 legend.

triloba (1 mismatch, 99.8% identity). However, at the time of the study there were no GenBank *rbcL* records for *A. officinalis*. Overall, at the time of the study about one-third of plant species listed on product labels lacked *rbcL* or *matK* records in GenBank. Reflecting incomplete representation of intraspecific variants, more than half of *C. sinensis* tea products yielded an *rbcL* barcode 100% identical to congeneric species *C. oleifera* and *C. sasanqua* but with one mismatch compared to the *C. sinensis rbcL* record.

Barcode identifications were incompatible with listed ingredients for some products, including 21/60 (35%) herbal and 3/70 (4%) CS teas (Table 1). Some of the non-label DNAs matched plants used in other tea products, some matched common weeds or other non-food plants, and some could not be identified. The most common non-label ingredient, found in seven products, was chamomile (Matricaria recutita). Four herbal teas yielded sequences identified as tea plant (C. sinensis), although none listed ingredients in the tea family (Theaceae). Regarding non-food plants, a product labeled "St. John's wort (Hypericum perforatum)," a flowering plant, yielded an *rbcL* sequence identical to that of several fern species. A barcode from an herbal tea matched Poa annua, a widely cultivated meadow grass. Four products yielded barcodes closely matching plants in Apiaceae, the parsley family, although the particular species could not be determined. Apiaceae includes many food plants and ubiquitous wild relatives, but for the products in question none of the listed ingredients were in this family.

Taxonomic resolution. For most *rbcL* haplotypes, the differences between closest match, nearest neighbor (NN) in the same genus, and NN in a different genus were modest or absent. Among the 48

haplotypes, the average percent identity was 99.9% for closest, 99.8% for congeneric NN, and 99.2% for NN in a different genus, or about 0.6, 1.1, and 4.6 nucleotide differences respectively (Fig. 1; additional details in Supplementary Table S1 online). Of 32 *rbcL* haplotypes with 100% match, 15 were also identical to one or more congeneric species and eight were identical to one or more species in a different genus.

For *matK*, the average identities were 99.5% for closest match, 99.5% for NN congeneric, and 98.1% for NN different genus, or about 3.8, 3.8, and 14.3 nucleotide differences (Fig. 2; additional details in Supplementary Table S2 online). Of 14 haplo-types with a 100% match, three were also identical to one or more congeneric species, and none were identical to species in a different genus.

C. sinensis rbcL nucleotide sequence polymorphism. We observed nucleotide variation (A or C) in CS *rbcL* sequences at a site corresponding to position 68 of the coding region (gi 7525012:54958-56397 was used as a reference), with the predicted predicted amino acid being either asparagine (68A) or threonine (68C). The 68A sequence was identical to the *C. sinensis rbcL* GenBank record, whereas the 68C variant was identical to *rbcL* sequences of several congeneric species (*C. albogigas, C. granthamiana, C. japonica, C. oleifera, C. sasanqua*) and a related species *Tutcheria hirta*. Among tea products for which geographic or tea type information was available, the 68C variant was associated with products from India as compared to China (94% vs. 31%, p < 0.0001) and with black vs. green tea (93% vs. 19%, p < 0.0001). Among vouchered specimens, the 68C variant was strongly associated with *C. sinensis*



Table 1 | DNA barcode identification of unlisted ingredients.

Product:	Gl
Label:	apple pieces, vitamin C, citric acid, natural flavor
Comment:	G1 <i>matK</i> 100% identity to <i>Camellia sinensis</i> , familiy Theaceae. No listed ingredients in Theaceae.
Product:	G2 apple pieces, orange peel, rosehins, hibiscus, comflower blossoms, clove, cinnamon, anise, pepper, natural flavor
Non-label DNA:	chamomile (Matricaria recutita)
Comment:	G2 rbcL 99.6% identity Pentzia incana, G2 matK 98.9% match Achillea millefolium, both family Asteraceae. rbcL and matK sequences most likely represent chamomile (Matricaria recutita), based on 100% match to partial M. recutita sequences in GenBank and recovery of identical or nearly identical sequences from products listing chamomile as sole ingredient. Listed ingredient cornflower is not an approved name for chamomile ³⁵ . It refers to Centaurea cyanus, family Asteraceae. Compared to closest match, G2 sequence is more distant from C. cyanus rbcL (96.8%) (AB530955); no C. cyanus matK sequences in GenBank. Other ingredients in different families.
Product:	G4
Label:	raspberry pieces, apple pieces, orange peel, rosehips, hibiscus, lemongrass, vitamin C, natural raspberry tlavor
Comment:	G4 matK 98.9% match Achillea millefolium, family Asteraceae, most likely representes chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Asteraceae.
Product:	G6
Label:	Prunella vulgaris
Comment:	G6 matK 98.9% match Achillea millefolium, family Asteraceae, most likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Asteraceae.
Product:	G10 honourualle flaver
Non-label DNA:	chamomile (Matricaria recutita)
Comment:	G10 matK 98.9% identity Achillea millefolium, family Asteraceae, most likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Asteraceae.
Product:	K21
Non-label DNA	tea (Camillia sinensis)
Comment:	K21 <i>rbcL</i> and <i>matK</i> 100% match <i>Camellia sinensis</i> , family Theaceae. No listed ingredients in Theaceae.
Product:	K22
Label:	rosehips, orange peel, chamomile flowers, lemongrass, lemon myrtle, hibiscus flowers, nana mint, natural citrus flavors and other natural flavors
Non-label DNA: Comment:	Taiwanese cheesewood (Pittosporum pentandrum) K22 rbcL and matK 100% match Pittosporum pentandrum (Taiwanese cheesewood), family Pittosporaceae. No listed ingredients in Pittosporaceae.
Product:	K24
Label:	ginger root, natural tlavors, linden, lemon peel, blackberry leaves, lemongrass, citric acid
Comment:	K24 rbcL 100% match to Poa annua (annual bluegrass), family Poaceae. Listed ingredient lemongrass, Citropogon citratus, is also in Poaceae. However, compared to closest match, K24 sequence is more distant from C. citratus (94.0%) (GQ436383). No other ingredients in Poaceae.
Product:	K27 eleuthera peppermint cinnamon ainger chamomile west indian lemonarass licarice cathin tilia flowers natural lemon flavor
	hops, vitamins B6 and B12 white accesses (Chancendum album)
Comment:	In GenBank, K27 matK 99.2% match Rhagodia baccata and Chenopodium album, both family Amaranthaceae. In BOLD, K27 matK is 100% match to Chenopodium album. No listed ingredients in Amaranthaceae.
Product:	R8
Label: Non-label DNA	mate, licorice, rosehips, mint, pineapple chunks, natural flavors tea <i>l'Camellia</i> sinensist, parslev family (Apiaceae)
Comment:	R8 <i>rbcL</i> 1 100% match to <i>Camellia oleifera</i> , family Theaceae. No listed ingredients in Theaeceae. R8 <i>rbcL</i> 2 99.8% and <i>matK</i> 99.7% match <i>Pimpinella saxifraga</i> , family Apiaceae. No listed ingredients in Apiaceae.
Product:	R10
Label: Non-label DNA	rea, iemongrass, iemon verbena, spearmint, natural tiavors chamomile (Matricaria recutita), skullcap (Scutellaria barbata)
Comment:	R10 rbcL 99.6% match Pentzia incana, R10 matK1 98.9% match Achillea millefolium, both family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed Ingredients in Asteraceae. R10 matK2 99.7% match Scutellaria barbata, family Lamiaceae. Listed Ingredient spearmint (Mentha spicata) is also in Lamiaceae. However compared to closest match, R10 matK2 is relatively distant from M. spicata (91.8%) (GU381684). No other ingredients in Lamiaceae.
Product:	KIS black teg plus regibes black pepper, cardament cingament cingar, organic care sugar, patient flavors
Non-label DNA:	chamomile (Matricaria recutita), alfalfa (Medicaao sativa)
Comment:	R15 rbcL99.6% match Pentzia incana and R15 matK1 98.9% match Achillea millefolium, both family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No ingredients in Asteraceae. R15 matK2 99.7% match Medicago sativa, family Fabaceae. Listed ingredient rooibos (Aspalathus linearis) is also family Fabaceae. No A. linearis matK sequences in GenBank for comparison, but A. linearis and M. sativa rbcL sequences show limited identity (93.6%). Other ingredients in different families

