Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*)[1]

Iris E. Peralta[2] and David M. Spooner[3]

Vegetable Crops Research Unit, USDA, Agricultural Research Service, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, Wisconsin 53706-1590 USA

Eight wild tomato species are native to western South America and one to the Galapagos Islands. Different classifications of tomatoes have been based on morphological or biological criteria. Our primary goal was to examine the phylogenetic relationships of all nine wild tomato species and closely related outgroups, with a concentration on the most widespread and variable tomato species *Solanum peruvianum,* using DNA sequences of the structural gene granule-bound starch synthase (GBSSI, or *waxy*). Results show some concordance with previous morphology-based classifications and new relationships. The ingroup comprised a basal polytomy composed of the self-incompatible green-fruited species *S. chilense* and the central to southern Peruvian populations of *S. peruvianum, S. habrochaites,* and *S. pennellii.* A derived clade contains the northern Peruvian populations of *S. peruvianum* (also self-incompatible, green-fruited), *S. chmielewskii,* and *S. neorickii* (self-compatible, green-fruited), and the self-compatible and red- to orange- to yellow-fruited species *S. cheesmaniae, S. lycopersicum,* and *S. pimpinellifolium.* Outgroup relationships are largely concordant with prior chloroplast DNA restriction site phylogenies, support *S. juglandifolium* and *S. ochranthum* as the closest outgroup to tomatoes with *S. lycopersicoides* and *S. sitiens* as basal to these, and support allogamy, self-incompatibility, and green fruits as primitive in the tomato clade.

Key words: GBSSI; *Lycopersicon*; phylogeny; Solanaceae; *Solanum* sect. *Lycopersicon*; tomato; *waxy* gene.

Wild tomatoes (*Solanum* L. subsect. *Lycopersicon,* an autonym of *Solanum* section *Lycopersicon* (Mill.) Wettst.) are entirely American in distribution, growing in the western South American Andes from central Ecuador through Peru to northern Chile and in the Galapagos Islands, where the endemic species *S. cheesmaniae* grows (Fig. 1; see Table 1 for taxonomic authorities). In addition, *S. lycopersicum,* the wild ancestor of cultivated tomatoes, is more widespread and perhaps more recently distributed into Mexico, Colombia, Bolivia, and other South American countries (Rick and Holle, 1990). Wild tomatoes grow in a variety of habitats, from near sea level to >3300 m in elevation (Rick, 1973; Taylor, 1986). These habitats include the arid Pacific coast to the mesic uplands in the high Andes. All species are diploid (2*n* = 24;

Rick, 1979). Breeding systems vary from allogamous self-incompatible to facultative allogamous and self-compatible to autogamous and self-compatible (Rick, 1963, 1979, 1984, 1986; Table 1). The self-incompatibility system in tomatoes is gametophytic and controlled by a single, multiallelic *S* locus (Tanksley and Loaiza-Figueroa, 1985).

Linnaeus (1753) first recognized tomatoes in the genus *Solanum,* but Miller (1754) recognized tomatoes in the genus *Lycopersicon* Miller. Miller's recognition of *Lycopersicon* continues to be used by the majority of botanists and plant breeders today. However, a minority of authors (MacBride, 1962; Seithe, 1962; Heine, 1976; Fosberg, 1987; Child, 1990; Spooner, Anderson, and Jansen, 1993; Bohs and Olmstead, 1999; Knapp and Spooner, 1999; Olmstead et al., 1999) treat tomatoes in *Solanum* (Peralta and Spooner, 2000). Molecular data from chloroplast DNA restriction site and sequence data (Spooner, Anderson, and Jansen, 1993; Bohs and Olmstead, 1997, 1999; Olmstead and Palmer, 1997; Olmstead et al., 1999) firmly place tomatoes and potatoes as sister taxa, justifying this treatment on phylogenetic grounds. The continued maintenance of *Lycopersicon* can be justified only on convenience and the maintenance of nomenclatural stability. Börner (1912) also recognized the close affinity between tomatoes and potatoes and proposed a new genus *Solanopsis* to include them. D'Arcy (1987) and Lester (1991) also recognized the close relationship of the genera *Solanum* and *Lycopersicon* but maintained them as distinct for nomenclatural stability.

The two most recent and complete taxonomic treatments of tomatoes are those of Müller (1940) and Luckwill (1943; Fig. 2). Müller (1940) considered *Lycopersicon* as a distinct genus and divided it into two subgenera, *Eulycopersicon* C. H. Müll., with two species possessing glabrous and red- to orange-colored fruits, flat seeds, bractless inflorescences, and leaves without pseudostipules, and *Eriopersicon* C. H. Müll., with four
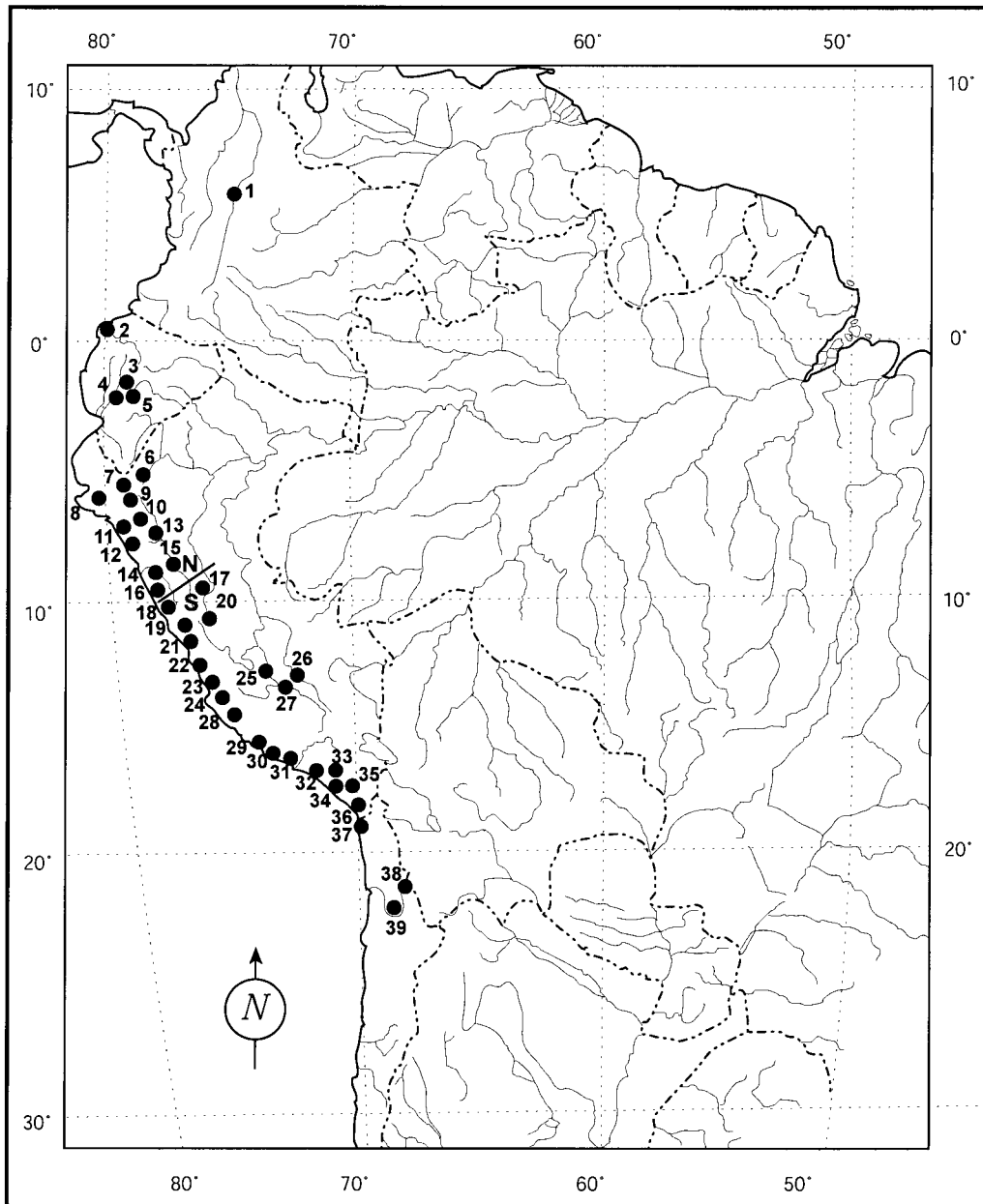
Fig. 1. Distribution of the wild tomato accessions used in this study. Map numbers correspond to generalized map localities in Table 1. A line drawn at ~10° S indicates the northern (N) and southern (S) edges of distribution of "north" and "south" *Solanum peruvianum* populations that reflect the phylogenetic relationships showed in Fig. 5.

