Rep. INT-GTR-372. U.S. Department of Agriculture, Forest Service, Intermountain Research Station, Ogden, UT.
Roundy, B.A. 1999. Lessons from historical rangeland revegetation for todays restoration. p. 33–38. In L.K. Holzworth and R.W. Ray (comps.) Revegetation with native species: Proceedings. RMRS-P-8. USDA, Forest Service, Rocky Mountain Research Station, Ogden, UT.
Soreng, R.J. 1990. Chloroplast-DNA phylogenetics and biogeography in a reticulating group: study in *Poa* (Poaceae). Am. J. Bot. 77: 1383–1400.
Soreng, R.J. 1991. Systematics of the Epiles group of *Poa* (Poaceae). Syst. Bot. 16:507–583.
Stebbins, G.L., Jr., and R.M. Love. 1941. A cytological study of California forage grasses. Am. J. Bot. Syst. Bot. 11:559–566.
Stark, R.H., A.L. Hafenrichter, and W.A. Moss. 1950. Adaptation of grasses for soil and water conservation at high altitudes. Agron. J. 42:124–127.
Usberti, J.A., and S.K. Jain. 1978. Variation in *Panicum maximum*; a comparison of sexual and asexual populations. Bot. Gaz (Chicago) 139:112–116.
Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hornes, A. Frijters, L. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407–4414.

Analysis of Genetic Diversity in *Theobroma cacao* with Emphasis on Witches' Broom Disease Resistance

J. M. Marita,* J. Nienhuis, J. L. Pires, and W. M. Aitken

ABSTRACT

To facilitate the identification of *Theobroma cacao* L. that possess desirable traits to meet changing production and market conditions, there is a need to understand the genetic relationships among *T. cacao* germplasm. In addition, new cultivars are needed to provide more broadly based resistance to devastating diseases such as witches' broom disease [*Crinipellis perniciosa* (Stahel) Singer]. A subset of 270 *T. cacao* accessions based on (i) witches' broom disease resistance data, (ii) genetic characterization experiments, and (iii) a random sampling of recently acquired accessions was selected from the extensive germplasm collection at the Centro de Pesquisa do Cacau (CEPEC; Itabuna, Bahia, Brazil) in collaboration with Fazenda Almirante, a division of M&M Mars Incorporated in Itajuipe, Bahia, Brazil. Estimates of genetic distance based on random amplified polymorphic DNA (RAPD) markers were used to evaluate the genetic diversity among the selected accessions. Incorporation of recently acquired accessions (Accessions 181–270) did not increase the breadth of the distribution of genetic diversity already present within the "original" group sampled (Accessions 1–180) suggesting new accessions collected from already sampled geographic regions do not increase the existing genetic diversity in the germplasm collection. In addition, differences in RAPD marker frequencies were associated with accessions that had a high threshold of tolerance to witches' broom disease. Most accessions exhibiting tolerance to witches' broom disease were from the Upper Amazon region, with the exception of SGU 26, a hybrid from Guatemala. This suggests that the Upper Amazon is not the only region to have genes for resistance to witches' broom disease and stresses the need for further collection and examination of germplasm from other regions.

Theobroma cacao, the source of chocolate, is the predominant cultivated species in the genus and is grown throughout the tropics of Central and South America, Asia, and Africa. The three main groups within *T. cacao*, which are based on geographic location and morphological characteristics, are Criollo, Forastero, and Trinitario (Cheesman, 1944). The Criollo group represents the first domesticated cacaos, and is composed of trees with thick, white or rosy beans that yield the most flavored and finest chocolate. Criollo was originally cultivated in Central America but because of its susceptibility to diseases is infrequently cultivated today (Soria, 1970). Because of higher yields and greater disease resistance, Forastero is currently the most widely cultivated group accounting for over 80% of the world's production of cacao. The Forastero group is further subdivided into Lower and Upper Amazon Forastero with the Upper Amazon group considered more genetically diverse and exhibiting superior agronomic performance (Bartley, 1969; Lockwood, 1976). The Trinitario group represents hybrid forms from crosses between the Criollo and Forastero groups (Wood and Lass, 1985).

The first attempts to collect and conserve cacao germplasm began in the early 20th century (Kennedy et al., 1987). These early collections were based on phenotypic selection within populations of Forastero, Criollo, and Trinitario and were later referred to as "land variety" collections (Kennedy et al., 1987). In later years, additional expeditions were undertaken for collection of phenotypically superior germplasm that explored under-represented regions of the upper Amazon (Baker, 1953). Currently, the International Plant Genetic Resources Institute (IPGRI, Rome, Italy) considers geographic areas that are under-represented in existing collections and areas of variability that are under the greatest threat as the two most important criteria in determining priorities for collection (Williams, 1984). With these objectives in mind, present day collecting expeditions have expanded to regions of Brazil, Colombia, Mexico, and Venezuela.

Witches' broom disease, one of the most serious threats to cacao production, was first described by Went (1904), and the fungus (*C. perniciosa*) responsible for

J.M. Marita, Dep. of Forestry, 1925 Linden Drive, USDFRC, Madison, WI 53706; J. Nienhuis, Dep. of Horticulture, 1575 Linden Drive, Univ. of Wisconsin, Madison, WI 53706; J.L. Pires, Centro de Pesquisa do Cacau-Cepec, Km 22 Rod, CEP45600-000, Itabuna, Bahia, Brazil; W.M. Aitken, Almirante Cacau Agrícola Comercio E Exp Ltda., Caixa Postal 55, CEP 45630-000, Itajuipe, Bahia, Brazil. Received 12 April 2000. *Corresponding author (jmarita@facstaff.wisc.edu).

Published in Crop Sci. 41:1305–1316 (2001).

Abbreviations: CB, cushion brooms; CEPEC, Centro de Pesquisa do Cacau; IPGRI, International Plant Genetic Resources Institute; MDS, multi-dimensional scaling; RAPD, random amplified polymorphic DNA; VB, vegetative brooms.

Table 1. Passport information for the 270 accessions from the Centro de Pesquisa do Cacau germplasm collection included in the RAPD marker analysis.