•		14	
			1
			L
			٢.
		Ζ.	4
	1.1	1	
		5.4	

Product: K1 Lobel: corob poet, indian sarsaparilla root, ginger root, kava root, cinnamon bark, stevia leaf, cardamom seed, natural flavors, barley mail, essential olis Nonlabel DNA: K1 / fct, 99.8% identity to Camellia olaifera, family Theoceaee. No listed ingredients in Theoceaee. Nonlabel DNA: pertod, coronge essone, ginger, licorica, natural flavors, orange peel, sofficwers, hibiscus flowers, rose petts, coronge essone, ginger, licorica, natural flavors, orange peel, sofficwers, hibiscus flowers, rose petts, coronge essone, ginger, licorica, natural flavors, branch Pimpinella saxifraga, both family Apiaceae. No listed ingredients in Apiaceaee. Product: T1204 Comment: Nat factor 90.1% motch Heteromorpha arborescens, matK 99.7% match Pimpinella saxifraga, both family Apiaceaee. No listed ingredients in Polypodiaceaee. Product: T1204 Comment: T1204 Comment: T1204 Comment: T1207 Comment: T1207 Condoct: T1207 Comment: T1207 Comment: T1207 Colos happing herb [Exphrasia officinalis) Nonlobel DNA: red bortsin [Colonthaceae. The or to a color family sequences in GenBank for direct comparison. However, the closest Exphrasia sequences in GenBank is releatively altoriffor recovered sequence. T1207 rbcl is 97.3% match to Colonthes explain family Corboanchaceaee.	Table 1 (Continue	Table 1 (Continued)				
Lobel: corrob pod, indian sarcaparilla root, ginger root, kava root, cinnamon bark, stevia leaf, cardamom seed, natural flavors, barley mall, essanicial ouis Non-label DNA: teo [Camellia a inensis] Comment: K41 / Ket 99 %: identity to Camellia oleifara, family Theocesee. No listed ingredients in Theocesee. Product: K45 Lobel: lemongross, blackberry leaves, citric acid, rose hips, spearmint, natural flavors, orange peel, sofflowers, hibiscus flowers, rose partisy formit (Apiaceee). Non-label DNA: T1204 S1. John's vort [garia] part] [Hypericum parforatum] Non-label DNA: Non-label DNA: T1204 S1. John's vort [garia] part] [Hypericum parforatum] Non-label DNA: Comment: T005 moth to several species in genus Terpischere, family Polypodiaceae. No listed ingredients in Polypodiaceae. Non-label DNA: experight herb [Exphrosis officinalis] Constent: T207 /bcl 100% and the several species in genus Terpischere, family Polypodiaceae. Non-label DNA: red bortis (Johnties verm) Comment: T1207 /bcl 100% and the several species in genus Terpischere, family Polypodiaceae. Conduct: T1207 /bcl 100% and the several species in genus Terpischere, family Corbonchaceae. Cobolub tred bortis/s AV849664, and T1207 markt is	Product:	R41				
Nonlabel DNA: teol (Camellia sinensis) Comment: K41 / fxcl 99 % identify to Camellia deifera, family Theocesee. No listed ingredients in Theocesee. Product: K45 Label: temongrass, blackberry leaves, citric ocid, rose hips, spearmint, natural flavors, orange peel, safflowers, hibiscus flowers, rose peticls, orange essence, ginger, licarice, natural flavors Product: T1204 Label: temongrass, blackberry leaves, citric ocid, nose hips, spearmint, natural flavors, orange peel, safflowers, hibiscus flowers, rose peticls, orange essence, ginger, licarice, natural flavors Product: T1204 Label: st. Join's wort (acrial part) [<i>Hypericum perforatum</i>] Nonlabel DNA: for (arreline) Comment: T1204 Lobel: epeirigh herb [<i>Exphrasia officinalis</i>] Nonlabel DNA: red bartis (downlifes errorul) Comment: T1207 / rbcl 100% andch be averal species in genus Terpischore, family Polypodiaceae. No listed ingredients <i>Exphrasia officinalis</i> is ados in family Orobonchoceae. There are no <i>Enfinalis</i> (species) for familar of direct comparison. However, the class tells aphrasia species with sequences in Centificalis stead ingredients in Rosaceae. Product: T7210 Lobel: roots (Apadathus lineari), lemongrass (Cymbopagan cirinatu), steria (Steria rebaudian)	Label:	carob pod, indian sarsaparilla root, ginger root, kava root, cinnamon bark, stevia leaf, cardamom seed, natural flavors, barley malt, essential oils				
Comment: R45 Label: Include: R45 Label: Includ	Non-label DNA:	tea (Camellia sinensis)				
 Induct. Induct.	Comment: Product:	R41 rbcL99.8% identity to Camellia oleitera, tamily Theaceae. No listed ingredients in Theaceae.				
Non-Jabel DNA: paraley family (Apicacea) Comment: R45 fcd2 91.% motch Heleromorpha arborescens, mark 99.7% match Pimpinella soxifraga, both family Apiaceae. No listed ingredients in Apicaceae. Product: T1204 Lobel: St. John's wort (earial part) (Hypericum perforatum) Non-Jabel DNA: fem (Tergischere sp. indet). Comment: T1204 rote1 100% match to several spaces in genus Terpischere, family Polypodiaceae. No listed ingredients in Polypodiaceae. Comment: T1207 rbcl. 100% ond mark 99.5% match to Odontifes versus, family Orobanchaceae. Listed ingredient Euphrasia afficialitis sequences in GenBank is relatively distant from recovered sequence: T1207 rbcl. is 97.3% match to E. spectabilits AY849603. Product: T1210 Lobel: robick [Appidritus [meoria], lemongrass [Cymbopogon citratus], stevia (Stevia rebaudianc) Non-Jobel DNA: blockberry (Rubus sp. indet.) Comment: T1210 Lobel: robick [Appidritus [meoria], lemongrass [Cymbopogon citratus], stevia (Stevia rebaudianc) Non-Jobel DNA: blockberry (Rubus sp. indet.) Comment: T1210 Lobel: robick [Carcin papary] Non-Jobel DNA: paparys (Carcin papary) Non-Jobel DNA: paparys (Carcin papary) <	Label:	lemongrass, blackberry leaves, citric acid, rose hips, spearmint, natural flavors, orange peel, safflowers, hibiscus flowers, rose petals, orange essence, ginger, licorice, natural flavors				
Comment: R45 rbcl. 99.1% match Heteromorphic arborescens, mark 99.7% match Pimpinella soxifraga, both family Apiaceae. No listed ingredients in Apiaceae. Product: T1204 Information of the part (Hypericum perforatum) Interpret Comment: T1204 rbcl. 100% match to several species in genus Terpischore, family Polypodiaceae. No listed ingredients in Polypodiaceae. Product: T1207 rbcl. 100% match to several species in genus Terpischore, family Polypodiaceae. No listed ingredients in Polypodiaceae. Product: T1207 rbcl. 100% match to several species in genus Terpischore, family Orobanchaceae. Listed ingredient Euphrosia officinalis is also in family Orobanchaceae. There are no E. afficinalis sequences. In GenBank for direct comparison. However, the closest Euphrasis species with sequences in GenBank is relatively distant fram recovered sequence: T1207 rbcl. is 97.3% match to E. spectabilis AY849864, and T1207 matk is 93.7% match to E. spectabilis AY849603. T210 rbcl. 100% ond matk 99.5% match to Rubus discolor, family Rosaceae. No listed ingredients in Rosaceae. T1213 matk 100% match to Rubus discolor, family Rosaceae. No listed ingredients in Rosaceae. Product: T1213 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: T1225 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: T1225 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: T1225 rbcl. 99.8% match to Melsas afficinalis, family lamiaceae. No listed ingredients in Verbenaceae. Product: T1225 rbcl. 99.8% match to Melsas afficinalis, family Lamiaceae. No listed ingredients in Verbenaceae. Product: MGm17 rbcl. 99.6% identify Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MGm17 rbcl. 99.8% identify Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: WGm17 rbcl. 99.7% match to Melsas afficinalis, family Lamiaceae. R23 matk 98.9% mat	Non-label DNA:	parsley family (Apiaceae)				