species bearing pubescent green fruits, thick seeds, bracteate inflorescences, and leaves usually with pseudostipules. Three years later, Luckwill (1943) adopted the same supraspecific categories but recognized different infraspecific taxa and five species in the subgenus *Eriopersicon* (Fig. 2). These treatments have become outdated as the number of species and races collected from South America has increased (Rick, 1971, 1991; Holle, Rick, and Hunt, 1978, 1979; Taylor, 1986).

Rick (1960, 1979) proposed an infrageneric classification based mainly on crossing relationships. He recognized nine wild tomato species, classified into two complexes. The Esculentum complex includes seven species, mainly self-compatible, and easily crossed with the cultivated tomato. Within

this group three species, *L. esculentum, L. pimpinellifolium,* and *L. cheesmaniae,* have mostly glabrous, pigmented fruits, while the others have pubescent, green fruits. The Peruvianum complex includes the self-incompatible species *S. peruvianum* and *S. chilense,* which have pubescent green fruits that seldom cross with *L. esculentum.*

In a recent taxonomic treatment, Child (1990; Fig. 2) placed the tomatoes in *Solanum* subgenus *Potatoe* (G. Don) D'Arcy, section *Lycopersicon,* subsection *Lycopersicon,* and segregated them into three series: *Eriopersicon* (C. H. Müll.) Child, *Lycopersicon,* and *Neolycopersicon* (Correll) Child. He also hypothesized that *Solanum* subsection *Lycopersicoides* Child and section *Juglandifolium* (Rydb.) Child are the closest relatives

TABLE 1. Accessions of wild tomatoes (*Solanum* sect. *Lycopersicon* [Mill.] Wettst. Subsection *Lycopersicon*) and outgroups examined for morphological and GBSSI gene sequence variation. Vouchers are deposited at BM, MERL, PTIS, and WIS. We used the *Solanum* equivalents of the *Lycopersicon* species as recognized by Rick, Laterrot, and Philouze (1990), with the addition of the varieties and forms of *S. cheesmaniae*, *S. habrochaites*, *S. pennellii*, and *S. peruvianum* as listed in the database of germplasm at the C. M. Rick Tomato Genetics Resource Center (http://tgrc.ucdavis.edu).

Taxon[a]	Map locality[b]	Data set[c]	Accession number[d]	Collector[e]	Breeding system[f]	Locality[g]
Ingroup, *Solanum* sect. *Lycopersicon* subsection *Lycopersicon*						
chs	G	M, W	LA166	Unknown	Aut-SC	Ecuador. Galapagos Islands: Santa Cruz, Barranco, N of Punta Ayora, 1 km N of Punta Ayora, 50 m.
chs	G	M, W	LA1450	Unknown	Aut-SC	Ecuador. Galapagos Islands: Isabela, Bahia San Pedro, 1/2 km from coast, 15 m.
chsm	G	M, W	LA317	Unknown	Aut-SC	Ecuador. Galapagos Islands: Bartolome, 15 m.
chl	29	M, W	LA1930	*Rick et al. 3172*	Allo-SI	Peru. Arequipa: Quebrada Calapampa, Río Acarí, 500 m, 15°39′ S, 74°39′ W.
chl	36	M, W	LA1963	*Rick et al. 3204*	Allo-SI	Peru. Tacna: Río Caplina, 13 km W of Tacna, 3 km from Panamerican Hwy, toward Boca del Río, 200 m, 18°14′ S, 70°33′ W.
chl	38	M, W	LA2884	*Rick et al. 7609*	Allo-SI	Chile. Antofagasta: Ayaviri, 22°02′ S, 68°07′ W.
chm	25	M, W	LA1306 (PI 379029)	*Rick et al. 390*	Fac-SC	Peru. Ayacucho: Tambo, 12°56′ S, 74°01′ W, 1 km E of Tambo, 3100 m.
chm	27	M, W	LA1327	*Rick et al. 411*	Fac-SC	Peru. Apurímac: Soracata, 1980 m.
chm	26	M, W	LA2663	*Rick et al. 6101*	Fac-SC	Peru. Cusco: Tujtohaiya (Upper Apurimac), 2500 m.
lyp	S	W	Cultivar 'Mountain Spring'		Aut-SC	Cultivated, obtained from Tomato Growers Supply Company, Fort Meyer's Florida, USA.
lypc	5	M, W	LA1226 (PI 379045)	*Rick et al. 390*	Aut-SC	Ecuador. Morona-Santiago: 3 km S of Sucua, 1000 m.
lypc	21	M, W	LA1673	*Rick et al. 1985*	Aut-SC	Peru. Lima: Ñaña, 11°59′ S, 76°50′ W, along W wall of the Estancia.
lypc	32	M, W	LA1334 (PI 379005)	*Rick et al. 418*	Aut-SC	Peru. Arequipa; Pescaderos, 50 m.
hab	11	M, W	LA1353 (PI 365934)	*Rick et al. 437*	Allo-SI	Peru. Cajamarca: Contumazá, 2650 m, 7°22′ S, 78°49′ W.
hab	24	M, W	LA1928	*Rick et al. 3170*	Fac-SC	Peru. Ica: Ocaña, Río Ingenio, 76 km from Panamerican Hwy, 2660 m, 14°39′ S, 75°15′ W.
habg	3	M, W	LA1223 (PI 365903)	*Rick et al. 301*	Fac-SC	Ecuador. Chimborazo: Alausí, 2°12′ S, 78°50′ W, 2200 m.
neo	7	M, W	LA2200	*Rick et al. 5045*	Aut-SC	Peru. Amazonas: Choipiaco, Río Suche, 6°13′ S, 77°51′ W, 16 km E of Chachapoyas, 1900 m.
neo	17	M, W	LA247	*Ochoa 1071*	Aut-SC	Peru. Huánuco: Chavinillo, 9°47′ S, 76°35′ W.
neo	20	M, W	LA1326 (PI 379033)	*Rick et al. 410*	Aut-SC	Peru. Apurímac: Río Pachachaca, 1980 m, 11°39′ S, 76°01′ W.
pen	19	M, W	LA1376 (PI 379017)	*Rick et al. 460*	Allo-SI	Peru. Lima: Sayán, 1000 m, Agua perdida (Río Ingenio), 11°06′ S, 77°39′ W.
pen	31	M, W	LA716 (PI 246502)	Unknown	Fac-SC	Peru. Arequipa: Atico, Near Atico from Chala along Panamerican Hwy, 16°14′ S, 73°39′ W.
penp	28	M, W	LA1926	*Rick et al. 3168*	Allo-SI	Peru. Ica: 25 km from Panamerican Hwy, 1200 m, 14°38′ S, 75°04′ W.
per	6	M, W	LA2185	*Rick et al. 5030*	Allo-SI	Peru. Amazonas: Pongo de Rentema, in the gap of the Marañon, E of Coral Quemado, 400 m, 5°26′ S, 79°39′ W.
per	9	M, W	LA2172	*Rick et al. 5017*	Allo-SI; possibly SC or segregating SI/SC	Peru. Cajamarca: Cuyca, 161 km from Olmos junction with the Panamerican Hwy, 11°31′ S, 75°07′ W.
per	9	M, W	LA2163	*Rick et al. 5008*	Allo-SI	Peru. Cajamarca: between Cochabamda and Yamaluc, 2–5 km N of Cochabamda, 1800–1900 m.