Acc. #	Clone name	Country origin	Region within country	Map location
1	AMAZON 2-1	Peru	Loreto, Rio Amazonas	73.1°W 3.4°S
2	AMAZON 3-2	Peru	Loreto, Rio Amazonas	73.1°W 3.4°S
3	CAB 5.003-23	Brazil	Roraima	–
4	CAB 36	Brazil	Para, Rio Acara	48.0°W 3.0°S
5	CAB 15	Brazil	Roraima, Rio Jiparana	61.5°W 11°S
6	CCN 10	Ecuador	–	–
7	CCN 34	Ecuador	–	–
8	CCN 51	Ecuador	–	–
9	C SUL 8	Brazil	Acre, Rio Jurua, Floresta/Guajara	72.5°W 8.1°S
10	C SUL 3	Brazil	Acre, Rio Jurua, Ilha Guajara	72.5°W 8.1°S
11	C SUL 4	Brazil	Acre, Rio Jurua, Ilha Guajara	72.5°W 8.1°S
12	SCA 6	Peru	Loreto	73.0°W 3.4°S
13	SCA 12	Peru	Loreto	73.0°W 3.4°S
14	TSA 516	Trinidad/Tobago	–	–
15	TSA 641	Trinidad/Tobago	–	–
16	TSH 1188	Trinidad/Tobago	–	–
17	TSH 565	Trinidad/Tobago	–	–
18	EET 376	Ecuador	Los Rios	79.5°W 1°S
19	EET 390	Ecuador	Los Rios	79.5°W 1°S
20	IAC 1	Brazil	–	69°W 10°S
21	CEPEC 38	Brazil	Bahia, Ilheus	39°W 15°S
22	CEPEC 46	Brazil	Bahia, Ilheus	39°W 15°S
23	CEPEC 89	Brazil	–	–
24	CEPEC 92	Brazil	–	–
25	NA 33	Peru	Loreto, Rio Nanay	73.0°W 3.3°S
26	NA 312	Peru	Loreto, Rio Nanay	73.0°W 3.3°S
27	NA 727	Peru	Loreto, Rio Nanay	73.0°W 3.3°S
28	MA 16	Brazil	Amazonas, Rio Amazonas, Terra Nova	60.0°W 3.0°S
29	MA 13	Brazil	Amazonas, Santo Antonio, Terra Nova	60.0°W 3.0°S
30	MA 15	Brazil	Amazonas, Santo Antonio, Terra Nova	60.0°W 3.0°S
31	IMC 67	Peru	Loreto, Iquitos	73.0°W 3.5°S
32	IMC 27	Peru	Loreto, Iquitos	73.0°W 3.5°S
33	IMC 47	Peru	Loreto, Iquitos	73.0°W 3.5°S
34	MOQ 216	Ecuador	Hda. Moquique	–
35	MOQ 417	Ecuador	Hda. Moquique	–
36	CA 1	Brazil	Amazonas, Bac. do Solimoes, Careiro, Ilha do Parana	60.0°W 3.5°S
37	CA 5	Brazil	Amazonas, Bac. do Solimoes, Careiro	60.0°W 3.5°S
38	CA 2	Brazil	Amazonas, Bac. do Solimoes, Careiro, Ilha do Parana	60.0°W 3.5°S
39	OC 77	Venezuela	Aragua, Ocumare de la Costa	67.4°W 10.3°S
40	OC 66	Venezuela	Aragua, Ocumare de la Costa	67.4°W 10.3°S
41	SGU 26	Guatemala	E.E. Los Brillantes, Santa Cruz, Mulúa, Retalhuleu	–
42	SGU 50	Guatemala	E.E. Los Brillantes, Santa Cruz, Mulúa, Retalhuleu	–
43	SGU 54	Guatemala	E.E. Los Brillantes, Santa Cruz, Mulúa, Retalhuleu	–
44	EET 45	Ecuador	Los Rios	79.5°W 1.0°S
45	EET 62	Ecuador	Los Rios	79.5°W 1.0°S
46	EET 59	Ecuador	Los Rios	79.5°W 1.0°S
47	SIAL 84	Brazil	Bahia, Jucurucu	39.0°W 17.0°S
48	SIAL 70	Brazil	Bahia, Jucurucu	39.0°W 17.0°S
49	SIC 24	Brazil	Bahia, Urucuca	39.0°W 15°S
50	SIC 662	Brazil	Bahia, Urucuca	39.0°W 15°S
51	EEG 29	Brazil	Espirito Santo, E.E. Goitacaces	40°W 19.5°S
52	EEG 50	Brazil	Espirito Santo, E.E. Goitacaces	40°W 19.5°S
53	ICS 9	Trinidad/Tobago	Hda. Phips	–
54	ICS 1	Trinidad/Tobago	River Estate	–
55	ICS 39	Trinidad/Tobago	River Estate	–
56	UF 221	Costa Rica	Limon Almirante	–
57	UF 667	Costa Rica	Atlantic Coast	–
58	UF 677	Costa Rica	Atlantic Coast	–
59	CAS 2	Brazil	Para, Santarém	55.0°W 3.0°S
60	MOCO 1	Brazil	Para, Santarém	55.0°W 3.0°S
61	P 19	Peru	Loreto, Iquitos	73.0°W 3.5°S
62	P 4 B	Peru	Loreto, Iquitos	73.0°W 3.5°S
63	P 11	Peru	Loreto, Iquitos	73.0°W 3.5°S
64	SPA 12	Colombia	Valle del Cacau, Palmira	76.5°W 3.5°N
65	SPA 5	Colombia	Valle del Cacau, Palmira	76.5°W 3.5°N
66	SPA 17	Colombia	Valle del Cacau, Palmira	76.5°W 3.5°N
67	PA 51	Peru	Loreto, Rio Marañón, Parinari	74.6°W 4.6°S
68	PA 150	Peru	Loreto, Rio Marañón, Parinari	74.6°W 4.6°S
69	PA 13	Peru	Loreto, Rio Marañón, Parinari	74.6°W 4.6°S
70	Be 4	Brazil	Para, Belem, Rio Guama, Abaetetuba	48.0°W 1.3°S
71	Be 6	Brazil	Para, Belem, Rio Guama, Fuio Grande	48.0°W 1.3°S
72	SPEC 54-1	Colombia	Vaupes, Rio Papuri	69.0°W 1.0°S
73	SPEC 138-8	Colombia	Valle del Cacau, Palmira	76.2°W 1.3°N
74	RB 36	Brazil	Acre, Isla Amapa, Rio Acre	67.5°W 9.3°S
75	RB 39	Brazil	Acre, Isla Amapa, Rio Acre	67.5°W 9.3°S
76	RB 37	Brazil	Acre, Isla Amapa, Rio Acre	67.5°W 9.3°S
77	GS 36	Grenada	Hda Bouloque	61.5°W 12.1°N

Continued next page.

Table 1. Continued.

Acc. #	Clone name	Country origin	Region within country	Map location
78	GS 29	Grenada	Hda Bouloque	61.5°W 12.1°N
79	SC 5	Colombia	Valle del Cacau, Palmira	76.5°W 3.5°N
80	SC 49	Colombia	Valle del Cacau, Palmira	76.5°W 3.5°N
81	RIM 76	Mexico	Chiapas, Tuxtla Gutierrez	–
82	RIM 52	Mexico	Chiapas, Tuxtla Gutierrez	–
83	RIM 15	Mexico	Chiapas, Tuxtla Gutierrez	–
84	21 P (J)	Mexico	Tobasco	–
85	8 (P)	Mexico	Tobasco	–
86	CC 41	Costa Rica	Cartago, Turrialba	–
87	CC 11	Costa Rica	Limon	–
88	CC 10	Costa Rica	Limon	–
89	CJ 7	Brazil	Para, Jari	52.4°W 0.1°S
90	CJ 4	Brazil	Para, Jari	52.4°W 0.1°S
91	EET 377	Ecuador	–	–
92	EET 392	Ecuador	–	–
93	CEPEC 90	Brazil	–	–
94	C SUL 7	Brazil	Acre, Rio Jurua, Floresta/Guajara	72.5°W 8.1°S
95	C SUL 10	Brazil	Acre, Rio Jurua, S. Luiz, Guajara	72.5°W 8.1°S
96	EQX 107	Ecuador	E.E. Pichilingue, Quevedo, Los Rios	79.5°W 1°S
97	CEPEC 94	Brazil	–	–
98	TSA 644	Trinidad/Tobago	–	–
99	CEPEC 25	Brazil	Espirito Santo, Linhares	40.0°W 19°S
100	PLAYA ALTA 4	Venezuela	Rio Orinoco delta	61.0°W 9.0°S
101	CHUAO 120	Venezuela	Valle de Chuao	–
102	OC 67	Venezuela	Aragua, Ocumare de la Costa	67.4°W 10.3°N
103	C 87.56	Trinidad/Tobago	River Estate	–
104	CAB 4	Brazil	Amazonas, Rio do Solimoes	60.0°W 3.5°S
105	RB 38	Brazil	Acre, Isala Amapa, Rio Acre	67.5°W 9.3°S
106	CEPEC 86	Brazil	Bahia	
107	CEPEC 11	Brazil	Bahia, Ilheus	39.0°W 15°S
108	CEPEC 12	Brazil	Bahia, Ilheus	39.0°W 15°S
109	CEPEC 13	Brazil	Bahia, Ilheus	39.0°W 15°S
110	CEPEC 14	Brazil	Bahia, Ilheus	39.0°W 15°S
111	CEPEC 15	Brazil	Bahia, Ilheus	39.0°W 15°S
112	CEPEC 533	Brazil	Bahia, Ilheus	39.0°W 15°S
113	CEPEC 538	Brazil	Bahia, Ilheus	39.0°W 15°S
114	CEPEC 550	Brazil	Bahia, Ilheus	39.0°W 15°S
115	SIAL 20	Brazil	Bahia, Jucurucu	39.0°W 17°S
116	SIC 19	Brazil	Bahia, Urucuca	39.0°W 15°S
117	EET 53	Ecuador	Los Rios	79.5°W 1.0°S
118	EET 94	Ecuador	Tenguel, Prov. Guayas	80.5°W 2.0°S
119	P 7	Peru	Loreto, Iquitos	73.0°W 3.5°S
120	P 16	Peru	Loreto, Iquitos	73.0°W 3.5°S
121	PA 4	Peru	Loreto, Rio Marañón, Parinari	74.6°W 4.6°S
122	OB 52	Brazil	Para, Obidos	55.0°W 2.0°S
123	SIAL 512	Brazil	Bahia, Jucurucu	39.0°W 17°S
124	EEG 14	Brazil	Espirito Santo, E.E. Goitacaces	40°W 19.5°S
125	ICS 8	Trinidad/Tobago	Hda. Phips	–
126	ICS 60	Trinidad/Tobago	–	–
127	Be 3	Brazil	Para, Balem, Rio Guama	48°W 1.3°S
128	CC 34	Costa Rica	Limon, Los Diamantes	–
129	EET 397	Ecuador	E.E. Pichilingue, Quevedo, Los Rios	–
130	CEPEC 42	Brazil	Bahia, Ilheus	39.0°W 15°S
131	SIAL 283	Brazil	Bahia, Jucurucu	39.0°W 17°S
132	Be 5	Brazil	Para, Balem, Rio Guama	48.0°W 1.3°S
133	C SUL 2	Brazil	Acre, Rio Jurua, Faz. S. Luiz	72.5°W 8.1°S
134	C SUL 5	Brazil	Acre, Rio Jurua, Ilha Guajara	72.5°W 8.1°S
135	C SUL 9	Brazil	Acre, Rio Jurua, S. Luiz Guajara	72.5°W 8.1°S
136	CEPEC 87	Brazil	Bahia	–
137	SPA 7	Colombia	Valle del Cacau, Palmira	76.5°W 3.5°N
138	EET 399	Ecuador	–	–
139	RB 30	Brazil	Acre, Xapuri, Rio Branco	60.2°W 3.5°S
140	CAS 3	Brazil	Para, Santarém	55.0°W 3.0°S
141	Be 8	Brazil	Para, Balem, Rio Guama	48.0°W 1.3°S
142	SIAL 407	Brazil	Bahia, Jucurucu	39.0°W 17°S
143	PA 169	Peru	Loreto, Rio Marañón, Parinari	74.6°W 4.6°S
144	ICS 98	Trinidad/Tobago	Hda Poole, Rio Claro	–
145	ICS 89	Trinidad/Tobago	–	–
146	EET 61	Ecuador	Los Rios	79.5°W 1.0°S
147	CA 6	Brazil	Amazonas, Careiro	60.0°W 3.5°S
148	CEPEC 16	Brazil	Bahia, Ilheus	39.0°W 15°S
149	EEAT 228	Ecuador	Guayas	80.5°W 2.0°S
150	SIAL 543	Brazil	Bahia, Jucurucu	39.0°W 17°S
151	SIAL 505	Brazil	Bahia, Jucurucu	39.0°W 17°S
152	CA 3	Brazil	Amazonas, Bac do Solimoes, Careiro, Ilha do Parana	60.0°W 3.5°S
153	IMC 76	Peru	Loreto, Iquitos	73.0°W 3.5°S
154	APA 4	Colombia	Valle del Cauca, Palmira	76.5°W 3.5°N

Continued next page.