Product: T1204 Label: St. John's wort (aerial part) (<i>Hypericum perforatum</i>) Non-label DNA: fern (Terpsichore sp. indet.) Comment: T1204 feat 100% match to several species in genus Terpischore, family Polypodiaceae. No listed ingredients in Polypodiaceae. Ibabl: eyobright herb (Euphrasia officinalis) Non-label DNA: red bartsia (Odonties verus) Comment: T1207 robanchaceae. There are no E. <i>Editicualis</i> sequences in Gen8ank for diverse sequences: T1207 rbcl is 97.3% match to E. spectabilis AY849864, and T1207 matk is 93.7% match to E. spectabilis AY849603. Product: T1210 T1210 tabaches junction in the sequences in Gen8ank is relatively distant from recovered sequence: T1207 rbcl is 97.3% match to E. spectabilis AY849603. Non-label DNA: backbarry (Rubus gp. indet.) Label: roaibas (Japaelahnus linearis), lemongrass (Cymbopogon citratus), stevia (Stevia rebaudiand) Non-label DNA: backbarry (Rubus gp. indet.) Label: roaibas (Japaelahnus linearis), lemongrass (Cymbopogan citratus), stevia (Stevia rebaudiand) Non-label DNA: backbarry (Rubus discolor, family Rosaceae. No listed ingredients in Rosaceae. Product: T1213 Uabel: yepsyce (Carica papoya) Comment: T1225 matk 100% match to Carica papaya, family Car	Comment:	R45 rbcL 99.1% match Heteromorpha arborescens, matK 99.7% match Pimpinella saxifraga, both family Apiaceae. No listed ingredients in Apiaceae.				
Label: S). John's wort (serial part) (Hypericum perforatum) Non-Label DNA: fem (Terpsichore sp. indet.) Comment: TT204 rbcl. 100% match to several species in genus Terpischore, family Polypodiaceae. No listed ingredients in Polypodiaceae. Non-Label DNA: red bartisi (Odonities verus) Comment: TT207 rbcl. 100% and matK 99.5% match to Odonities verus, family Orobanchaceae. Listed ingredient Euphrasia officinalis Non-Label DNA: infoil Orobanchaceae. There are no E. Enficinalis sequences in GenBank for direct comparison. However, the closes Product: TT210 TD21 robits and AV8864, and TT207 matK is 93.7% match to E. spectabilis AV849864. Non-Label DNA: blackberry (Rubus sp. indet.) Comment: TT210 TD21 robits (Aspalathus linearis), lemongrass (Cymbopogon citratus), stevia (Stevia rebaudiana) Non-Label DNA: blackberry (Rubus sp. indet.) Comment: TT213 TD21 robits (Graic papaya) Comment: TT213 TD23 reflex contrass of ficinalis, family Caricaceae. No listed ingredients in Caricaceae. Product: TT225 TD21 reflex contrass of ficinalis, family Lamiaceae. No listed ingredients in Lamiaceae. Product: TT225 </td <td>Product:</td> <td>TT204</td>	Product:	TT204				
Nonhabel DNA: Intern [<i>Irps/chore sp. inder.</i>] T207 Comment: T207 Irbo/La 100% match in several species in genus <i>Terpischore</i> , family Polypodiaceae. No listed ingredients in Polypodiaceae. T207 Irbo/Label: eyebright herb [<i>Euphrasia officinalis</i>] Comment: T207 Irbo/Label DNA: red bartsia (<i>Octonities vernus</i>) Comment: T207 Irbo/Label DNA: red bartsia (<i>Octonities vernus</i>) Comment: T207 Irbo/Label DNA: red bartsia (<i>Octonities vernus</i>) Irbo/Label DNA: red bartsia (<i>Octonities vernus</i>) Comment: T207 Irbo/Label DNA: red bartsia (<i>Octonities vernus</i>) Irbo/Label DNA: red bartsia (<i>Octonities vernus</i>) Irbo/Label DNA: red bartsia (<i>Octonities vernus</i>) Irbo/Label DNA: Iblo/Label: Irbo/Label DNA: Iblo/Label DNA: Irbo/Label DNA: Iblo/Label DNA: Iplo/Label DNA: Iplo		St. John's wort (aerial part) (Hypericum perforatum)				
 Traduct: Trading of the function of the service in period representing for the productive in the intermination of the productive in the intermediate intermediate	Commont:	tern (<i>Terpsichore sp. Inder.)</i> TT204 rhol 100% match to soveral species in gonus Terpischere, family Polypodiaceae. No listed ingredients in Polypodiaceae				
 Indexistical Contribution of Control Contro Control Control Control Control Control Control Control Contr	Product:	TI207 TI207				
 Hotobor Viel Teodomics Products Trador Action (1997) and matk 99.5% match to Odontites versus, family Orobanchaceae. Listed ingredient Euphrasia afficinalis is also in family Orobanchaceae. There are no E. afficinalis sequences in GenBank for direct comparison. However, the closest Euphrasia species with sequences in GenBank is relatively distant from recovered sequence: TT207 rbcL is 97.3% match to E. spectabilis AY849864, and TT207 matk is 93.7% match to E. spectabilis AY849603. Troduct: TT210 toobs (Aspalathus linearis), lemongrass (Cymbopogon citratus), stevia (Stevia rebaudiana) blackberry (Rubus sp. indet.) Comment: TT210 matk 99.9% match to Rubus discolor, family Rosaceae. No listed ingredients in Rosaceae. Product: TT213 Index Yeng (Rubus sp. indet.) Comment: TT213 matk 100% match to Rubus discolor, family Caricaceae. No listed ingredients in Caricaceae. Product: TT225 Comment: TT225 rbcL 99.8% match to <i>Aubus discolar</i>, family Lamiaceae. No listed ingredients in Lamiaceae. Product: H225 Index 100% match to <i>Kubus discolar</i>, family Lamiaceae. No listed ingredients in Lamiaceae. Product: MGm17 Intaoa (Lamiana sp. indet.) Comment: TT225 rbcL 99.8% match to <i>Melissa officinalis</i>, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRof 7 Label: orange, mango, cinnamon Non-label DNA: lantana (Lantana sp. indet.) Comment: MGm17 rbcL 90.9% identity Lantana camara, family Verbenaceae. No listed ingredient chicory (<i>Cichorium intybus</i>) is also in family Asteraceae, but in a different tribe, Cichoreae. Na dis sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. Product: R23 Label: ginger, chicory MRof Cel. 100% identity Stevia rebaudiana, family Asteraceae, Product: R23 also in f	Non-label DNA	eyebright herb (cuphrasia officinalis) red bartsia (Odantites vernus)				
Product: 11210 Label: rooibos (Aspalathus linearis), lemongrass (Cymbopogon citratus), stevia (Stevia rebaudiana) Non-label DNA: blackberry (Rubus sp. indet.) Comment: TT213 Label: yellowdock root (Rumex crispus) Non-label DNA: popaya (Carica papaya) Comment: TT213 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: TT225 Label: hipiricao (Hypericum perforatum) Non-label DNA: lemon balm (Melissa officinalis, family Lamiaceae.No listed ingredients in Lamiaceae. Product: TT225 rbcl 99.8% match to Melissa officinalis, family Lamiaceae. No listed ingredients in Lamiaceae. Product: MGm17 Label: orange, mango, cinnamon Non-label DNA: lantana (Lantana sp. indet.) Comment: MGm17 rbcl 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRoó ginger, chicory Non-label DNA: lantana (Lantana sp. indet.) Comment: MRoó rbcl. 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRoá sequence has lower identity to C. intybus (97.4%) [L	Comment:	TT207 rbcL 100% and matK 99.5% match to Odontites vernus, family Orobanchaceae. Listed ingredient Euphrasia officinalis is also in family Orobanchaceae. There are no E. officinalis sequences in GenBank for direct comparison. However, the closest Euphrasia species with sequences in GenBank is relatively distant from recovered sequence: TT207 rbcL is 97.3% match to E. spectabilis AY849864, and TT207 matK is 93.7% match to E. spectabilis AY849603.				
Label: Toolbos (Aspatiations lineorgis), lemongross (Cymbopogon citratus), stevia (Stevia rebaudiana) Non-label DNA: blackberry (Rubus sp. indel.) Comment: TT210 matk 99.9% match to Rubus discolar, family Rosaceae. No listed ingredients in Rosaceae. Product: TT213 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: TT225 Label: hipiricao (Hypericum perforatum) Non-label DNA: lemon balm (Melissa officinalis) Comment: TT225 rbcl 99.8% match to Melissa officinalis, family Lamiaceae. No listed ingredients in Lamiaceae. Product: MGm17 Label: orange, mango, cinnamon Non-label DNA: lantona (Lantana sp. indet.) Comment: MGT 7 rbcl 99.8% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRaó Label: orange, rango, cinnamon Non-label DNA: lentona (Lantana sp. indet.) Comment: MGAó rbcl. 100% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRaó Jabel: ginger, chicory Non-label DNA: sevia (Stevia rebaudiana) Label: prebudiana sequence.	Product:	TT210				