TABLE 1.   Continued.

Taxon[a]	Map locality[b]	Data set[c]	Accession number[d]	Collector[e]	Breeding system[f]	Locality[g]
per	9	M, W	LA2157	*Rick et al. 5002*	Fac-SC	Peru. Cajamarca: Tunel Chotano, 23 km N of Chota, 1600 m.
per	11	M, W	LA392	*Rick et al. 141*	Allo-SI	Peru. Cajamarca: Llallan, 78 km from Car. Panamerican Hwy, 12 km W of Chile, 900 m.
per	11	W	LA1351 (PI 365951)	*Rick et al. 435*	Allo-SI	Peru. Cajamarca: Rupe, 1800–2100 m.
perh	11	M, W	LA2152	*Rick et al. 4997*	Allo-SI	Peru. Cajamarca: Río Jequetepeque, 140 km from Panamerican Hwy, 2 km W of San Juan, 2150 m, 7°17′ S, 78°29′ W.
perh	11	M, W	LA2548	PE-29	Allo-SI	Peru. Cajamarca: La Muyuna, Río Jeque-tepeque, 870 m.
per	10	M, W	LA1396	*Rick et al. 480*	Allo-SI	Peru. Amazonas: Balsas, 4°25′ S, 80°17′ W.
per	13	M	LA2553	PE-31	Allo-SI	Peru. Cajamarca: Balconcillo-San Mar-cos, 2650 m, 7°20′ S, 78°11′ W.
per	12	M, W	LA441	*Rick et al. 128*	Allo-SI	Peru. La Libertad: Cerro Campana, W of Panamerican Hwy N of Trujillo, 350 m, 5°05′ S, 79°17′ W.
per	12	W	LA378	*Rick et al. 125*	Allo-SI	Peru. Cajamarca: Cascas, 1200 m, 7°29′ S, 78°49′ W.
per	12	M, W	LA1984	*Rick et al. 3225*	Allo-SI	Peru. La Libertad: Otuzco (Río Moche), 2 km on side road E of Otuzco, 2800 m, 7°54′ S, 78°35′ W.
per	13	M, W	LA2328	*Rick et al. 5172*	Allo-SI	Peru. La Libertad: Aricapampa, 2300–2500 m, 7°48′ S, 77°43′ W.
per	14	M	LA1982	*Rick et al. 3223*	Allo-SI	Peru. Ancash: Huallanca, Río Santa, near the hydroelectric plant, ~4 km beyond LA1981, 1400 m, 8°49′ S, 77°52′ W.
per	15	M, W	LA1626	Unknown	Allo-SI	Peru. Ancash: mouth of Río Rupac, 1700 m.
per	16	M	LA1360 (PI 365952)	*Rick et al. 444*	Allo-SI	Peru. Ancash: Pariacoto, 1490 m, 9°31′ S, 77°53′ W.
per	18	W	LA153 (PI 126944)	Unknown	Allo-SI	Peru. Ancash: Culebras, 60–90 m, 9°55′ S, 78°12′ W.
per	18	M, W	LA111	*Rick et al. 41*	Allo-SI	Peru. Lima: Supe, along Panamerican Hwy at km 158, 50 m, 10°48′ S, 77°44′ W.
per	18	M, W	LA110	*Rick et al. 39*	Allo-SI	Peru. Ancash: Cajacay, 5 km W of Caja-cay, 2500 m, 10°10′ S, 77°26′ W.
per	18	M	LA1365 (PI 365953)	*Rick et al. 449*	Allo-SI	Peru. Ancash: Caranquilloc, 2450 m.
per	19	M, W	LA1379 (PI 379018)	*Rick et al. 463*	Allo-SI	Peru. Lima: Caujul, 1500 m, 10°48′ S, 77°02′ W.
per	21	M, W	LA1274 (PI 365940)	*Rick et al. 356*	Allo-SI	Peru. Lima: Pacaibamba, 1440 m.
per	21	M, W	LA370	*Rick et al. 115*	Allo-SI	Peru. Lima: Huampani Quebrada, 200 m.
per	21	W	LA1556	Unknown	Allo-SI	Peru. Lima: Rímac Valley, Hacienda Higuereto, Río Surco, 250 m.
perg	21	M	LA364	*Rick et al. 109*	Allo-SI	Peru. Lima: 9 km W of Canta, 2100 m, 12°35′ S, 69°09′ W.
perg	21	M, W	LA1283 (PI 365942)	*Rick et al. 365*	Allo-SI	Peru. Lima: Santa Cruz de Laya, 2000 m.
per	21	M, W	LA103	*Rick et al. 28*	Allo-SI	Peru. Lima: Cajamarquilla, On both sides of dry wash, 100 m, 6°25′ S, 79°14′ W.
per	22	W	LA1609	*Rick et al. 1889*	Allo-SI	Peru. Lima: between Asia and El Pinon, 9 km E of Panamerican Hwy, 60 m, 12°47′ S, 76°30′ W.
per	22	M, W	LA107	*Rick et al. 33*	Allo-SI	Peru. Lima: Hacienda San Isidro, 5°33′ S, 80°49′ W.
per	23	M	LA1339 (PI 365949)	*Rick et al. 423*	Allo-SI	Peru. Lima: Capillucas, 1700 m, 13°39′ S, 12°45′ W.
per	23	M, W	LA444	*Rick et al. 198*	Allo-SI	Peru. Ica: Chincha, southern outskirts, 100 m, 17°20′ S, 71°00′ W.
per	24	M, W	LA1305 (PI 379015)	*Rick et al. 389*	Allo-SI	Peru. Huancavelica: 20 km E of Ticrapo, 2960 m.
per	24	M, W	LA1910	*Rick et al. 3152*	Allo-SI	Peru. Huancavelica: Ticrapo, Tambillo, (Río Ica), 74 km from Ica, 2300 m.

TABLE 1.   Continued.