Table 1. Continued.

Acc. #	Clone name	Country origin	Region within country	Map location
155	PA 148	Peru	Loreto, Rio Marañón, Parinari	74.6°W 4.6°S
156	RIM 10	Mexico	Chiapas, Tuxtla Gutiérrez	–
157	ICS 6	Trinidad/Tobago	River Estate	–
158	EEG 65	Brazil	Espirito Santo, E. E. Goitacaces	40°W 19.5°S
159	SIC 328	Brazil	Bahia, Urucuca	39.0°W 15°S
160	TSA 654	Trinidad/Tobago	–	–
161	TSH 774	Trinidad/Tobago	–	–
162	CEPEC 523	Brazil	Bahia, Ilhues	39.0°W 15°S
163	CEPEC 541	Brazil	Bahia, Ilheus	39.0°W 15°S
164	ICS 16	Trinidad/Tobago	Hda. Phips	
165	CEPEC 73	Brazil	Bahia, Ilheus	39.0°W 15°S
166	JA 546	Ecuador	Hda. Javilla	–
167	C 13.5	Nicaragua	–	–
168	CCN 2	Ecuador	–	–
169	MO 9	Peru	Loreto, Rio Morona	76.8°W 4.8°S
170	CJ 8	Brazil	Para, Jari	52.4°W 0.1°S
171	10 (P)	Mexico	Tabasco	–
172	UF 12	Costa Rica	Limon Almirante	–
173	CEPEC 519	Brazil	Bahia, Ilheus	39.0°W 15°S
174	CEPEC 48	Brazil	Bahia, Ilheus	39.0°W 15°S
175	CEPEC 30	Brazil	Bahia, Ilheus	39.0°W 15°S
176	UF 296	Costa Rica	Atlantic Coast	–
177	ICS 75	Trinidad/Tobago	–	–
178	CEPEC 532	Brazil	Bahia, Ilheus	39.0°W 15°S
179	RIM 105	Mexico	Chiapas, Tuxtla Gutiérrez	–
180	22 (P)	Brazil	Tabasco	–
181	CAB 2	Brazil	Amazonas, Rio Solimoes	66.0°W 3.5°S
182	CEPEC 95	Brazil	Bahia	–
183	CEPEC 108	Brazil	Bahia, Belmonte	39.0°W 16°S
184	CEPEC 131	Brazil	Bahia, Canavieras	39.0°W 16°S
185	CEPEC 171	Brazil	Bahia, Santa Luzia	–
186	CAB 262	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3.0°S
187	CAB 263	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3.0°S
188	CAB 148	Brazil	Acre, Rio Acre	67.5°W 9.3°S
189	CAB 94	Brazil	Amazonas, Rio Purus	67.4°W 8.0°S
190	CAB 520	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3.0°S
191	CAB 157	Brazil	Amazonas, Rio Acre	67.5°W 9.3°S
192	LCTEEN 28s1	Ecuador	Itaya, Rio Capocuy	76.34°W 0.2°S
193	GU 125 C	French Guiana	Haut Camopi, Rio Camopi	56.1°W 2.3°N
194	GU 136 H	French Guiana	Haut Camopi, Rio Camopi	56.1°W 2.3°N
195	CAB 414	Brazil	Rondônia, Rio Jamari	63.2°W 8.5°S
196	CAB 61	Brazil	Para, Rio Tapajos	55.4°W 10°S
197	CEPEC 144	Brazil	Bahia, Itabuna	39.0°W 15°S
198	CEPEC 148	Brazil	Bahia, Mata de São João	38.0°W 11°S
199	CAB 275	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3.0°S
200	CAB 53	Brazil	Para, Salgado (Zona)	48.0°W 1.0°S
201	CAB 65	Brazil	Acre, Rio Iaco	69.0°W 9.0°S
202	CEPEC 159	Brazil	Bahia, Mucuri	40.0°W 18°S
203	CEPEC 125	Brazil	Bahia, Canavieras	39.0°W 16°S
204	H 28	Peru	Loreto, Rio Huallago	75.5°W 5.5°S
205	LCTEEN 37F	Ecuador	Rio Napo, Rio Anangu	76.34°W 0.2°S
206	CAB 194	Brazil	Amazonas, Rio Purus	67.4°W 8.0°S
207	CAB 103	Brazil	Amazonas, Rio Purus	67.4°W 8.0°S
208	CEPEC 151	Brazil	Bahia, Mata de São João	38.0°W 11.°S
209	H 7	Peru	Loreto, Rio Huallaga	75.5°W 5.5°S
210	CAB 486	Brazil	Amazonas, Rio Japura	68.0°W 1.5°S
211	H 39	Peru	Loreto, Rio Huallaga	75.5°W 5.5°S
212	CAB 21	Brazil	Para, Rio Maicuru	54.1°W 2.0°S
213	CAB 155	Brazil	Acre, Rio Acre	67.5°W 9.3°S
214	CEPEC 158	Brazil	Bahia, Mucuri	40.0°W 18°S
215	H 17	Peru	Loreto, Rio Huallaga	75.5°W 5.5°S
216	CEPEC 166	Brazil	Bahia, Mucuri	40.0°W 18°S
217	CEPEC 147	Brazil	Bahia, Mata de São João	38.0°W 11°S
218	H 9	Peru	Loreto, Rio Huallaga	75.5°W 5.5°S
219	CEPEC 150	Brazil	Bahia, Mata de São João	38.0°W 11°S
220	CEPEC 136	Brazil	Bahia, Canavieras	39.0°W 16°S
221	U 14	Peru	Loreto, Rio Ucayali	74.0°W 5°S
222	CAB 312	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3°S
223	CAB 201	Brazil	Acre, Rio Tarauaca	71.3°W 8.3°S
224	CAB 505	Brazil	Amazonas, Rio Japura	68.0°W 1.5°S
225	CAB 299	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3°S
226	CAB 531	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3°S
227	CAB 108	Brazil	Acre, Rio Acre	67.5°W 9.3°S
228	CAB 68	Brazil	Acre, Rio Iaco	69.0°W 9°S
229	CAB 165	Brazil	Amazonas, Rio Acre	67.5°W 9.3°S
230	CAB 130	Brazil	Acre, Rio Purus	67.4°W 8.0°W
231	SA 3	Brazil	Rondônia	
232	CAB 382	Brazil	Rondônia, Rio Jamari	63.2°W 8.5°S

Continued next page.

Table 1. Continued.