Non-Jobel DNA: Didx.sergin (Nubber 99.9% match to Rubus discolor, family Rosaceae. No listed ingredients in Rosaceae. Product: T12 13 Label: yellowdock root (Rumex crispus) Non-Jabel DNA: papaya (Carica papaya) Comment: T12 13 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: T1225 Label: hipiricao (Hypericum perforatum) Non-Jabel DNA: lemon balm (Melissa officinalis, family Lamiaceae.No listed ingredients in Lamiaceae. Product: MGm 17 Label: orange, mango, cinnamon Non-Jabel DNA: lemtona (Lantana sp. indet.) Comment: MCm 17 cbcl 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRa6 Label: ginger, chicory Non-Jabel DNA: stevia (Stevia rebaudiana) Comment: MRa6 track to to 00% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) [L13652] than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa aolong tea	Label: Non label DNA:	rooibos (Aspalathus linearis), lemongrass (Cymbopogon citratus), stevia (Stevia rebaudiana) blackberg: (Bubus sp. indet l				
IT213 IT213 Label: yellowdock root (Rumex crispus) Non-label DNA: papaya (Carica papaya) Comment: TT213 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: TT225 Label: hipiricao (Hypericum perforatum) Non-label DNA: lemon balm (Melissa officinalis, family Lamiaceae. No listed ingredients in Lamiaceae. Comment: TT225 rbcl 99.8% match to Melissa officinalis, family Lamiaceae. No listed ingredients in Lamiaceae. Product: MGm17 Label: orange, mango, cinnamon Non-label DNA: lantona (Lantana sp. indet.) Comment: MGG 72.4% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MR6 Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MK6 /bcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is a also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: formosa oolong tea Non-label DNA: <	Comment:	TT210 <i>matk</i> 99.9% match to <i>Rubus discolor</i> , family Rosaceae. No listed ingredients in Rosaceae.				
Non-label DNA: papaya (Carica papaya) Non-label DNA: papaya (Larica papaya) Roduct: TT213 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: TT225 Label: hipiricao (Hypericum perforatum) Non-label DNA: lemon balm (Melissa officinalis) Comment: TT225 rbcl 99.8% match to Melissa officinalis, family Lamiaceae. No listed ingredients in Lamiaceae. Product: MGm17 Label: orange, mango, cinnamon Non-label DNA: lantana (Lantana sp. indet.) Comment: MRa6 / L99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRa6 Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRa6 fact 100% identity Stevia rebaudiana, family Asteraceae. Tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: head II (Prunella vulgaris), chamomile (Matricaria recutita) Comment:		11213 vellowdock root (Rumex crispus)				
Comment: TT213 matk 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae. Product: TT225 Label: hipiricao (Hypericum perforatum) Non-label DNA: lemon balm (Melissa officinalis) Comment: TT225 rbcl 99.8% match to Melissa officinalis, family Lamiaceae. No listed ingredients in Lamiaceae. Product: MGm17 Label: orange, mango, cinnamon Non-label DNA: lantana (Lantana sp. indet.) Comment: MRa6 TrbcL 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRa6 Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) [L13652] than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: head all (Prunella vulgaris, chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matrica	Non-label DNA:	papaya (Carica papaya)				
Product: T1225 Label: hipiricao (Hypericum perforatum) Non-label DNA: lemon balm (Melissa officinalis) Comment: T1225 rbcl 99.8% match to Melissa officinalis, family Lamiaceae.No listed ingredients in Lamiaceae. Product: MGm17 Label: orange, mango, cinnamon Non-label DNA: lantana (Lantana sp. indet.) Comment: MGm17 rbcl 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRaó Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRaó rbcl 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRaó sequence has lower identity to C. intybus (97.4%) [L13652] than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa colong tea Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 tbcl 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) Comment: R33 R43 sichuan tea <td>Comment:</td> <td>TT213 matK 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae.</td>	Comment:	TT213 matK 100% match to Carica papaya, family Caricaceae. No listed ingredients in Caricaceae.				
Label: hipricac (Hypericum pertoratum) Non-label DNA: lemon balm (Melissa officinalis) Comment: TT225 rbcL 99.8% match to Melissa officinalis, family Lamiaceae.No listed ingredients in Lamiaceae. Product: MGm17 Label: orange, mango, cinnamon Non-label DNA: lantana (Lantana sp. indet.) Comment: MGm17 rbcL 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRaó Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRaó focl. 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteracee, but in a different tribe, Cichoreae. MRaó sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 tob.1 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) Comment: R33 Product: R33 Label: Sichuan tea <td>Product:</td> <td>Π225</td>	Product:	Π225				
 Intendable DNA: International (Meinssa anticinalis) Comment: TT225 theL 99.8% match to Melissa officinalis, family Lamiaceae. No listed ingredients in Lamiaceae. Product: MGm17 Label: orange, mango, cinnamon Non-label DNA: Iantana (Lantana sp. indet.) Comment: MGm17 theL 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRa6 Label: ginger, chicory Non-label DNA: steria (<i>Steria rebaudiana</i>) Comment: MRa6 theL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (<i>Cichorium intybus</i>) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to <i>C. intybus</i> (97.4%) (L13652) than to <i>S. rebaudiana</i> sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: heal all (<i>Prunella vulgaris</i>), chamomile (<i>Matricaria recutita</i>) Comment: R23 tol. 100% match to <i>Prunella vulgaris</i> family Lamiaceae. R23 <i>matK</i> 98.9% match to <i>Achillea millefolium</i>, family Asteraceae, ikely represents chamomile (<i>Matricaria recutita</i>) Comment: R33 Comment: R33 Comment: R33 <i>matK</i> 99.7% match <i>Pimpinella saxifraga</i>, family Apiaceae. No listed ingredients in Apiaceae. Product: R34 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 <i>matK</i> 99.7% match <i>Pimpinella saxifraga</i>, family Apiaceae. No listed ingredients in Apiaceae. 	Label:	hipiricao (Hypericum pertoratum)				
Orment: MCm17 Label: orange, mango, cinnamon Non-label DNA: Iantana (Lantana sp. indet.) Comment: MGm17 rbcL 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRa6 Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 rbcl 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) Comment: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea Non-label	Comment:	TT225 rbcl 99.8% match to Melissa officinalis, family lamiaceae No listed inaredients in lamiaceae				
Label: orange, mango, cinnamon Non-label DNA: Iantana (Lantana sp. indet.) Comment: MGm17 rbcL 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRa6 Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: Heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) Comment: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae)	Product:	MGm17				