Taxon[a]	Map locality[b]	Data set[c]	Accession number[d]	Collector[e]	Breeding system[f]	Locality[g]
per	28	M, W	LA1913	*Rick et al. 3155*	Allo-SI	Peru. Ica: Tinguiayog, (Río Santa Cruz), 15 km from Panamerican Hwy, 900 m.
per	28	M, W	LA1331 (PI 365946)	*Rick et al. 415*	Allo-SI	Peru. Ica: Nazca, 1500 m, 14°50′ S, 74°57′ W.
per	29	M, W	LA1937	*Rick et al. 3179*	Allo-SI	Peru. Arequipa: Quebrada Torrecillas, (Río Chaparra), 21 km from Chaparra, 64 km from Panamerican Hwy, 2500 m, 15°55′ S, 74°07′ W.
per	30	M, W	LA448	*Rick et al. 202*	Allo-SI	Peru. Arequipa: Chala, 15°52′ S, 74°16′ W.
per	30	W	LA1336 (PI 365947)	*Rick et al. 420*	Allo-SI	Peru. Arequipa: Atico, Near Atico from Chala along Car. Panamerican, 10 m, 16°14′ S, 73°39′ W.
per	30	M	LA1945	*Rick et al. 3186*	Allo-SI	Peru. Arequipa: Caravelí, 71 km from Panamerican Hwy, 8 km before reaching Caravelí, 1940 m, 15°46′ S, 73°22′ W.
per	32	M	LA1474	Unknown	Allo-SI	Peru. Arequipa: Lomas de Camaná, 1300 m, 16°22′ S, 73°01′ W.
per	33	M, W	LA1973	Unknown	Allo-SI	Peru. Arequipa: Yura, 16°12′ S, 71°42′ W.
per	33	W	LA451	*Rick et al. 205*	Allo-SI	Peru. Arequipa: 18 km from Arequipa, 3200 m, 16°24′ S, 71°33′ W.
per	35	M, W	LA3156	Unknown	Allo-SI	Peru. Moquegua: Omate Agricultural Valley.
per	34	M, W	LA1954	*Rick et al. 3195*	Allo-SI	Peru. Arequipa: 2–4 km W of Mollendo, 50 m, 17°02′ S, 72°01′ W.
per	34	M	LA454	*Rick et al. 208*	Allo-SI	Peru. Arequipa: 8 km S of Río Tambo crossing, 1200 m.
per	36	M, W	LA2964	Unknown	Allo-SI	Peru. Tacna: Quebrada de Los Burros.
per	37	M, W	LA2744	*Rick et al. 7502*	Allo-SI	Chile. Tarapaca: Sobraya, (Azapa), 31 km from Arica, 400 m.
pim	2	M, W	LA1237 (PI 365910)	*Rick et al. 318*	Aut-SC	Ecuador. Esmeraldas: Atacames, 5 m.
pim	7	W	LA2184	*Rick et al. 5029*	Aut-SC	Peru. Amazonas: Bagua (Casual), 450 m, 5°40′ S, 78°31′ W.
pim	8	M, W	LA1581	*Rick et al. 1861*	Fac-SC	Peru. Lambayeque: Punto Cuatro, 1 km W of Panamerican Hwy.
pim	8	W	LA722	Line 160	Aut-SC	Peru. La Libertad: Trujillo, 8°07′ S, 79°02′ W.
pim	23	M, W	LA1606	*Rick et al. 1886*	Aut-SC	Peru. Ica: Tambo de Mora, Huaucas, 13°28′ S, 76°12′ W.

Outgroup, *Solanum* sect. *Basarthrum* (Bitter) Bitter

mur	S	W		*Anderson 844*[h]	Likely aut-SC	Ecuador

Outgroup, *Solanum* sect. *Etuberosum* (Buk. and Kameraz) Child

etb	S	W	PI 498311	*Contreras 1322*	Aut-SC	Chile. Bio-Bio: Estero Ponce, 615 m, 36°25′ S, 78°58′ W.
pal	S	W	PI 218228	*Erwin Baur Sortiment Genebank #388*	Aut-SC	Unknown (Chile or Argentina)

Outgroup, *Solanum* sect. *Lycopersicon* subsection *Lycopersicoides* Child

lyc	36	W	LA1964	Unknown	Allo-SI	Peru. Tacna.
lyc	36	W	LA1990 PI 255549	Unknown	Allo-SI	Peru. Tacna.
lyc	36	M, W	LA2386	*Ochoa 14248*	Allo-SI	Peru. Tacna: Chupapalca, 3180 m.
sit	39	M, W	LA2876	*Rick et al. 7601*	Allo-SI	Chile. Antofagasta: Chuquicamata.
sit	39	W	LA2878	*Rick et al. 7603*	Allo-SI	Chile. Antofagasta: Mina La Exotica.

Outgroup, *Solanum* sect. *Juglandifolium* (Rydb.) Child

jgl	4	W	LA2134	*Rick et al. 4974*	Allo-SI	Ecuador. Zamora-Chinchipe: Tinajillas, S side of the road, 2200 m.
jgl	1	W	LA2788	*Rick et al. 7546*	Allo-SI	Colombia. Antioquia: Quebrada la Buena.
och	26	W	LA2682	*Rick et al. 6120*	Allo-SI	Peru. Cusco: 3 km below Chinchaypuyo, 2500 m, 13°38′ S, 72°14′ W.
och	27	W	LA3650	*Rick et al. 7710*	Allo-SI	Peru. Apurímac: Choquemaray, on the road to Labrama, 10 km S Matara, 2440 m.

Outgroup, *Solanum* sect. *Petota* Dumort

blb	S	W	PI 347757	Tarn 153	Allo-SI	Mexico. Michoacán, 2060 m, 19°31′ S, 100°15′ W.
jam	S	W	PI 458424	Ugent 10–78	Allo-SI	U.S.A. New Mexico, 2150 m, 33°55′ S, 108°28′ W.

TABLE 1.   Continued.

Taxon[a]	Map locality[b]	Data set[c]	Accession number[d]	Collector[e]	Breeding system[f]	Locality[g]
tbr	S	W		Unknown	Aut-SC	Cultivated potato used to clone the GBSSI gene (Van der Leij et al., 1991)

[a] Species abbreviations (and *Lycopersicon* synonyms) are: blb = *Solanum bulbocastanum* Dunal; chl = *S. chilense* (Dunal in DC) Reiche [*Lycopersicon chilense* Dunal in DC]; chm = *S. chmielewskii* (C. M. Rick, E. Kesickii, J. E. Fobes, and M. Holle) D. M. Spooner, G. J. Anderson, and R. K. Jansen (*L. chmielewskii* C. M. Rick, E. Kesickii, J. E. Fobes, and M. Holle); chs = *S. cheesmaniae* (Riley) Fosb. [*L. cheesmaniae* Riley]; chsm = *S. cheesmaniae* f. *minor*; etb = *S. etuberosum* Lindl.; hab = *S. habrochaites* S. Knapp and D. M. Spooner [*L. hirsutum* Dunal]; habg = *S. habrochaites* f. *glabratum* C. H. Müll; jam = *S. jamesii* Torr.; jgl = *S. juglandifolium* Dunal; lyc = *S. lycopersicoides* Dunal; lyp = S. lycopersicum L.; lypc = *S. lycopersicum* L. var. *cerasiforme* (Dunal) D. M. Spooner, G. J. Anderson, and R. K. Jansen [*L. eculentum* Miller var. *cerasiforme* Dunal]; mur = *S. muricatum* Aiton; neo = *S. neorickii* D. M. Spooner, G. J. Anderson, and R. K. Jansen [*L. paviflorum* C. M. Rick, E. Kesickii, J. E. Fobes, and M. Holle]; och = *S. ochranthum* Dunal; pal = *S. paluste* Poepp. ex Schltdl.; pen = *S. pennellii* Correll [L. pennellii (Correll) D'Arcy]; penp = *S. pennellii* var. *puberulum* Correll; per = *S. peruvianum* L. [*L. peruvianum* (L.) Miller]; perg = *L. peruvianum* f. *glandulosum* C. H. Müll; perh = *L. peruvianum* var. *humifusum* C. H. Müll; pim = *S. pimpinellifolium* B. Juss. In Linnaeus [*L. pimpinellifolium* (B. Juss.) Miller]; sit =*S. sitiens* I. M. Johnst. (*S. rickii* Correll); tbr = *S. tuberosum* L.

[b] Map numbers correspond to Fig. 1; G and S refer to unmapped accessions, G from the Galapagos Islands located ~1000 km west of Ecuador and S to continental areas north and south of this map.

[c] M = morphological study (unpublished data); W = GBSSI gene sequence study. The morphological study accessions are listed here for ease of comparison to this paper (unpublished data).

[d] LA numbers are from the C. M. Rick Tomato Genetics Resource Center; PI numbers are the same United States Plant Introduction Number duplicate accessions held at the United States Department of Agriculture genebank in Geneva, New York, USA (Geneva Plant Genetic Resources Unit, NE-9).

[e] The collector abbreviation in the database of the C. M. Rick Tomato Genetics Resource Center is SAL, an abbreviation used by Charles Rick for his collections alone or with others to indicate "South American *Lycopersicon*."

[f] Allo = allogamous, Aut = autogamous, Fac = facultative allogamous, SC = self-compatible, SI = self-incompatible. Data from Rick (1963, 1979, 1984, 1986) and htp://tgrc.ucdavis.edu.