Acc. #	Clone name	Country origin	Region within country	Map location
233	CAB 223	Brazil	Rondônia, Rio Jamari	63.2°W 8.5°S
234	CAB 231	Brazil	Rondônia, Rio Jamari	63.2°W 8.5°S
235	CAB 44	Brazil	Mato Grosso, Maranhao Goiás, Rio Paranaita	56.4°W 10°S
236	CAB 283	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3°S
237	CAB 353	Brazil	Amazonas, Rio Curuca	72.0°W 4.5°S
238	CAB 460	Brazil	Rondônia, Rio Jiparana	61.5°W 11°S
239	CAB 224	Brazil	Rondônia, Rio Jamari	63.2°W 8.5°S
240	U 32	Peru	Loreto, Rio Ucayali	74.0°W 5°S
241	COCA 3370-5	Ecuador	Rio Napo, Rio Coca	77.0°W 0.2°S
242	AMAZON 15	Peru	Loreto, Rio Amazonas	72.0°W 3.0°S
243	GU 221 C	French Guiana	Haut Camopi, Rio Camopi	56.1°W 2.3°N
244	CAB 121	Brazil	Acre, Rio Chandless	70.0°W 10°S
245	CAB 305	Brazil	Amazonas, Rio Solimoes, B. Japura	65.0°W 3.0°S
246	CAB 380	Brazil	Rondônia, Rio Jamari	63.2°W 8.5°S
247	CAB 389	Brazil	Rondônia, Rio Jamari	63.2°W 8.5°S
248	CAB 252	Brazil	Rondônia, Rio Jamari	63.2°W 8.5°S
249	LCTEEN 7A	Ecuador	Sucumbíos, Rio Aguarico	76.3°W 0.2°S
250	GU 154 C	French Guiana	Haut Camopi, Rio Camopi	56.1°W 2.3°N
251	PA 175	Peru	Loreto, Rio Marañón, Parinari	74.6°W 4.6°S
252	LCTEEN 163A	Ecuador	Rio Payamino	77.05°W 0.2°S
253	U 2	Peru	Loreto, Rio Ucayali	74.0°W 5°S
254	GNV 225	USA	–	–
255	IMC 83	Peru	Loreto, Iquitos	73.0°W 3.5°S
256	GNV 31	USA	–	–
257	GNV 111	USA	–	–
258	EET 19	Ecuador	Guayas	80.5°W 2.0°S
259	EET 58	Ecuador	Los Rios	79.5°W 1.0°S
260	P 18	Peru	–	–
261	SC 10	Colombia	Valle del Cauca, Palmira	76.5°W 3.5°N
262	EQXZ	Ecuador	Rio Payamino	–
263	SPEC 160.9	Colombia	Antioquia Rio Cauca	75.5°W 6.4°N
264	GU 121	French Guiana	Haut Camopi, Rio Camopi	56.1°W 2.3°N
265	GU 133 C	French Guiana	Haut Camopi, Rio Camopi	56.1°W 2.3°N
266	LCTEEN 241	Ecuador	Para, Rio Villano	77.27°W 1.3°S
267	SC 3	Colombia	Valle del Cauca, Palmira	76.5°W 3.5°N
268	SCA 2	Peru	Loreto	73.0°W 3.4°S
269	P 5 C	Peru	–	–
270	MO 20	Peru	Loreto, Rio Morona	77.0°W 4.0°S

the disease by Stahel (1915) in Surinam. It has since spread to many other cacao producing countries (Pereira et al., 1990). Vegetative flushes are the most abundant infection sites for the witches' broom pathogen (Sreenivasan and Dabydeen, 1989); however, petioles, pulvini, and stems are also susceptible to infection (Cronshaw and Evans, 1978). In Bahia, Brazil, where 85% of Brazilian cocoa grows in the understory of the rainforest under humid conditions favoring germination of the fungus' basidiospores (Frias et al., 1991; Pereira et al., 1990), the pathogen has coevolved with cacao and targets the cushion of the cacao tree. Loss of production due to witches' broom disease in important cacao growing areas such as Bahia, Brazil, has generated a strong demand for resistant varieties.

Although considerable genetic diversity exists in cacao germplasm collections, the genotypes used in production have a very narrow genetic base derived from few genotypes (Cope, 1976). The impact of diseases such as witches' broom on cacao yield and production has highlighted the need to characterize existing germplasm collections and examine the genetic diversity within the collections for sources of resistance. In addition, the expense associated with the maintenance of cacao collections suggests the need for accurate and detailed accession characterization to manage collections efficiently. Understanding the structure of genetic diversity within the collections can permit more efficient sampling of germplasm resources to maximize diversity and minimize redundancy. Better understanding of the genetic relationships among cacao germplasm collections will facilitate the identification of germplasm closely related to resistant accessions for use as additional parental sources of resistance.

Several studies have demonstrated the usefulness of RAPDs in studying cacao. RAPDs have been applied in studying cacao genetic diversity (Whitkus et al., 1998; Lerceteau et al., 1997: Figueira et al., 1994; N'Goran et al., 1994; Laurent et al., 1994), creating genetic linkage maps (Crouzillat et al., 1996; Lanaud et al., 1995), and DNA fingerprinting (Wilde et al., 1992). Molecular marker techniques, specifically using RAPDs, allow an effective and accurate method of characterizing germplasm resources. The extensive collection held at CEPEC provides an opportunity to sample and accurately characterize a significant portion of the cacao germplasm available to breeders and breeding programs.

The objective of this study was to characterize the genetic diversity within the germplasm collection at CEPEC by means of RAPD marker based estimates of genetic relationships among a sample of 270 cacao accessions. A second objective was to assess disease tolerance relationships among a subset of 180 cacao accessions to identify the relationship between the struc-

ture of genetic diversity and the distribution of accessions resistant to witches' broom disease.

MATERIALS AND METHODS

Germplasm

A subset of 270 cacao accessions (Table 1; Marita et al., 2000) from the more than 1000 accessions housed at CEPEC was sampled for germplasm evaluation. Accessions examined formed a diverse collection of selections from normally cultivated germplasm, trees grown on farms heavily infected with witches' broom disease, and unadapted material. The first set of accessions sampled (1–90; Table 1) included two to three accessions from many important agronomic "series" selected by the cacao breeder at CEPEC, representing unique clones with possible resistance to witches' broom disease. A "series" represents accessions categorized under a single acronym. For example, C SUL accessions are unique clones collected from Cruzeiro do Sol in the Upper Amazon region of Brazil. Many series represent accessions collected during some of the first collection expeditions and form the parental basis of early cacao breeding programs. The second set (91–180; Table 1) included a sample of accessions included in genetic characterization experiments conducted by CEPEC. The third set (181–270; Table 1) represented randomly selected accessions from among newly acquired accessions not available initially. Three accessions classified as USA (254, 256, and 257) referenced the location where the crosses were made and not the origin of the genetic material involved in the crosses.

DNA Extraction and RAPD Reactions

Fresh leaf tissue was supplied from CEPEC's living field cacao collection where leaf tissue from 10 clonal cacao trees was pooled for each accession. DNA extractions were done in cooperation with Fazenda Almirante. DNA was extracted from approximately 50 mg of mature leaf tissue per individual cacao tree. The DNA isolation followed a modified method from Doyle and Doyle (1990). The DNA was subsequently purified following the protocol from the Prep-A-Gene DNA Purification Systems kit (BIO-RAD Laboratories, Hercules, CA). The RAPD reaction mixtures followed Skroch and Nienhuis (1995a) while RAPD cycling conditions followed Johns et al. (1997). A total of 38 RAPD markers were used. RAPD primers A15, A17, A19, B1, B5, B6, B8, B12, B15, C1, C2, C8, C9, C13, E11, E18, E19, F2, F9, F16, I11, J10, J19, K7, K8, L5, M17, Z6, AJ2, and AP2 (Operon Technologies, Inc., Alameda, CA) were selected for screening on the basis of reports of RAPD polymorphisms in cacao (Ronning et al., 1995; N'Goran et al., 1994; Figueira et al., 1994). To maximize sampling of the entire cacao genome, additional primers (C15, M10, P14, Q4, R19, AP4, AW18, and AX15; Operon Technologies, Inc., Alameda, CA) were selected on the basis of dispersal on published linkage maps (Crouzillat et al., 1996; Lanaud et al., 1995). Data were scored as the presence (1) or absence (0) of a fragment for each polymorphic marker (Marita, 1998).

RAPD Analysis

Genetic distances were calculated among all pairwise combinations of the 270 accessions by means of the complement to the simple matching coefficient (Gower, 1985).

$$\text{Genetic Distance } (i,j) = \sum n(i \neq j)/[\sum n(i \neq j) + \sum n(i = j)],$$

where i and j represented a pairwise combination of accessions and n represented the number of markers scored for those two accessions. Genetic distance was the ratio of discordant comparisons to the total number of comparisons. A genetic distance equaling 0 or 1 indicated maximum similarity or difference between the pair of accessions, respectively. To visualize the relationships among the accessions, the genetic distance matrix was converted to two-dimensional coordinates by the monotonic multidimensional scaling procedure (MDS) by the Kruskal scaling option in Systat ver. 5.2 (Wilkinson et al., 1992).

Thirteen randomly selected accessions were sampled twice and each separate DNA extraction treated as a separate entry. The reproducibility error was calculated as the failure of accessions to be scored consistently over replications because of random error in the generation and scoring of data.

Germplasm Characterization

Differences in the distribution of genetic diversity present within the original group sampled (Accessions 1–180) compared with the newly acquired accessions (181–270) were estimated as RAPD marker diversity for each subgroup by Nei's genetic diversity at a locus,

$$h = (1 - \sum x_i^2)n_i/(n_i - 1) = (2pqn)/(n - 1),$$

where x_i was the allele frequency at the ith locus, p was the frequency of presence, and q was the frequency of absence of RAPD bands among n accessions for the ith RAPD marker (Nei, 1987). The two groups could be defined genetically by allele frequencies such that changes in these frequencies reflect genetic changes in the population (Nei, 1987). In addition, we examined whether the breadth of genetic diversity represented by the origianl group increased with the inclusion of newly acquired accessions. The unique clustering of each group was determined by differences in RAPD marker frequency between the same two groups. The significance level of each comparison was determined by randomization tests (Edgington, 1980) on the basis of 10 000 permutations appropriate to each comparison.