Non-label DNA: Iantana (Lantana sp. indet.) Comment: MGm17 rbcL 99.6% identity Lantana camara, family Verbenaceae. No listed ingredients in Verbenaceae. Product: MRa6 Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) [L13652] than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) Comment: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea Non-label DNA: gunpowder tea Non-label DNA: gunpowder tea <	Label:	orange, mango, cinnamon				
Comment: MGm17 rbcL 99.5% identity Lantana camara, tamily Verbenaceae. No listed ingredients in Verbenaceae. Product: MRa6 Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) [L13652] than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa colong tea Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) Comment: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 ubel: gunpowder tea Non-label DNA: gunpowder tea Non-label DNA: gunpowder tea Non-label DNA: gunpowder tea Non-label DNA	Non-label DNA:	lantana (Lantana sp. indet.)				
Index Middo Label: ginger, chicory Non-label DNA: stevia (Stevia rebaudiana) Comment: MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Lamiaceae or Asteraceae. Product: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R36 Label: gingre family (Apiaceae) Comment: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea </td <td>Comment:</td> <td>MGm1/ rbcL 99.6% identity Lantana camara, tamily Verbenaceae. No listed ingredients in Verbenaceae.</td>	Comment:	MGm1/ rbcL 99.6% identity Lantana camara, tamily Verbenaceae. No listed ingredients in Verbenaceae.				
Non-label DNA:gingle, i, rebaudiana,Stevia (Stevia rebaudiana)Comment:MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae.Product:R23 Label:Non-label DNA:heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment:Comment:R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Lamiaceae or Asteraceae.Product:R33 Label:Don-label DNA:parsley family (Apiaceae) Comment:R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.Product:R36 Label:Unable:gunpowder tea Non-label DNA:Non-label DNA:parsley family (Apiaceae) Comment:R36 Label:gunpowder tea Non-label DNA:Non-label DNA:parsley family (Apiaceae) R36Comment:R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.R36 Label:gunpowder tea Non-label DNA:Non-label DNA:parsley family (Apiaceae) Comment:R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.		Mikao ginger chicory				
Comment: MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae. Product: R23 Label: Formosa oolong tea Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae. Product: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 R36 gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. N	Non-label DNA:	stevia (Stevia rebaudiana)				
also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652) than to S. rebaudiana sequence. No other ingredients in Asteraceae.Product:R23Label:Formosa oolong tea Non-label DNA:Non-label DNA:heal all (Prunella vulgaris), chamomile (Matricaria recutita) R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Lamiaceae or Asteraceae.Product:R33 Label:Non-label DNA:parsley family (Apiaceae) Comment:R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.Product:R36 Label:Unpowder tea Non-label DNA:Product:R36 Label:R36 Label:gunpowder tea Non-label DNA:Non-label DNA:gunpowder tea Non-label DNA:R36 Label:gunpowder tea Non-label DNA:Non-label DNA:parsley family (Apiaceae) Comment:R36 Label:gunpowder tea Non-label DNA:Non-label DNA:parsley family (Apiaceae) R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae.Non-label DNA:parsley family (Apiaceae) R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae.Non-label DNA:parsley family (Apiaceae) R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae.Non-label DNA:parsley family (Apiaceae) R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae.Non-label DNA:parsley family (Apiaceae) R36 matK 99.7%	Comment:	MRa6 rbcL 100% identity Stevia rebaudiana, family Asteraceae, tribe Eupatorieae. Listed ingredient chicory (Cichorium intybus) is				
than to S. rebaudiana sequence. No other ingredients in Asteraceae.Product:R23Label:Formosa oolong teaNon-label DNA:heal all (Prunella vulgaris), chamomile (Matricaria recutita)Comment:R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Lamiaceae or Asteraceae.Product:R33Label:Sichuan teaNon-label DNA:parsley family (Apiaceae)Comment:R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.Product:R36Label:gunpowder teaNon-label DNA:parsley family (Apiaceae)Comment:R36Label:gunpowder teaNon-label DNA:parsley family (Apiaceae)Comment:R36Label:gunpowder teaNon-label DNA:parsley family (Apiaceae)Comment:R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.Non-label DNA:parsley family (Apiaceae)Comment:R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.		also in family Asteraceae, but in a different tribe, Cichoreae. MRa6 sequence has lower identity to C. intybus (97.4%) (L13652)				
Product: K23 Label: Formosa oolong tea Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Lamiaceae or Asteraceae. Product: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Devel at	than to <i>S. rebaudiana</i> sequence. No other ingredients in Asteraceae.				
Non-label DNA: heal all (Prunella vulgaris), chamomile (Matricaria recutita) Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Lamiaceae or Asteraceae. Product: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Non-label DNA: parsley family (Apiaceae) Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.		KZS Formosa oolona tea				
Comment: R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Lamiaceae or Asteraceae. Product: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Non-label DNA:	heal all (Prunella vulgaris), chamomile (Matricaria recutita)				
Product: R33 Label: Sichuan tea Non-label DNA: parsley family (Apiaceae) Comment: R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae. Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Comment:	R23 rbcL 100% match to Prunella vulgaris, family Lamiaceae. R23 matK 98.9% match to Achillea millefolium, family Asteraceae, likely represents chamomile (Matricaria recutita) (see Comment under G2). No listed ingredients in Lamiaceae or Asteraceae.				
Label:Sichuan teaNon-label DNA:parsley family (Apiaceae)Comment:R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.Product:R36Label:gunpowder teaNon-label DNA:parsley family (Apiaceae)Comment:R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Product:	R33				
Non-label DNA:parsley tamily (Apiaceae)Comment:R33 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.Product:R36Label:gunpowder teaNon-label DNA:parsley family (Apiaceae)Comment:R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Label:	Sichuan tea				
Comment: K33 matk 99.7% match Pimpinella saxifraga, tamily Apiaceae. No listed ingredients in Apiaceae. Product: R36 Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 matk 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Non-label DNA:	parsley tamily (Apiaceae)				
Label: gunpowder tea Non-label DNA: parsley family (Apiaceae) Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Comment: Product:	K33 matk YY./% match Pimpinella saxitraga, tamily Apiaceae. No listed ingredients in Apiaceae.				
Non-label DNA: parsley family (Apiaceae) Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Label:	aunpowder tea				
Comment: R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.	Non-label DNA:	parsley family (Apiaceae)				
	Comment:	R36 matK 99.7% match Pimpinella saxifraga, family Apiaceae. No listed ingredients in Apiaceae.				