[g] Longitude and latitude data added by the authors.

[h] Accession provided by Gregory Anderson.

of section *Lycopersicon.* Different criteria used in classification, morphology, and crossability, therefore, have led to different numbers of species, subspecies, and varieties and to conflicting hypotheses of interspecific relationships (Fig. 2).

Molecular data provide numerous and independent characters useful for plant phylogenetic inference (Sytsma, 1990). Single or low-copy nuclear genes are of great interest in molecular systematics because they are potentially most informative for species-level comparisons. Nuclear gene introns, such as in the genes *Phy, Pgi, adh-1,* and *GapA,* evolve faster and have been considered useful for angiosperm phylogenetic studies (Soltis and Soltis, 1998). In the perennial relatives of soybean (*Glycine*) sequences from the histone H3-D intron

provided useful characters for assessing relationships at the species level (Doyle, Kanazin, and Shoemaker, 1996).

The phylogenetic utility of the structural nuclear gene for granule-bound starch synthase (GBSSI or *waxy* gene) was first explored by Mason-Gamer and Kellogg (1996) in grasses. The intron–exon regions revealed an interesting structure with potential utility at different taxonomic levels. The GBSSI introns provided more phylogenetically informative characters, and higher levels of variation, than the ITS region in grasses (Mason-Gamer, Weil, and Kellogg, 1998). A similar sequence variation was found among *Ipomoea* species comparing ITS with three exons (9, 10, 11; Fig. 3) and two introns, comprising in total 651 nucleotide sites, of the 3′ end of GBSSI (Miller,

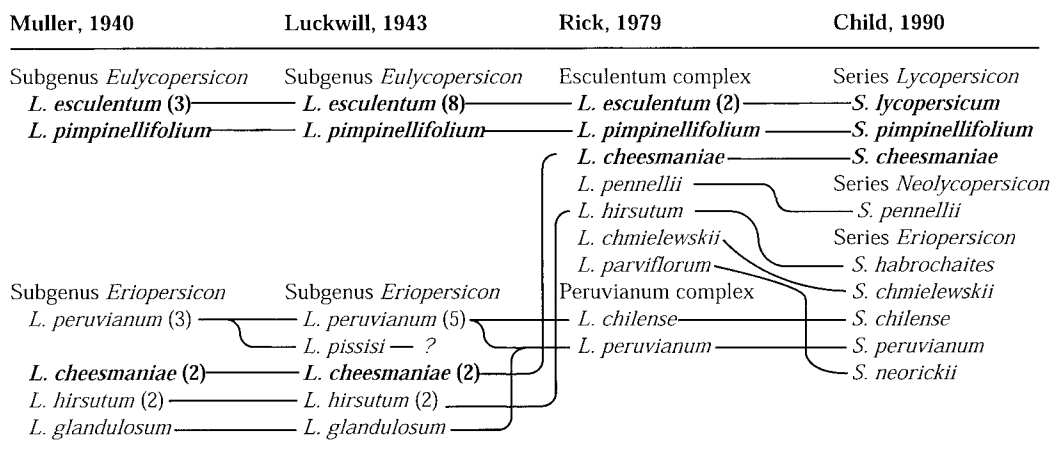| Muller, 1940	Luckwill, 1943	Rick, 1979	Child, 1990
Subgenus *Eulycopersicon*	Subgenus *Eulycopersicon*	Esculentum complex	Series *Lycopersicon*
L. esculentum (3)	**L. esculentum (8)**	L. esculentum (2)	S. lycopersicum
L. pimpinellifolium	**L. pimpinellifolium**	L. pimpinellifolium	S. pimpinellifolium
		L. cheesmaniae	S. cheesmaniae
		L. pennellii	Series *Neolycopersicon*
		L. hirsutum	S. pennellii
		L. chmielewskii	Series *Eriopersicon*
		L. parviflorum	S. habrochaites
Subgenus *Eriopersicon*	Subgenus *Eriopersicon*	Peruvianum complex	S. chmielewskii
L. peruvianum (3)	L. peruvianum (5)	L. chilense	S. chilense
	L. pissisi — ?	L. peruvianum	S. peruvianum
L. cheesmaniae (2)	**L. cheesmaniae (2)**		S. neorickii
L. hirsutum (2)	L. hirsutum (2)		
L. glandulosum	L. glandulosum		

Fig. 2.   Comparison of classifications of *Solanum* L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon.* Species in boldface have red to orange to yellow fruits and those without boldface have green fruits. The numbers in parentheses represent numbers of infraspecific ranks (subspecies, varieties, and forms). The lines connect synonymous taxa.
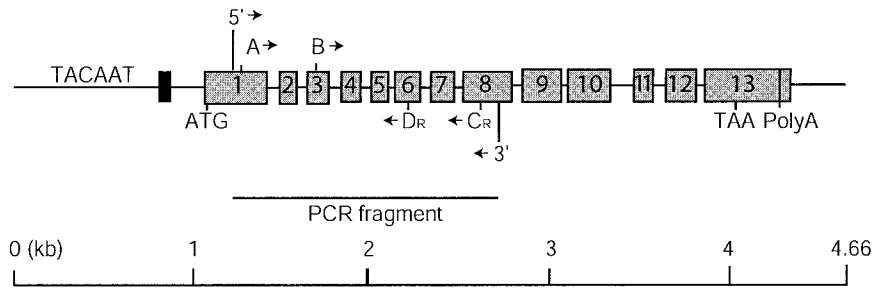
Fig. 3. Schematic of the structural potato GBBSI (*waxy*) gene analyzed in this study. The promoter sequence (TACAAAT), translational start (ATG), stop codon (TAA), and polyadenylation sites (polyA) are indicated. The black box represents an untranslated exon, shaded boxes the 13 translated exons and interconnecting lines introns. The fragment obtained by the polymerase chain reaction comprises ~1525 base pairs (from the first to the eighth exon) and is indicated with a line below the schematic. Arrow heads indicate position and direction of primers 5′ and 3′ that were used to amplify the fragment and forward primers A and B and reverse primers CR and DR that were used for cycle sequencing. The scale represents the total length of complete genomic nucleotide sequence of potato GBBSI (modified from van der Leij et al., 1991).

Raush, and Manos, 1999). Nevertheless, ITS sequences showed more variation than GBSSI sequences within the genus *Physalis,* although only 700 base pairs (bp) of GBSSI were analyzed (Whitson and Manos, 1999).

Many studies documented GBSSI as a single-copy gene in cereals (Shure, Wessler, and Fedoroff, 1983; Klösgen et al., 1986; Rohde, Becker, and Salamini, 1988; Wang et al., 1990; Clark, Robertson, and Ainsworth, 1991) and in dicots (van der Leij et al., 1991; Dry et al., 1992; Mérida et al., 1999; Wang, Yeh, and Tsai, 1999). It has been shown that GBSSI is a low-copy number gene in allotetraploid cassava, but it was unclear whether it is a single gene with different alleles at a locus or has more than one GBSSI locus (Salehuzzaman, Jacobsen, and Visser, 1993). The Rosaceae (Evans et al., 2000) GBSSI contains more than one copy of the GBSSI locus. The GBSSI gene was cloned in potato (van der Leij et al., 1991) and comprises 4663 bp, has 13 introns, and encodes a 58.2 kilodalton mature protein with 540 amino acids. Because potatoes and tomatoes are supported as sister groups by morphology and cpDNA data, the GBSSI gene sequence information of van der Leij et al. (1991) was used to design primers in tomatoes. Our study is the first phylogenetic application of GBSSI in the Solanaceae using the 5′ portion of this nuclear gene, from the first until the eighth exon, including seven introns and comprising 1311 nucleotides.