To assess the clustering of accessions within an individual country of origin compared with all other countries, differences in mean RAPD marker frequencies were analyzed. Those countries with fewer than five accessions were excluded from the analysis since results were difficult to interpret. The frequency of each RAPD marker was calculated as the frequency of the presence of RAPD amplification among accessions in a subgroup defined by a country of origin. Comparisons were made between accessions within an individual country of origin and all remaining accessions. To test the difference in mean RAPD marker frequencies between a subgroup and remaining accessions, the sum of the absolute value of the difference between RAPD marker frequencies of comparison groups was divided by the number of RAPD markers evaluated. The significance level for each paired comparison was determined by a randomization test on the basis of 10 000 random permutations appropriate to each comparison (Edgington, 1980).

Witches' Broom Disease Resistance

Witches' broom disease data were taken at CEPEC on the basis of natural infestation for Accessions 1 to 180 (Marita, 1998). The total number of vegetative brooms per plant and the total number of cushion brooms per plant averaged over 10 plants were reported and Pearson product moment correlations calculated. Accessions 4 and 146, were excluded from

analysis because of discrepancies in age of tissue and plant location on the farm, respectively. A MDS plot representing accessions tolerant to vegetative broom infection and tolerant to cushion broom infection was examined to evaluate relationships between the two categories of infection.

Differences in RAPD marker frequency of groups defined as accessions within the lower 15% of the total infected broom distribution were compared with (i) remaining accessions and (ii) accessions within the upper 15% of the distribution. These analyses were based on variation in polymorphic marker frequency between comparison sets. The significance of these observed differences was tested using a chi-square test for goodness of fit (Snedecor and Cochran, 1992).

Differences in RAPD marker frequency for individual RAPD loci evaluated were analyzed by means of two-way cross tabulation tables (PROC FREQ; SAS Institute, 1990). An individual RAPD marker frequency was equal to $\Sigma(p_i)/n$ where p_i is the frequency of the presence of RAPD bands among accessions in a subpopulation for the ith RAPD marker and n is the number of accessions. Comparison groups were identical to those used to test differences in RAPD marker frequency of subgroups defined above. The low frequency of RAPD alleles in some cases required that the significance tests be based on exact probabilities rather than approximations. Thus, all P-values were computed by Fisher's exact t-test and compared at the 5% significance level (P-values ≤ 0.025) (Steel and Torrie, 1980).

Two comparisons were performed (i) to test whether Upper Amazon accessions formed a unique cluster, and (ii) to test whether Upper Amazon accessions and resistant accessions formed the same cluster by testing the difference in mean RAPD marker frequencies between the two groups (see Germplasm Characterization). For comparison purposes, an accession collected west of the 58° longitude was classified as Upper Amazon and an accession within the lower 15% of the total infected broom distribution was classified as tolerant.

RESULTS AND DISCUSSION

RAPD Analysis

The 38 RAPD primers resulted in 133 scorable polymorphic RAPD markers. Between one and seven polymorphic markers per primer were scored with a mean of 3.5 markers per primer. This is consistent with previously reported average number of markers per primer (N'Goran et al., 1994; Figueira et al., 1994). The use of prior publications and linkage maps to help select markers that were not clustered genetically did not result in a reduction in observed polymorphism.

The distribution of simple linear correlations between 133 RAPD markers among 270 cacao accessions (8778 pairwise combinations) resulted in a normal distribution with a mean (±SD) of 0.022 (±0.167). No two RAPD markers had identical amplification patterns across any pair of accessions, suggesting the RAPD markers used in this study were distributed with minimal clustering on the cacao genetic linkage map.

Genetic distances among the 36 315 pairwise combinations of accessions ranged from 0.000 to 0.5691 with a mean (±SD) genetic distance of 0.310 (±0.09). The mean reproducibility error calculated from 13 replicates was 1.26% (Table 2). Genetic distances less than 1.26% (0.0126) could not be discriminated between different accessions. Fewer than 0.08% of the pairwise comparisons were indistinguishable. Since scoring errors are non-randomly distributed with a large portion distributed within a relatively small subset of RAPD markers (Skroch and Nienhuis, 1995b), the removal of 12 RAPD markers (8%) with inconsistent amplification across replications was warranted. In general, the low reproducibility error suggested that genetic distance estimates based on RAPD markers were highly reliable (and repeatable) among cacao accessions and the markers used in this study could uniquely distinguish among most accessions.

Table 2. Estimate of reproducibility error based on the failure of 13 replicated accessions to be scored consistently.

Acc. Number	Clone Name	Number of inconsistencies	Number of RAPD bands scored†	Percent Error
33	IMC 47	1	144	0.7
39	OC 77	0	145	0.0
50	SIC 662	1	144	0.7
60	Mocorongo 1	0	145	0.0
82	RIM 52	2	143	1.3
83	RIM 15	2	143	1.3
95	C SUL 10	4	141	2.8
111	CEPEC 15	3	142	2.1
122	OB 52	2	143	1.3
126	ICS 60	2	143	1.3
146	EET 61	3	142	2.1
151	SIAL 505	1	144	0.7
173	CEPEC 519	3	142	2.1
Mean reproducibility error				1.26

† Number of comparisons excluding bands with missing data. Total possible is 145 RAPD markers.

Germplasm Genetic Relationships

Analyzing 270 accessions from CEPEC's more than 1000 accessions was a critical step towards understanding the genetic relationships of accessions that comprise the entire collection. The genetic relationships among all cacao accessions were displayed as a MDS plot of the 270-by-270 genetic distance matrix (Fig. 1). No difference was observed between the original group (Accessions 1–180) sampled and the "new" group (Accessions 181–270) on the basis of analysis of RAPD frequency. In addition, differences based on Nei's genetic diversity were not observed between the two groups. The incorporation of the new accessions did not increase the magnitude of genetic diversity already present within the original group sampled. This may be a result of resampling geographic origins already represented in the original collection, since many of the new accessions came from the same geographic regions or were hybrids between parents from regions sampled in the original collection (Table 1). This suggests that new collections of accessions within already sampled geographic regions did not increase the existing genetic diversity in the cacao germplasm, and additional accessions need to be evaluated to assess whether they represent unique genetic diversity compared with accessions already present in the collection.

The number of accessions represented by each country varied from 1 (Nicaragua) to 134 (Brazil). A total of 134 accessions were classified as "Brazilian" but these represented both unique accessions and hybrids be-

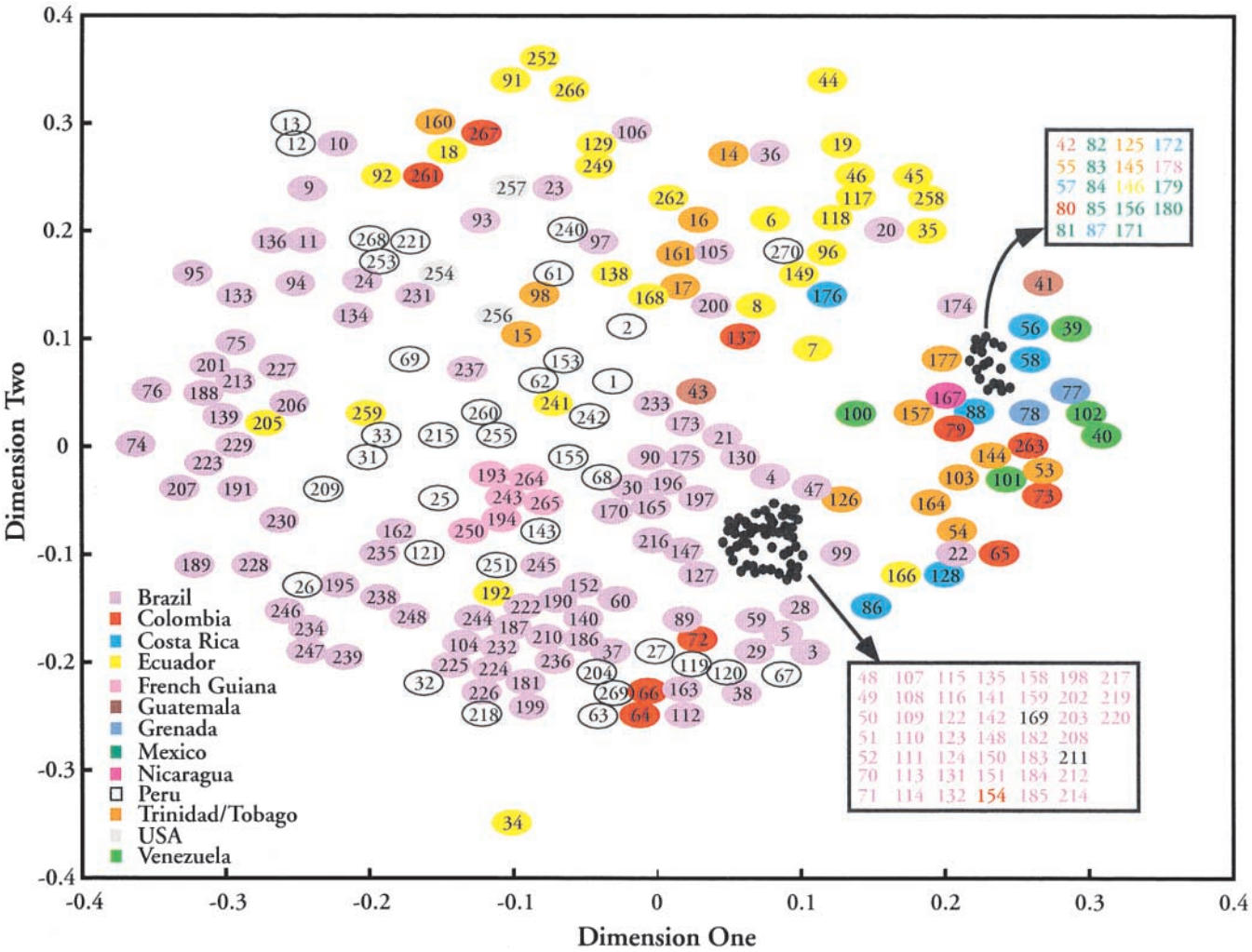

Fig. 1. A multidimensional scaling plot of 270 cacao accessions analyzed from CEPEC's germplasm collection determined on the basis of RAPD data.