var. assamica vs. C. sinensis var. sinensis (71% vs. 12%, p = 0.0002) (additional details in Supplementary Table S3 online).

Discussion

Reliable DNA identification of species requires recovery of a barcode sequence from the sample, representation of relevant species in the reference database, and sufficient nucleotide sequence variability to distinguish among closely-related species²⁶. Regarding the first requirement, we recovered *rbcL* or *matK* barcodes from 90% of commercial tea products using a single set of primers for each region. Success was less frequent with herbal as compared to CS teas (84% vs 96%), which may reflect primer mismatch, *Taq* inhibition, or DNA

degradation in some of the diverse plant materials in herbal teas. In terms of markers, *rbcL* was recovered from a broader taxonomic range of plants than *matK* (42 species in 24 families vs. 25 species in 16 families; Figs. 1,2). These results are consistent with general observation that *rbcL* is more easily amplified from wide range of species than is $matK^{19,20}$.

The second condition for DNA identification of species is representation of relevant taxa in the reference database, in our case GenBank. As in most practical applications of barcoding, our specimens were morphologically unrecognizable, thus representation cannot be assessed directly. About one-third of the plant species listed on labels lacked GenBank records for *rbcL*, *matK*, or both at



the time of the study. A more precise indicator of species representation is whether the recovered sequences are identical to any in the database. 62% of our barcode haplotypes did not have an identical match in GenBank (Figs. 1,2). This indicates that many plant species found in tea products are either not represented, have undocumented intraspecific variation, or that a sequencing error has occurred.

The third requirement for identifying species by barcode is biological: there must be sequence differences that discriminate among closely-related species. We can determine how well this condition is met for our specimens by comparing the best match and the congeneric nearest neighbor for each haplotype. For *rbcL*, these differed by only 1 site on average, and for *matK* these differed by only 2 sites on average (Figs. 1,2; see also Supplementary Tables S1,S2 online). Our results are consistent with the estimated 70%-85% species discrimination using *rbcL* + *matK* barcodes, and highlight the relatively small number of positions that distinguish many closely-related plant species^{19,23,24}. Differences between congeneric species in this study are similar to those reported for intraspecific variation and are also the same magnitude as sequencing error. Thus a barcode that differs from its closest reference database sequence at just one or a few sites plausibly represents an unrecorded variant for that species, a closely-related species not in the reference database, or sequencing error.

Our results highlight a need for improved algorithms for assigning taxonomic names to plant barcode sequences, particularly if barcoding is to be applied by non-specialists, which is one of the goals of the effort^{1,12,25}. Algorithms that place search results in the context of plant taxonomy and current database representation of related plants will be helpful. Character-based approaches may assist in distinguishing closely-related species, particularly if supported by expert annotation that flags diagnostic nucleotide positions^{27,28}. In addition, although employing two markers adds precision to plant barcode identifications, it also generates a need for algorithms that integrate database search results. In our data, most extractions that yielded both markers gave discordant results, that is, the *rbcL* and *matK* barcodes matched different species in GenBank, largely reflecting differences in representation of species or intraspecific variants for the two markers.

A large fraction (35%) of herbal products yielded one or more barcodes that pointed to non-label ingredients. Possible explanations include database errors (e.g. sequences with incorrect species names), limitations of search algorithm (e.g. relevant sequences not recognized by BLAST), laboratory error (e.g. PCR contamination, sample mix-up), or presence of unlisted ingredients. The disproportionate number of discordant sequences recovered from herbal specimens and the finding of species not listed on other products and not under study in the laboratory points to unnamed constituents. This could reflect inadvertent introduction, such as from harvested plant material mixed with unrecognized species, residual products in processing machinery, or as part of unspecified flavorings listed on some products. The relative amount of such potential material in our samples is unknown and is beyond the scope of this study. The finding of unlisted chamomile (M. recutita) or tea plant (C. sinensis) in multiple products suggests the possibility of addition or substitution to improve taste, appearance, or for economic reasons²⁹.

To our knowledge, the polymorphism at *rbcL* position 68 is the first described plastid marker that differs among *C. sinensis* varieties, regions of cultivation, and tea processing types⁵⁻¹¹. Our results are consistent with marketplace trends—India and Sri Lanka, largely devoted to cultivation of *C. sinensis* var. *assamica*, are the dominant global exporters of black tea, whereas China, largely cultivating *C. sinensis* var. *sinensis*, has become the dominant exporter of green tea, with 75% of world market³⁰. Our findings may help inform future research on the geographic origin and diversity of wild and cultivated CS resources^{5,31}.