The main goal of our research was to examine the phylogenetic relationships of all nine wild tomato species and closely related outgroups, with a concentration on the most widespread and variable tomato species, *Solanum peruvianum,* using DNA sequences of the structural gene granule-bound starch synthase. A secondary goal was to compare these results to previous systematic treatments in wild tomatoes that use morphology (Müller, 1940; Luckwill, 1943; Child, 1990), crossing relationships (Rick, 1963, 1979, 1986), isozyme (Rick and Fobes, 1975a, b; Rick and Tanksley, 1981; Rick, 1983, 1986; Bretó, Asins, and Carbonell, 1993), or molecular data (Palmer and Zamir, 1982; McClean and Hanson, 1986; Miller and Tanksley, 1990).

MATERIALS AND METHODS

Plants—In the ingroup (*Solanum* section *Lycopersicon* subsection *Lycopersicon*) three accessions per species were analyzed for *S. cheesmaniae, S. chmielewskii, S. chilense, S. habrochaites, S. lycopersicum* var. *cerasiforme, S. neorickii, S. pennellii*; five accessions for *S. pimpinellifolium*; and 39 for *S. peruvianum*. Many more accessions of *S. peruvianum* were analyzed because this

species is the most widespread and has the greatest degree of morphological, molecular, and crossability variation (Rick, 1963, 1979, 1986; Miller and Tanksley, 1990). One accession of a cultivated tomato, *S. lycopersicum* CV 'Mountain Spring,' was also included. The outgroups were selected according to the cpDNA results of Spooner, Anderson, and Jansen (1993). Like the ingroup, they are all diploid at $2n = 24$ (Table 2). Two accessions per species were included in the outgroup taxa *S. juglandifolium, S. lycopersicoides, S. ochranthum, S. sitiens* and one accession per species in the outgroup taxa *S. bulbocastanum, S. etuberosum, S. jamesii, S. muricatum* and *S. palustre*. In total, 79 accessions were used for the phylogenetic analysis (Table 1). All the tomato (*Solanum* sect. *Lycopersicon* subsection *Lycopersicon*) accessions and outgroups in sect. *Lycopersicon* subsection *Lycopersicoides* and sect. *Juglandifolium* were obtained from the Tomato Genetics Resource Center, Department of Vegetable Crops, University of California, Davis. All potato accessions (*Solanum* sect. *Petota*) and members of sect. *Etuberosum* were obtained from the United States Potato Introduction Station, National Research Support Program-6 (NRSP-6), Sturgeon Bay, Wisconsin, USA (Bamberg et al., 1996); one accession of sect. *Basarthrum* was obtained from Dr. Gregory Anderson, University of Connecticut, Storrs, Connecticut, USA. Dr. Charles Rick, curator of the Tomato Genetics Resource Center, provided advice for the choice of tomato accessions based on geographic distribution, morphology, genetic diversity, and breeding behavior. Vouchers for these accessions were deposited at DAV, MERL, PTIS, and WIS.

DNA isolation and purification—Total genomic DNA was isolated from young leaves of single plants using a mini-extraction procedure kindly provided by Dr. Dario Bernacchi from Dr. Stephen Tanksley's laboratory at Cornell University, Ithaca, New York, USA. Three to eight small leaflets were transferred to a 1.5-mL plastic tube, frozen with liquid nitrogen ground, and then 750 μL of extraction buffer was added. The DNA extraction buffer was prepared with 1 volume of buffer A (1275.4 g Sorbitol, 242.2 g Tris, 37.22 g EDTA Na$_2$ salt, 20 L of ddH$_2$O, pH = 8.26, and 3.802 g/L of Na-bisulfite [the pH was adjusted to 7.5 after the addition of Na-bisulfite]), 1 volume of buffer B (nuclei lysis buffer: 200 mL of 1 mol/L Tris, 200 mL of 0.25 mol/L EDTA, 400 mL of 5 mol/L NaCl, 20 g of CTAB, 200 mL of ddH$_2$O, pH 7.5), and 0.4 volumes of 5% Sarcosyl (50 g/L of N-lauroylsarcosine). The sample was incubated in a 65°C water bath for 30–60 min. The tube was filled with chloroform : isoamyl alcohol (24 : 1), vortexed vigorously, and microcentrifuged at 10 000 rpm for 5 min. The aqueous phase was transferred into a new 1.5-mL tube, and 2/3–1 volume of cold isopropanol was added to precipitate the DNA. The sample was microcentrifuged at 10 000 rpm for 5 min, the supernatant was discarded, and the DNA was washed with 70% ethanol. The DNA was resuspended in 50 μL of TE (10 mL of 1 mol/L Tris, pH 8, 4 mL of 0.25 mol/L EDTA, pH 7, 986 mL of ddH$_2$O). The crude total DNAs were cleaned using acetate salts (7.5 mol/L ammonium acetate and 2.5 mol/L sodium acetate). Purified DNAs were used as templates for PCR amplification using the DNAeasy Plant Mini Kit (QIAgen Valencia, California, USA).

TABLE 2. Number of distinct GBSSI sequences generated from different clones obtained form a single PCR product of a single individual. Each individual represents one of the seven accessions listed below. The number of identical clones is indicated in parentheses. Numbers indicate the exon or intron position as illustrated in Fig. 3. A, C, G, T are nucleotides, and nucleotide substitutions are indicated with a slash. For example, all 12 clones sequenced for GBSSI for *S. chilense* LA 1930 were identical. However, for *S. peruvianum* LA110, we obtained six distinct GBSSI sequences types in 10 clones and four of those were identical clones for a sequence with a G/A transition in intron 7, a T/A transversion in intron 8, and an ATA deletion in intron 3.

Accessions	No. distinct GBSSI sequences (no. identical clones)	Exons		Introns			
		Transition	Transversion	Transition	Transverison	Deletions	Isertions
S. chilense LA1930	1 (12)						
S. cheesmaniae LA1450	1 (10)						
S. peruvianum LA110	1 (4)	7 G/A			8 T/A	3 ATA	
	2 (2)		8 A/C		8 C/G	3 T	
	3 (1)	7 G/A	8 A/C		8 C/G	3 ATA	
	4 (1)				8 T/A	3 T	
	5 (1)	7 G/A	6 A/C		5 T/A; 8 T/A	3 T	
	6 (1)		8 A/C				
S. peruvianum LA111	1 (5)						
	2 (2)		7 A/C				
	3 (1)	2 T/C; 4 T/C			2 G/T		
S. peruvianum LA1336	1 (3)	4 C/T	7 T/A			3T	
	2 (3)			5 C/T		3 ATA	
	3 (1)	3 T/C		5 C/T		3 T	
	4 (1)	4 C/T					
S. peruvianum LA1556	1 (5)				4 C/A; 7 A/T		
	2 (1)				5 A/T		4 T
S. peruvianum LA2152	1 (4)	4 C/T	4 T/A	3 C/T		7 T	
	2 (2)	4 C/T	4 T/A	3 C/T			
	3 (1)				3 C/G	$3 (AT)_3$	

Generation of GBSSI sequences—The oligonucleotide primers were developed based on the GBSSI sequences of *S. tuberosum* obtained from GenBank. Two oligonucleotides were designed as external primers (Fig. 3) for the polymerase chain reaction (PCR), in the first exon region at position 136 (5′-GATGGGCTCCAATCAAGAACTAAT-3′) and in the eighth exon at position 1639 (5′-GCCATTCACAATCCCAGTTATGC-3′) of the published sequence of GBSSI (van der Leij et al., 1991), using the software program Oligo (Rychlik, 1998). Within this region (1311 bp) there are seven introns (from second to eighth introns) of lengths 84, 84, 101, 83, 80, 90, and 83 bp, respectively (total of 602 bp; van der Leij et al., 1991). This region was amplified using PCR for all accessions. In all cases, a single PCR product was obtained (Fig. 3).