tween accessions from Brazil and other countries. As a result, accessions categorized as "Brazil" were distributed throughout the MDS plot (Fig. 1). Peru, Ecuador, Trinidad–Tobago, Mexico, Costa Rica, Venezuela, and French Guiana formed unique but nondiscrete clusters in the MDS plot (Table 3; Fig. 1). Colombian accessions did not form a distinct cluster but were dispersed among Brazilian hybrids and the clusters formed by Peru, Venezuela, and Trinidad–Tobago (Table 3; Fig. 1).

Table 3. Difference in mean RAPD marker frequencies between individual countries of origin and all other countries determined by 10 000 random permutations of a randomization test.

Country of Origin	# of accessions	Mean frequency of country	Mean frequency of all other countries	*P*-value
Brazil	–	–	–	–
Peru	38	0.4235	0.4334	<0.000
Colombia	12	0.4572	0.4309	=0.0619
Ecuador	30	0.4750	0.4269	<0.000
French Guiana	6	0.3697	0.4335	=0.0007
Trinidad/Tobago	18	0.4727	0.4293	<0.000
Venezuela	5	0.4881	0.4310	=0.0063
Mexico	9	0.5759	0.4272	<0.000
Guatemala	–	–	–	
Costa Rica	9	0.5452	0.4281	=0.0003
Grenada	–	–	–	–
Nicaragua	–	–	–	–

Listed sources of germplasm carried many contradictions. In some cases, a region within a country listed an experiment station (e.g., E.E. Los Brillantes for Accessions 41–43) as opposed to a river, farm, or village location as its collection source. In such cases, the map coordinates identified the city where the experiment station was located and not the map coordinate of collection (Table 1). When an experiment station was listed under a region within a country, often the accession listed was a hybrid (i.e., many SGU accessions from Guatemala, OC accessions from Venezuela, and CEPEC accessions from Brazil). In addition, most EET accessions from Ecuador, some CC accessions from Costa Rica, and TSA, TSH, or ICS accessions from Trinidad–Tobago were hybrids categorized by a single country of origin. Furthermore, certain accession origins were unknown, such as several accessions listed under the CEPEC designation (Accessions 86–92). Some of these accessions were from seed confiscated at quarantine stations and then incorporated into CEPEC's collection.

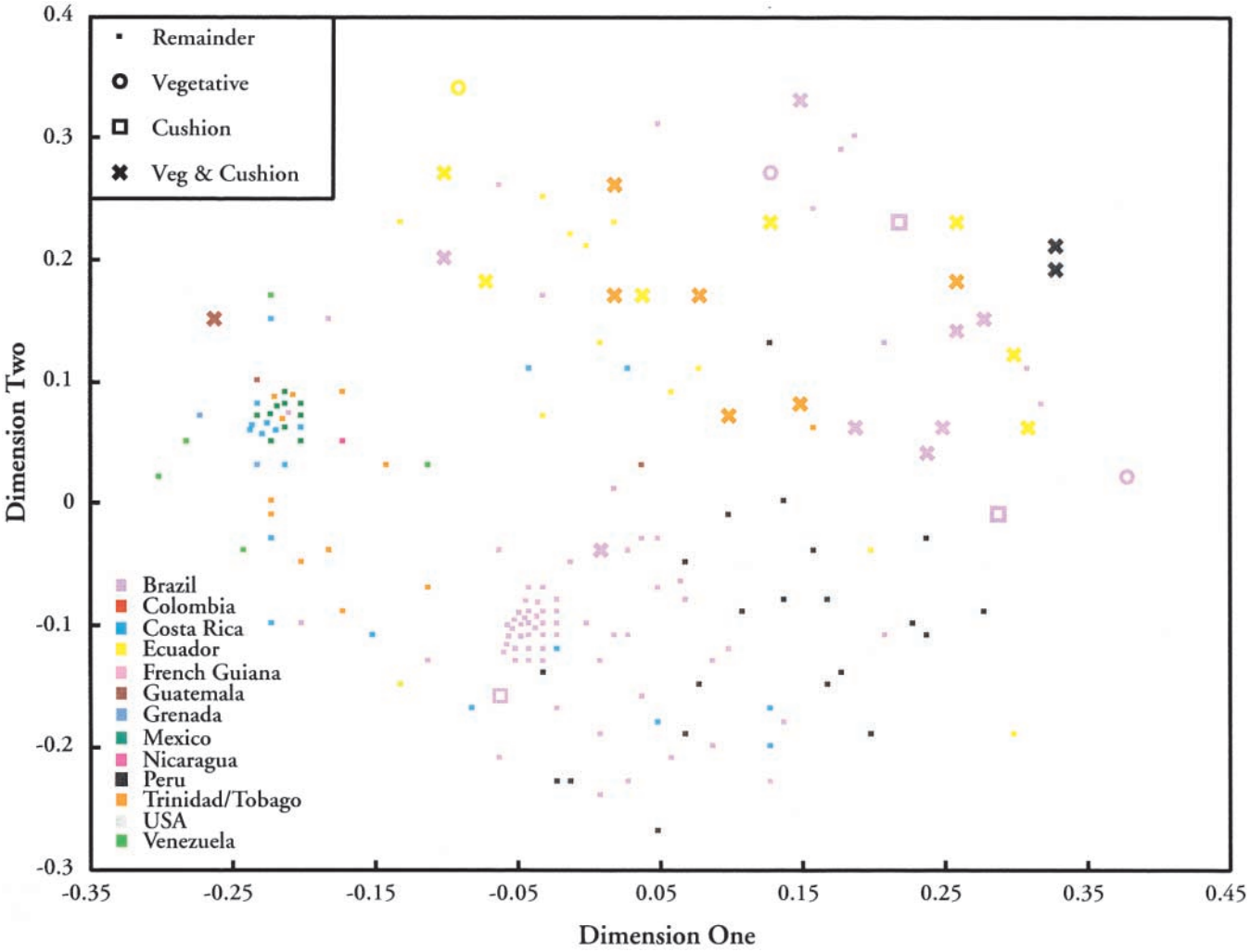

Fig. 2. A multidimensional scaling plot representing accessions in the lower 15% of vegetative and cushion broom distribution determined on the basis of RAPD data. Accessions listed in the lower 15% of both categories are represented by an X.

Other accessions may have been misnamed during handling. Identification of closely related accessions and/or probable identification of such accessions was aided by our analyses. CEPEC 87 (136) was identified as most closely related to C SUL accessions collected along the Juruá River in the northwest region of Brazil. In addition, CEPEC 86 (106) was identified as most closely related to RB or CAB accessions collected along the Acre River in the state of Acre (Table 1; Fig. 1). Accessions designated as CEPEC typically lay along with SIAL and SIC accessions within one cluster represented by the lower, middle, right-hand cluster on the full MDS plot (Fig. 1). As a group, these accessions were representative of early collection expeditions brought back to Bahia, Brazil, and farmed for many years. One might hypothesize that they form a unique cluster where cross- and self-pollination among accessions within this population has existed for many years without new sources of genetic material entering the gene pool of the cluster.

Since Brazil covers such a large area compared with all other country origins, particular attention was paid to where accessions were collected. Sites of collection within Brazil along the Amazon River and its minor rivers were reviewed (End et al., 1992). Many accessions analyzed were collected along main arteries of the Amazon River. Consequently, large gaps existed along minor rivers and lakes throughout the basin. A list was compiled of the rivers flowing into the Amazon River, beginning at the mouth of the river in Para, Brazil, and ending in Loreto, Peru, (Marita, 1998). The list contained approximately 272 rivers and lakes flowing into the Amazon River. With 270 cacao accessions analyzed (134 accessions listed from Brazil), only 22 (8%) of the more

Table 4. Band frequency comparisons between accessions of cacao most tolerant to witches' broom disease and two groupings of remaining accessions of cacao.