In summary, plant DNA barcodes can be recovered from most commercial tea products using a standard protocol. At the same time, interpreting DNA barcode identifications in relation to product labels is challenging. New algorithms that place search results in the context of standard plant names and character-based keys for distinguishing closely-related species are needed. With appropriate software to guide non-experts, DNA barcoding can offer an effective method to help provide more accurate ingredient labels to consumers, thereby improving safety of food and botanicals³². This is particularly pertinent in an increasingly global economy where longer and more complex market chains distance suppliers from the source of products and where regulatory agencies are becoming more stringent with food and botanical labeling^{33,34}.

Methods

Specimen collection. CS and herbal tea products from New York City stores, school dining halls, and homes of investigators were collected during October 2009-February 2010. 146 products were obtained from 25 locations, representing 33 manufacturers, 17 countries, and 82 plant common names. As this study was designed to enable comparison between CS and herbal teas and not among individual products or manufacturers, products were assigned an arbitrary alphanumeric code. 73 were *C. sinensis*, and 73 were herbal products prepared from other plant species. Five herbal products contained *C. sinensis* together with other plants. 44 herbal teas (60.3%) listed a single ingredient; the remainder named 2–10 different plants. When not specified on the label, scientific and common name equivalents were determined from the reference used by the U.S. Food and Drug Administration³⁵.

Reference samples. *C. sinensis* var. *assamica* specimens (n = 17) were collected in Yunnan, China by SA during 2007–2009. *C. sinensis* var. *sinensis* specimens (n = 24) collected in China (7), Taiwan (7), Japan (7), and Argentina (3) were obtained from the Kunming Institute of Botany, Kunming, China. Reference sample *rbcL* sequences and additional collection information were deposited in GenBank under accession codes JN009623-JN009663. GenBank accessions used for comparison of *C. sinensis rbcL* haplotypes included *C. albogigas* (AF380033), *C. granthamiana* (AF380034), *C. japonica* (AF380035), *C. oleifera* (GQ436637), *C. sasanqua* (AF380036), *C. sinensis* (AF380037), and *Tutcheria hirta* (AF380067).

DNA extraction and sequencing. DNA was isolated from 5–15 mg dried tissue using a DNeasy96 Plant kit (Qiagen). The manufacturer's protocol was modified as follows: tissue was disrupted and then incubated for 12–18 h with gentle mixing at 42°C in 600 μ L of the supplied AP1 buffer with 600 μ g of protease K added (630 μ L total volume). Polysaccharides were precipitated at 4°C with 200 μ L AP2. The remaining steps followed the manufacturer's protocol. For the 86% of specimens that appeared morphologically homogenous, a single extraction was performed. The remaining samples were divided into groups of morphologically homogeneous material (average 3, range 2–8), and separate extractions were performed with the aim of recovering individual components.

Individual amplifications of *matK* and *rbcL* took place in a 15 μ L volume containing: 1.5 μ L buffer [200 mM Tris pH 8.8, 100 mM KCl, 100 mM (NH₄)₂SO₄, 20 mM MgSO₄•7H₃O, 1% (v/v) Triton X-100, 50% (w/v) sucrose, 0.25% (w/v) cresol red], 0.2 mM dNTPs, 0.025 μ g/ μ L BSA, 0.5 (*rbcL*) or 1 (*matK*) μ M of each primer, 1 unit of *Taq*, and 0.5 μ L genomic DNA. For amplification and sequencing of *matK*, primers 3F (5'-CGT-ACA-GTA-CTT-TTG-TGT-TTA-CGA-G-3') and 1R (5'-ACC-CAG-TCC-ATC-TGG-AAA-TCT-TGG-TTC-3') ²⁷ were used with the following cycling conditions: 95°C 2.5 min; 10 cycles: 95°C 30 s, 56°C 30 s, 72°C 30 s; 25 cycles: 88°C 30 s, 56°C 30 s, 72°C 30 s; 72°C 10 min. For *rbcL* amplification and sequencing, primers F1 (5'-ATG-TCA-CCA-CAA-ACA-GAG-ACT-AAA-GC-3') ²² and R634 (5'-GAA-ACG-GTC-TCT-CCA-ACG-CAT-3')²⁰ were used with the following cycling conditions: 95°C 2.5 min; 35 cycles: 95°C 30 s, 58°C 30 s, 72°C 30 s; 72°C 10 min.

PCR products were treated with ExoSAP-IT and bi-directionally sequenced with BigDye 3.1 chemistry on an ABI 3730 sequencer (High–Throughput Genomics Unit, University of Washington).

Portable laboratory. A subset of specimens (10) were analyzed in a portable laboratory. Equipment included a thermal cycler (Techne), microcentrifuge (Eppendorf minispin), vortex mixer, heating block, pipettemen, and E-gel apparatus (Invitrogen), purchased used or reconditioned except for E-gel unit. DNA was isolated with DNeasy Plant Mini Kit (Qiagen) following manufacturer's instructions. PCR was performed using *rbcL* primers as described above except that 25 µl reaction volume, 0.5 units TaKaRa Ex *Taq*, and buffer supplied by manufacturer were used. DNA and PCR yields were assessed on an E-gel EX 1% with a blue-light excitable nucleic acid stain, products were cleaned with QIA quick PCR purification kit (Qiagen), and unidirectional sequencing was performed at a commercial facility (Macrogen).

Sequence files and data analysis. Trace files were assembled in MacVector 11.0, and sequences with greater than 2% ambiguous bases were discarded, using QV of 40 for



bi-directional reads and 20 for single reads. Sequences were aligned using ClustalW (*rbcL*) or MUSCLE v3.8.31 (*matK*). Sequence files are deposited in GenBank under accession codes HQ699082-HQ699129 (*rbcL*) and HQ699130-HQ699169 (*matK*). Fisher's exact test, two-tailed, was used for statistical comparisons.

Database searches. GenBank database was searched using megaBLAST during August-October 2010, with default parameters adjusted to retrieve 5000 sequences. To optimize correct identifications, the closest match for each *rbcL* and *matK* haplotype was defined as the target with highest percentage identity using an arbitrary cutoff of 90% or greater overlap with the query sequence. In most cases this corresponded to the sequence with the highest BLAST score. In other cases, the closest match was a shorter target with a higher percent identity. Ambiguous bases in query or target sequences were considered as matching. For queries that produced multiple identical matches, the target with a species name closest to a label ingredient was chosen when possible. A similar procedure was followed for BOLD searches, with the exception that the number of alignment results was 100, which is the maximum allowed. For consistency in reporting, the species of sequences deposited in GenBank and BOLD were used unaltered even though some may be in error or reflect outdated taxonomy.

- Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. Biological identifications through DNA barcodes. *Proc. Biol. Sci.*, 270, 313–321 (2003).
- Chang, H. T. A taxonomy of the genus Camellia. Acta Sci. National Uni.v Sunyatseni, Monog. Series 1, 1–180 (1981).
- 3. Chang, H. T. & Bartholomew, A. Camellias. London: Batsford, 211 p. (1984).
- Ming, T. & Zhang, W. The evolution and distribution of genus *Camellia. Acta Botanica Yunnanica* 18, 1–13 (1996).
 Beleveryungen T. Dive, D. K. Kumer, P. P. Mureleydheren, N. & Sheeper, A. J.
- Balasaravanan, T., Pius, P. K., Kumar, R. R., Muraleedharan, N. & Shasany, A. K. Genetic diversity among south Indian tea germplasm (*Camellia sinensis, C. assamica, and C. assamica* spp. *Lasiocalyx*) using AFLP markers. *Plant Sci.* 165, 365–372 (2003).
- Ni, S., Yao, M., Chen, L., Zhao, L. & Wang, X. Germplasm and breeding research of tea plant based on DNA marker approaches. *Front. Agric. China* 2, 200–207 (2008).
- Katoh, Y., Katoh, M., Takeda, Y. & Omori, M. (2003) Genetic diversity within cultivated teas based on nucleotide sequence comparison of ribosomal RNA maturase in plastid DNA. *Euphytica* 134, 287–295 (2003).
- Wachira, F. R., Powell, W. & Waugh, R. An assessment of genetic diversity among *Camellia sinensis* L. (cultivated tea) and its wild relatives based on randomly amplified polymorphic DNA and organelle-specific STS. *Heredity* 78, 603–611 (1997).
- Chen, J., Wang, P., Xia, Y., Xu, M. & Pei, S. Genetic diversity and differentiation of *Camellia sinensis* L. (cultivated tea) and its wild relatives in Yunnan province of China, revealed by morphology, biochemistry, and alloenzyme studies. *Genet. Resources Crop. Eval.* 52, 41–52 (2005).
- Singh, D. & Ahuja, P. S. 5S rDNA gene diversity in tea (*Camellia sinensis* (L.) O. Kuntze) and its use for variety identification. *Genome* 49, 91–96 (2006).
- Lai, J.-A., Yang, W.-C. & Hsia, J.-Y. An assessment of genetic relationships in cultivated tea clones and native wild tea in Taiwan using RAPD and ISSR markers. *Bot. Bull. Acad. Sin.* 42, 93–100 (2001).
- Li, H. L. The domestication of plants in China: ecogeographical considerations. In: Keightley, D. N., editor. *The Origins of Chinese Civilization*. Berkeley: University of California Press, pp. 21–64 (1982).
- Ceresa, M. (1996) Diffusion of tea-drinking habit in pre-Tang and early Tang period. Asiatica Venetiana 1, 19–25 (1996).
- 14. Magoma, G. N., Wachira, F. N., Obanda, M., Imbuga, M. & Agong, S. G. The use of catechins as biochemical markers in diversity studies of tea (*Camellia sinensis*). *Genet. Resources Crops. Evol.* 47, 107–114 (2000).
- Ahmed, S., *et al.* Pu-erh tea tasting in Yunnan, China: correlation of drinkers' perceptions to phytochemistry. *J. Ethnopharmacol.* 132, 176–185 (2010).
- Centers for Disease Control. Anticholinergic poisoning associated with an herbal tea–New York City, 1994. Morb. Mortal. Wkly. Rep. 44, 193–195 (1995).
- Kumana, C. R. *et al.* Herbal tea induced veno-occlusive disease: quantification of toxic alkaloid exposure in adults. *Gut* 26, 101–104 (1985).
- Toxicology and Clinical Pharmacology of Herbal Products, editor Cupp, M. J. Totowa: Humana Press, 325 p. (2000).

- CBOL Plant Working Group. A DNA barcode for land plants. Proc. Natl. Acad. Sci. U. S. A. 106, 12794–12797 (2009).
- 20. Fazekas, A. J. *et al.* Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* 7, e2802 (2008).
- 21. Seberg, O. & Petersen, G. How many loci does it take to barcode a crocus? *PLoS ONE* **4**, e4598 (2009).
- 22. Kress, W. J. & Erickson, D. L. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE* 2, e508 (2007).
- Fazekas, A. J. *et al.* Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Mol. Ecol. Res.* 9(S1), 130–139 (2009).
- Lahaye, R. et al. DNA barcoding the floras of biodiversity hotspots. Proc. Natl. Acad. Sci. U. S. A. 105, 2923–2938 (2008).
- Chase, M. W. et al. Land plants and DNA barcodes: short-term and long-term goals. Phil. Trans. R. Soc. B 360, 1889–1895 (2005).
- Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. Identification of birds through DNA barcodes. *PLoS Biol.* 2, e312 (2004).
- Little, D. P. & Stevenson, D. W. A comparison of algorithms for the identification of specimens using DNA barcodes: examples from gymnosperms. *Cladistics* 23, 1–21 (2007).
- 28. Hollingsworth, P. M., Graham, S. W. & Little, D. P. Choosing and using a plant DNA barcode. *PLoS ONE* **6**, e19254.
- Cole, M. R. & Fetrow, C. W. Adulteration of dietary supplements. Amer. J. Health-System Pharm. 60, 1576–1580 (2003)
- 30. Food and Agricultural Organization of the United Nations. Medium-term prospects for agricultural commodities. Projections to the year 2010. Rome, 2010. Accessed online at http://www.fao.org/docrep/006/y5143e/y5143e0z.htm
- 31. Lou, S. K., Wong, K. L., Li, M., But, P. P., Tsui, S. K. & Shaw, P. C. An integrated web medicinal materials DNA database: MMDBD (Medicinal Materials DNA Barcode Database). BMC Genomics 11, 402 (2010).
- Prince, L. M. & Parks, C. R. Phylogenetic relationships of Teaceae inferred from chloroplast DNA sequence data. Amer. J. Botany 88, 2309–2320 (2001).
- Ebeler, S. E., Takeoda, G. R. & Winterhalter, P., editors. Authentication of Food and Wine. United States of America: Oxford University Press, 364 p. (2007).
- 34. Chen, S.*et al.* Validation of the ITS region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE **5**, e6813 (2010).
- McGuffin, M., Kartesz, J. T., Leung, A. Y. & Tucker, A. O., editors. *Herbs of Commerce*, 2nd Edition. United States of America: American Herbal Products Association, 421 p. (2000).

Acknowledgments

We thank Jesse Ausubel (JA) for encouraging investigation, Trinity School for permission to work with their students, and JA and Paul Waggoner for helpful comments on an earlier version of the manuscript. DPL was supported by a grant from The Alfred P. Sloan Foundation (2010-06-02). SA's field work in Yunnan was supported by NSF grants DDEP OISE-0749961 and NSF EAPSI OISE-0714431. SA thanks the Kunming Institute of Botany, Chinese Academy of Sciences for permission to collect CS samples.

Author contributions

MYS and DPL designed the study; SA, CCG, RK, and GY contributed samples; MYS, CCG, RK, GY, and DPL performed experiments and analyzed data; and MYS and DPL wrote the manuscript with assistance from all authors.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/ scientificreports

Competing financial interests: The authors declare no competing financial interests.

License: This work is licensed under a Creative Commons

Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/

How to cite this article: Stoeckle, M.Y. *et al.* Commercial Teas Highlight Plant DNA Barcode Identification Successes and Obstacles. *Sci. Rep.* **1**, 42; DOI:10.1038/srep00042 (2011).