Amplification reactions were conducted in a total volume of 50 μL with a Perkin Elmer programmable thermal controller. Each reaction comprised Promega (Madison, Wisconsin, USA) reaction buffer (1× final concentration), MgCl (2.5 mmol/L), dNTPs (200 μmol/L each), the two external primers (0.4 mmol/L each), bovine serum albumin (0.4 mmol/L), template DNA (50–100 ng/μL), and Promega *Taq* DNA polymerase (2.5 units/μL).

The amplification reaction was conducted for 4 min at 94°C, then 5 cycles of 30 sec at 94°C, 1 min at 50°C, and 1.5 min at 72°C, then 30 cycles of 30 sec at 94°C, 1 min at 55°C, and 1.5 min at 72°C, with a 10-min extension at 72°C, ending with 4°C hold. To determine whether amplification was successful, samples were run on a 1.0–1.2% agarose minigel stained with ethidium bromide and viewed with ultraviolet light. As a control, a lambda clone of the GBSSI gene from potato, kindly provided by Dr. Richard Visser (Wageningen, The Netherlands), was used. A DNA ladder (Gibco BRL low mass DNA ladder; Life Technologies, Rockville, Maryland, USA) was included to estimate the length and concentration of the amplified products.

Successful reactions were purified for cycle sequencing using cleaning columns (QIAgen Valencia, California, USA). Concentration of amplified cleaned product was checked with a fluorimeter and adjusted to 10–30 ng per 20 μL reaction. Cycle sequencing reactions were carried out using the ABI Prism Dye Terminator Cycle Sequencing Kit Reaction with AmpliTaq DNA Polymerase FS, and with Big Dye (Perkin-Elmer, Applied Biosystems, Foster City, California, USA) according to manufacturer's instructions.

Four internal primers (Fig. 3) were designed based on potato Genbank sequence to match conserved exon regions, beginning at position 174 (GBSSI A: 5′-CAAGATGGCATCCAGAACTGAGA-3′), 622 (GBSSI B: 5′-CACTGCTATAAACGTGGGGTTGA-3′), 1555 (GBSSI CR: 5′-GGCATAGTATGGGCTCACTGTAA-3′), and 1126 (GBSSI DR 5′-GGAATGAGAGCTGTGTGCCAATC-3′) of the published sequence of GBSSI (van der Leij et al., 1991). Sequences were obtained from both the forward (A, B) and reverse (C$_R$ and D$_R$) primers.

Cycle sequencing was conducted for 25 cycles, each cycle consisting of 96°C for 30 sec, 46°C for 15 sec, and 60°C for 4 min. Sequencing products were precipitated in 50 μL of 95% ethanol and 2 μL of 3 mol/L sodium acetate, and then vacuum-dried. The samples were resuspended in formamide loading buffer (84% formamide, 4 mmol/L EDTA, 8 mg/mL Blue Dextran) and electrophoresed on a Perkin-Elmer Applied Biosystems 377XL automated DNA sequencer.

Some of the GBSSI sequences showed polymorphisms, particularly in the third intron, which was difficult to align. A similar protocol was used to amplify GBSSI in the problematic accessions, but in this case using *Pfu* polymerase (Promega, Madison, Wisconsin, USA), which produces less error during the PCR amplification than *Taq* polymerase. Cleaned PCR products were then cloned into Promega's pGEM-T Easy vector. The vector contain an A 3′ terminal thymidine overhang, and T7 and SP6 RNA polymerase promoters flanking multiple cloning regions including the α-peptide coding region of the enzyme β-galactosidase and multiple restriction sites. Because *Pfu* polymerase produces blunt ends, an A-tailing reaction was performed using *Taq* DNA polymerase before the ligation. Ligation, transformation, and plating

were done according to the protocols provided by the manufacturer. Insertional inactivation of the α-peptide allows recombinant clones to be directly identified by color screening on indicator plates. Plasmid preparations were done with alkaline lysis (Wizard plus plasmid purification systems; Promega, Madison, Wisconsin, USA). Purified plasmid DNA was digested with *EcoRI*, and cut and uncut products were checked on 0.9% agarose gel for presence of inserts. A lambda control of GBSSI was included to check the size of the band. Two molecular ladders were also included (Gibco BRL low and high mass DNA ladders; Life Technologies, Rockville, Maryland, USA), and concentrations of the inserts were estimated by visual comparisons of band intensity. Plasmid minipreps containing the GBSSI fragment were cycle sequenced as described above. Six to 12 clones were sequenced for each individual plant per accession (Table 2).

Sequence alignment—DNA sequences initially were edited using Sequencher software (Gene Codes Corporation, Ann Arbor, Michigan, USA). These initially aligned sequences were then imported into PAUP* version 4.0d64 (Swofford, 1998) and were further aligned visually.

Phylogenetic analysis—Parsimony analyses of the DNA sequences were conducted with PAUP* 4.0d64 (Swofford, 1998). *Solanum muricatum* was used to root the trees based on the results of Spooner, Anderson, and Jansen (1993) and Olmstead and Palmer (1997). The first analyses were performed using Fitch parsimony, which allows free reversibility (change of characters in any direction among the nucleotide states). All the characters were included and were equally weighted, and gaps were considered as missing data. An initial tree was generated by jackknife (36.8% deletion, 10 000 replicates, and fast stepwise addition) (Farris et al., 1996). This tree was used as a starting tree for further heuristic searches using tree-bisection-reconnection (TBR) swapping algorithm, with retention of all equally parsimonious trees (MUL-PARS) with stepwise addition and steepest descent. The amount of homoplasy was evaluated with the consistency index (CI; Kluge and Farris, 1969) and the retention index (RI; Farris, 1989). Consensus trees were generated from all most parsimonious trees. Bootstrap analysis (Felsenstein, 1985), using 1000 replicates, and jackknife, also with 1000 replicates (36.8% deletions; Farris et al., 1996), were conducted to estimate the internal relative support for each branch. Also, a decay analysis (Bremer, 1988; Donoghue et al., 1992) was calculated using an heuristic TBR search. Bootstrap and decay values are placed on the strict consensus tree (Fig. 5).

In the second analysis, a similar heuristic search was performed but the insertion and deletion characters (indels) were scored as multistate characters attached to the nucleotide site data (Baum, Sytsma, and Hoch, 1994). The CI and RI, as well as the bootstrap values and decay support, were calculated and compared with the previous trees.

RESULTS

GBSSI sequence variation within individuals—About 15% of the 79 accessions showed some sites with two, and sometimes three, overlapping nucleotide peaks in the same position in the GBSSI sequence. These polymorphic sites were always found within the intron regions, mainly at the beginning of the third intron. To ascertain the nature of these variants, the GBSSI DNA fragments were cloned and sequenced from seven of these individuals that displayed infraindividual polymorphism (*S. cheesmaniae* LA1450, *S. chilense* LA1930, *S. peruvianum* LA110, LA111, LA1336, LA1556, LA2152). Table 2 lists the number of GBSSI clones that were sequenced from these seven individuals, the number of distinct sequence types found, and their nucleotide variation and position. A putative microsatellite $(AT)_n$ was found in the third intron, being $n = 1$ in potato and $n = 2$ up to $n = 9$ in the other accessions.

For some accessions (e.g., *S. chilense* LA1930 and *S. cheesmaniae* LA1450) all clones were identical, yet these dif-

fered from the sequences of the PCR product generated using *Taq* polymerase. We assumed that the sequences obtained from the clones were the most accurate, because *Pfu* polymerase produces less error. These sequences were included in the cladistic analysis. Sequences from five accessions of *S. peruvianum* differed by substitutions in the exon and intron regions and deletions/insertions only in the intron regions (Table 2). We initially included all sequence types in a cladistic analysis, but most were so similar (differing only by 1–3 base pairs) that they made no difference in the cladistic results. Others differed by up to 3 base pairs and were entered separately in the final analysis (indicated in italic boldface, Figs. 4, 5). In all cases they remained in the same part of the cladogram. In addition, cloned sequences of *S. pennellii* LA1376 has complex polymorphisms suggesting to us it is a possible hybrid; we do not include these results here but are investigating this further.