Comparison	df	Chi-square	prob > chi-square
Lower 15%† vs remaining set‡	132	561.8	0.000
Lower 15% vs upper 15%§	130¶	316.7	0.000

† This group represents accessions most tolerant to witches' broom disease.
‡ Those accessions represented in sets # 1 and 2 minus those accessions in the lower 15% category.
§ This group represents accessions least tolerant to witches' broom disease.
¶ Two markers monomorphic (no band) across two comparison groups were excluded from the analysis.

Table 5. RAPD markers with significant differences in frequency among accessions in two sub-population comparisons; specifically, accessions within the lowest quartile of the marker distribution compared with the remaining accessions analyzed (comparison #1) and accessions within the lowest quartile of the marker distribution compared with accessions within the upper quartile of the marker distribution (comparison #2).

RAPD marker†	Comparison #1 lower‡	Freq rest‡	Fisher exact test *P*-values§	Comparison #2 lower‡	Freq upper‡	Fisher exact test *P*-values§
$B5_{875}$	0.48	0.02	0.000	0.48	0.00	0.000
$E18_{450}$	0.41	0.04	0.000	0.41	0.00	0.000
$E18_{1000}$	0.50	0.78	0.007	0.50	0.93	0.000
$B12_{1250}$	0.58	0.85	0.003	0.58	0.96	0.000
$C8_{1050}$	0.77	0.96	0.003	0.77	1.00	0.010
$C8_{2100}$	0.88	0.18	0.000	0.88	0.08	0.000
$I11_{900}$	0.67	0.88	0.009	0.67	0.96	0.011
$I11_{1050}$	0.19	0.05	0.023	–	–	–
$K7_{300}$	0.68	0.35	0.011	–	–	–
$K7_{500}$	–	–	–	0.11	0.54	0.001
$K7_{575}$	0.50	0.86	0.000	0.50	0.96	0.000
$K8_{275}$	0.59	0.03	0.000	0.59	0.00	0.000
$K8_{2200}$	0.26	0.08	0.013	–	–	–
$C1_{725}$	0.35	0.87	0.000	0.35	1.00	0.000
$C1_{800}$	0.54	0.80	0.010	–	–	–
$C1_{1200}$	0.55	0.87	0.000	0.55	0.96	0.001
$E19_{750}$	0.96	0.76	0.018	–	–	–
$E19_{1200}$	0.52	0.89	0.000	0.52	1.00	0.000
$F9_{800}$	–	–	–	0.44	0.11	0.014
$J19_{475}$	0.41	0.78	0.000	0.41	0.92	0.000
$J10_{600}$	0.50	0.15	0.000	0.50	0.11	0.003
$J10_{1100}$	–	–	–	0.20	0.00	0.020
$J10_{1250}$	0.04	0.25	0.011	0.04	0.52	0.000
$A15_{1300}$	–	–	–	0.15	0.48	0.018
$B6_{550}$	0.56	0.07	0.000	0.56	0.00	0.000
$B6_{1250}$	0.30	0.65	0.001	0.30	0.93	0.000
$B6_{1500}$	0.56	0.85	0.001	0.56	0.96	0.000
$C9_{600}$	0.26	0.09	0.018	0.26	0.00	0.010
$C9_{1200}$	0.85	0.50	0.000	0.85	0.52	0.018
$C13_{600}$	0.96	0.55	0.000	0.96	0.50	0.000
$F16_{1350}$	–	–	–	0.15	0.56	0.004
$L5_{700}$	0.92	0.36	0.000	0.92	0.19	0.000
$A7_{500}$	0.72	0.96	0.000	0.72	1.00	0.004
$A17_{700}$	0.25	0.59	0.003	0.25	0.64	0.010
$A17_{1550}$	0.93	0.46	0.000	–	–	–
$A17_{2000}$	–	–	–	0.11	0.48	0.006
$M17_{650}$	0.37	0.72	0.000	0.37	0.88	0.000
$B15_{700}$	0.39	0.68	0.011	–	–	–
$B15_{900}$	0.52	0.81	0.004	0.52	0.96	0.000
$B15_{1200}$	–	–	–	0.11	0.48	0.006
$E11_{1050}$	0.48	0.88	0.000	0.48	1.00	0.000
$M10_{1000}$	0.56	0.02	0.000	0.56	0.04	0.000
$M10_{1250}$	0.56	0.95	0.000	0.56	1.00	0.000
$P14_{475}$	–	–	–	0.15	0.48	0.018
$P14_{1150}$	0.41	0.05	0.000	0.41	0.04	0.002
$Q4_{925}$	0.27	0.07	0.005	0.27	0.04	0.024
$R19_{525}$	0.58	0.88	0.000	0.58	0.96	0.002
$R19_{700}$	0.41	0.83	0.000	0.41	0.89	0.000
$R19_{1100}$	0.33	0.09	0.002	–	–	–
$AP2_{600}$	0.62	0.93	0.000	0.62	1.00	0.000
$AP2_{1725}$	0.74	0.39	0.001	–	–	–
$AP2_{1775}$	0.67	0.41	0.020	–	–	–
$Z6_{750}$	0.04	0.24	0.018	–	–	–
$AW18_{500}$	–	–	–	0.15	0.48	0.018
$AW18_{775}$	0.41	0.04	0.000	0.41	0.00	0.000
$AW18_{1050}$	0.44	0.19	0.006	–	–	–
$AJ2_{1075}$	0.52	0.15	0.000	0.52	0.00	0.000
$AJ2_{1475}$	0.67	0.39	0.011	–	–	–
$AX15_{975}$	0.64	0.97	0.000	0.64	1.00	0.000
$AX15_{1000}$	0.46	0.13	0.000	0.46	0.04	0.000
$AP4_{475}$	–	–	–	0.58	0.19	0.005
$AP4_{1800}$	0.04	0.25	0.011	0.04	0.50	0.000

† RAPD marker size in subscripts estimated to the nearest 25 base pairs.

‡ Rapid marker frequency = $\Sigma(p_i)/n_i$ where p is the presence of a RAPD band among accessions and n is the number of accessions evaluated in a sub-population for the ith RAPD marker.

§ P-value computed by Fishers' exact test (two tailed). *P*-values ≤ 0.025 and ≤ 0.005 are significant at the 5% and 1% levels, respectively.

than 272 rivers and lakes were represented as sites of origin in the collection. This does not include any material designated CEPEC, SIC, or SIAL consisting of hybrids between accessions with different countries of origin. Considering the immense area the Amazon River flows through, major areas throughout the states of Para, Roraima, Rondônia, and Amazonas in Brazil were underrepresented. This reaffirms our earlier results that

sampling within areas already collected did not increase the genetic diversity, and future cacao collection expeditions should be supported to investigate nonrepresented areas to expand the genetic diversity that currently exists in cacao germplasm collections.

Witches' Broom Disease Resistance

One hundred eighty accessions (Accessions 1–180) of the 270 analyzed were characterized for resistance to witches' broom. The number of affected vegetative brooms (VB) and cushion brooms (CB) were counted and averaged for 10 clones of each accession examined (data not shown). The average number of vegetative brooms exhibiting infection was 19.9, whereas the average number of cushion brooms was 20.6. The correlation between VB and CB data was 0.83. Histograms of the number of affected VB and CB showed very similar distributions (Marita, 1998). The arbitrary cut-off for accessions with the most disease tolerance was the lower 15% of the distribution for total infected brooms. A MDS plot comparing accessions with infected VB and accessions with infected CB within the lower 15% of each respective distribution (Fig. 2) demonstrated that approximately 89% of accessions represented in each category are the same. These results suggest that analyses of either the VB or CB data would be equally applicable.

Significant differences in RAPD marker frequency existed between the lower 15% of the distribution for total infected brooms (accessions most tolerant to witches' broom disease) and the rest of the accessions analyzed in the witches' broom trial. Similarly, significant differences in marker frequency existed between the lower 15% and the upper 15% of the distribution (accessions least tolerant to witches' broom disease) for total infected brooms. Highly significant differences were confirmed between the marker frequencies in each comparison (Table 4). These results along with the MDS plot supported the clustering of genetic distance differences between the accessions that fall in the lower 15% of the total infected brooms distribution and the rest of the accessions analyzed.

Individual marker frequency differences were observed between tolerant accessions and (i) least tolerant accessions and (ii) remaining accessions (Table 5). Of a total of 133 RAPD markers, 52 RAPD markers (39%) were significant between groups of accessions in the first comparison and 49 RAPD markers (37%) were significant between groups of accessions in the second comparison, with 39 significant RAPD markers identical between the two sets of comparisons. Few of the significant markers listed in Table 5 have been mapped on the basis of previous studies, and those significant markers that were selected from previous studies do not cluster in the cocoa genome. Unique markers associated with genes affecting witches' broom disease resistance could be tested through segregation analysis with inoculation and careful disease evaluation. By identifying specific markers, a few PCR reactions on potentially thousands of accessions worldwide could be used to identify accessions with a high probability of having resistance or tolerance genes. Such studies seem worthwhile to pursue considering the eventual cost savings.

Resistance to Witches' Broom Disease with Respect to Cacao Germplasm

Many accessions exhibiting tolerance to witches' broom disease were from the Upper Amazon, such as SCA 6, SCA 12, C SUL 3, C SUL 4, C SUL 7, and CCN 10, or were hybrids with Upper Amazon accessions having Scavina germplasm in their pedigrees, specifically SCA 6 and SCA 12. The clustering of tolerance to witches' broom disease with accessions from the Upper Amazon region was not definitive and could be biased. Many factors influenced the results, including nonuniformity in age, inclusion of hybrids with similar parentage, and accession selection based on resistance to witches' broom disease. Two series, C SUL and SCA, represented 67% of the accessions considered tolerant to witches' broom disease. This was not a surprising result since resistance to witches' broom disease is thought to come from Upper Amazon derived accessions and the first 180 accessions analyzed were selected for resistance to witches' broom disease.

Tolerance to witches' broom disease has been identified in relatively few sources of cacao germplasm, with Scavina 6 and Scavina 12 being two of the most important. The analyses of the witches' broom resistance data identified accessions closely related to tolerant accessions, which potentially represent new sources of disease resistance. Many of the C. SUL accessions exhibited a high threshold of tolerance to the disease, comparable to Scavina accessions and hybrids with Scavina backgrounds. SGU 26, a hybrid from Guatemala, was the only accession among accessions identified as most tolerant to witches' broom disease that was not classified definitively as an Upper Amazon accession as indicated on the MDS plot (Fig. 1). These results are very important for cacao collection, suggesting the Upper Amazon is not the only region to have genes for resistance to witches' broom disease. Future disease screening programs should focus in part on clones collected from regions around the SCA and C SUL collection sites and careful attention should be given to examining SGU 26 as a new source of resistance to witches' broom disease. From a breeding perspective, attention should be given to evaluating these accessions and incorporating these accessions into resistant populations.

ACKNOWLEDGMENTS

This research was financially supported by Fazenda Almirante, a division of Mars, M&M Inc. in Itajuipe, Bahia, Brazil.

REFERENCES

Baker, R.E.D. 1953. Anglo-American cacao collecting expedition. Archives of Cocoa Res. 1:127–154.

Bartley, B.G.D. 1969. Twenty years of cacao breeding at the Imperial College of Tropical Agriculture. p. 29–34. Trinidad Int. Cacao Res. Conf. Mem. 1967. CEPLAC, Ilhens-Itabuna, Bahia, Brazil.

Cheesman, E.E. 1944. Notes on the nomenclature, classification and possible relationships of cocoa populations. Trop. Agric. 2:144–159.

Cope, F.W. 1976. *Theobroma cacao* (Sterculiacae). p. 285–289. *In* N.W. Simmonds (ed.) Evolution of crop plants. Longman, New York.
Cronshaw, D.K., and H.C. Evans. 1978. Witches' broom disease of cocoa (*Crinipellis perniciosa*) in Ecuador: II. Methods of infection. Ann. Appl. Biol. 89:193–200.
Crouzillat, D., E. Lerceteau, V. Petiard, J. Morera, H. Rodriguez, D. Walker, W. Phillips, C. Ronning, R. Schnell, J. Osei, and P. Fritz. 1996. *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. Theor. Appl. Genet. 93:205–214.
Doyle, J.J., and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13–15.
Edgington, E.S. 1980. Randomization tests. Marcel Dekker Inc., New York.
End, M.J., R.M. Wadsworth, and P. Hadley. 1992. International cocoa germplasm database. The Biscuit, Cake, Chocolate, and Confectionary Alliance, London, England.
Figueira, A., J. Janick, M. Levy, and P. Goldsbrough. 1994. Reexamining the classification of *Theobroma cacao* L. using molecular markers. J. Am. Soc. Hort. Sci. 119:1073–1082.
Frias, G.A., L.H. Purdy, and R.A. Schmidt. 1991. Infection biology of *Crinipellis perniciosa* on vegetative flushes of cacao. Plant Dis. 75:552–556.
Gower, J.C. 1985. Measures of similarity, dissimilarity, and distance. p. 297–405. *In* S. Kotz and N.L. Johnson (ed.) Encyclopedia of statistical sciences. Volume 5. Wiley, New York.
Johns, M.A., P.W. Skroch, J. Nienhuis, P. Hinrichsen, G. Bascur, and C. Munoz-Schick. 1997. Gene pool classification of common bean landraces from Chile based on RAPD and morphological data. Crop Sci. 37:605–613.
Kennedy, A.J., G. Lockwood, G. Mossu, N.W. Simmonds, and G.Y. Tan. 1987. Cocoa breeding: past, present and future. Cocoa Growers' Bull. 38:5–22.
Lanaud, C., A.M. Risterucci, A.K.J. N'Goran, D. Clement, M.H. Flament, V. Laurent, and M. Falque. 1995. A genetic linkage map of *Theobroma cacao* L. Theor. Appl. Genet. 91:987–993.
Laurent, V., A.M. Risterucci, and C. Lanaud. 1994. Genetic diversity in cocoa revealed by cDNA probes. Theor. Appl. Genet. 88:193–198.
Lerceteau, E., R. Robert, V. Pétiard, and D. Crouzillat. 1997. Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. Theor. Appl. Genet. 95:10–19.
Lockwood, G. 1976. A comparison of the growth and yield during a 20 year period of Amelonado and Upper Amazon hybrid cocoa (*Theobroma cacao*) in Ghana. Euphytica 24(3):647–658.
Marita, J.M. 1998. Characterization of *Theobroma cacao* using RAPD-marker based estimates of genetic distance and recommendations for a core collection to maximize genetic diversity. MS Thesis, Univ. of Wisconsin, Madison, WI.
Marita, J.M., J.M. Rodriguez, and J. Nienhuis. 2000. Development of an algorithm identifying maximally diverse core collections. Genet. Res. Crop Evol. 47(5):515–526.
N'Goran, J.A.K., V. Laurent, A.M. Risterucci, and C. Lanaud. 1994. Comparative genetic diversity studies of *Theobroma cacao* L. using RFLP and RAPD markers. Heredity 73:589–597.
Nei, M. 1987. Genetic variation within species. p. 176–207. *In* Molecular evolutionary genetics. Columbia University Press, New York.
Pereira, J.L., A. Ram, J.M. de Figueiredo, and L.C.C. de Almeida. 1990. First occurrence of witches' broom disease in the principal cocoa-growing region of Brazil. Trop. Agric. 67(2):188–189.
Ronning, C.M., R.J. Schnell, and D.N. Kuhn. 1995. Inheritance of random amplified polymorphic DNA (RAPD) markers in *Theobroma cacao* L. J. Am. Soc. Hort. Sci. 120(4):681–686.
SAS Institute. 1990. SAS/STAT User's Guide, ver. 6. 4th ed. SAS Inst., Cary, NC.
Skroch, P.W., and J. Nienhuis. 1995a. Qualitative and quantitative characterization of RAPD variation among snap bean (*Phaseolus vulgaris*) genotypes. Theor. Appl. Genet. 91:1078–1085.
Skroch, P.W., and J. Nienhuis. 1995b. Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. Theor. Appl. Genet. 91:1086–1091.
Snedecor, G.W., and W.G. Cochran. 1992. Statistical Methods. 8th ed. Iowa State University Press, Ames, IA.
Soria, J.V. 1970. Principal varieties of cocoa cultivated in tropical america. Cocoa Growers' Bull. 19:12–21.
Sreenivasan, T.N., and S. Dabydeen. 1989. Modes of penetration of young cocoa leaves by *Crinipellis perniciosa*. Plant Dis. 73:478–481.
Stahel, G. 1915. *Marasmius perniciosus* nov spec. Bull. Dep. Landbouw Suriname 33:1–26.
Steel, R.G.D., and J.H. Torrie. 1980. Principles and procedures of statistics: a biometrical approach. McGraw-Hill Inc., New York.
Went, F.A.F.C. 1904. Krulloten en versteende vruchten van de cacao in Suriname, Verhandelingen der K akademie van wetenschappen. Amsterdam 2(10):1–40.
Whitkus, R., M. de la Cruz, L. Mota-Bravo, and A. Gómez-Pompa. 1998. Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. Theor. Appl. Genet. 96:621–627.
Wilde, J., R. Waugh, and W. Powell. 1992. Genetic fingerprinting of Theobroma clones using randomly amplified polymorphic DNA markers. Theor. Appl. Genet. 83:871–877.
Wilkinson, L., M. Hill, S. Miceli, P. Howe, and E. Vang. 1992. SYSTAT for the Macintosh, ver. 5.2. SYSTAT Inc., Evanstan, IL.
Williams, J.T. 1984. A decade of crop genetic resources research. *In* J.H.W. Holden and J.T. Williams (ed.) Crop genetic resources, conservation and evaluation. Allen and Unwin, London, England.
Wood, G.A.R. and R.A. Lass. 1985. Cocoa. 4th ed. Longman Group Limited, New York.