Cladistic results—The aligned data matrix consisted of 1384 characters. Additionally we scored 20 indels as binary or multistate characters. Indels with potential phylogenetic information were more frequent in the introns, and few were detected in the exons. The GBSSI intron sequences are at base pair positions 22–185, 267–380, 480–622, 713–830, 895–983, 1085–1176, 1287–1374 of the aligned sequences (data available in Peralta [2000] and at: *http://ajbsupp. botany.org/v88/peralta.html*). Most gap characters were present in the outgroups (18 indels). In the phylogenetic analysis of the data, considering gaps as characters and using 134 phylogenetically informative characters, 15 000 equally most parsimonious trees were saved (length = 347, CI = 0.88, RI = 0.95; Fig. 4). Identical strict consensus trees (Fig. 5) were generated considering gaps as missing data or as characters.

Bootstrap analysis using 1000 replicates gave an excellent support for the outgroups (86–100%). Almost identical support values were found with the jackknife method. The ingroup had much less resolution and consisted of a basal polytomy, from which some branches were resolved, and a terminal clade. The basal polytomy was composed entirely of the self-incompatible green-fruited species *S. chilense* (map localities 29, 36, 38; Fig. 1), the central to southern Peruvian populations of *S. peruvianum* (24 populations, plus three additional sequence types, map localities 18, 19, 21, 22, 23, 24, 28, 29, 30, 32, 33, 34, 35, 36, 37), *S. habrochaites* (map localities 3, 11, 24), and *S. pennellii* (map localities 19, 28, 31). The latter two species are resolved within this basal polytomy, with bootstrap values of 86 and 96%, respectively. Three other branches of *S. peruvianum* were supported within this basal polytomy with bootstrap values of 62–99%. Two of these branches contain populations of *S. peruvianum* in the same generalized geographic area (map location 21) or contiguous generalized areas (21, 22, and 23).

The derived clade (70% bootstrap value) contains 14 northern Peruvian populations of *S. peruvianum* (also self-incompatible, green-fruited, map localities 6, 9, 10, 11, 12, 13, 15), and one central Peruvian population of *S. peruvianum* (LA1609, map locality 22), *S. chmielewskii* and *S. neorickii* (self-compatible, green-fruited, map localities 25, 26, 27 and 7, 17, 20, respectively), and the self-compatible and red- to orange-fruited species *S. cheesmaniae, S. lycopersicum* (map localities 5, 21, 32), and *S. pimpinellifolium* (map localities 2, 7, 8, 23). The latter three species cluster together as a polytomy within this advanced clade, but include one northern ac-

Fig. 4. One representative phylogram of 15 000 equally parsimonious trees (indels treated as characters), length = 347, consistency index = 0.88, retention index = 0.95. Different sequence types are indicated in italic boldface. Numbers indicate branch lengths, and unnumbered branches have a length of 1.

cession of *S. peruvianum* (LA2163, green-fruited, self-incompatible). The derived clade includes two groups of *S. peruvianum* with bootstrap values higher than 50%, and a branch corresponding to the red- to orange- to yellow-fruited self-

compatible species, with a bootstrap value of 24%, containing *S. cheesmaniae, S. lycopersicum, S. pimpinellifolium,* and one accession of *S. peruvianum.* For branches with <50% bootstrap support the decay index = 1, for branches between 50

Fig. 5. Strict consensus tree of 15000 most parsimonious trees (indels treated as characters) based on GBSSI sequences. Percentage of 1000 bootstrap replicates is given above branches. Decay values between 1 and 3 are in parentheses; the other branches in the outgroup are at least 13 steps longer. Distinct GBSSI sequence types from different clones obtained from single individuals of seven accessions are indicated in italic boldface and a number (see Table 2). Map localities correspond to Fig. 1 and Table 1; G represents accessions from the Galapagos Islands, and S represents unmapped continental accessions. "North per" refers to accessions from the northern range of *S. peruvianum,* while "south per" refers to accessions from central and southern populations of *S. peruvianum,* as shown in Fig. 1.

and 70% the decay index = 2, and for branches between 70 and 95% the decay index = 3 (Fig. 5).

DISCUSSION

The GBSSI sequences provided good resolution at the outgroup level and helped clarify relationships among five of the nine sections recognized by Child (1990) within *Solanum* subgenus *Potatoe*. Those results are entirely concordant with the cpDNA phylogeny of *Solanum* (Spooner, Anderson, and Jansen, 1993) regarding interspecific relationships of *Solanum* section *Lycopersicon* subsection *Lycopersicoides,* sect. *Juglandifolium,* sect. *Basarthrum,* sect. *Petota,* and sect. *Etuberosum.* The only difference is the cladistic separation of the three tuber-bearing members of sect. *Petota, S. tuberosum* on one branch and *S. bulbocastanum* + *S. jamesii* on a different branch, but this likely is a taxa sampling-density problem.

The GBSSI sequences provided less phylogenetic information at the ingroup than at the outgroup level, but still provided insights into taxonomic relationships among tomatoes. These results support *S. juglandifolium* and *S. ochranthum* as the closest outgroup to tomatoes, with *S. lycopersicoides* and *S. sitiens* as basal to these taxa, and support allogamy, self-incompatibility, and green fruits as primitive.

Mating system have played an important role in the evolution of wild tomatoes. Self-incompatibility has been considered as an ancestral condition to self-compatibility in tomatoes and probably never reversed to self-incompatibility (Rick, 1982). In this evolutionary context, changes from self-incompatibility to self-compatibility are expected to arise frequently and independently in different lineages (Rick, 1982). Interestingly, *S. habrochaites* and *S. pennellii* have both self-incompatible and self-compatible populations. The self-incompatible populations have higher genetic variation and occupy the center of their species geographic distribution. Self-compatible populations occur toward the northern and southern edges of distribution and have less genetic variation and smaller flower parts (Rick, Fobes, and Tanksley, 1979; Rick and Tanksley, 1981; Rick, 1984). The change from self-incompatibility to self-compatibility has been reported in only one population of *S. peruvianum* LA2157 from Chotano, Peru (Rick, 1986). Self-compatibility and the red to orange to yellow fruit color in *S. cheesmaniae, S. lycopersicum,* and *S. pimpinellifolium* have been considered as derived synapomorphies for this monophyletic group (Palmer and Zamir, 1982). The two closely related species, *S. chmielewskii* and *S. neorickii,* are also self-compatible. The latter species is exclusively autogamous with low population genetic variation and small flowers with stigmas included in the anther tube. In contrast, the facultative allogamous *S. chmielewskii* has larger flower parts and higher levels of heterozygosity. It has been postulated that *S. neorickii* evolved from *S. chmielewskii* (Rick et al., 1976). All the populations of *S. chilense,* and the most closely related species outgroups, *S. lycopersicoides, S. sitiens, S. ochranthum,* and *S. juglandifolium,* are exclusively self-incompatible. In *Solanum* sect. *Petota,* sect. *Etuberosum* and sect. *Basarthrum* self-compatibility most probably arose independently.

Within the ingroup, the phylogenetic analyses of GBSSI sequences identified *S. pennellii* and *S. habrochaites* as well-supported branches within the basal polytomy. *Solanum peruvianum* is supported as paraphyletic, with the taxa separating into two geographical groups, one in northern Peru ("northern group") and another into the central and southern Peru ("southern group"). All the red- to orange- to yellow-fruited self-compatible autogamous species (*S. cheesmaniae, S. lycopersicum,* and *S. pimpinellifolium*) and the green-fruited facultative allogamous species (*S. chmielewskii* and *S. neorickii*) are derived and on the same clade as these northern populations of *S. peruvianum.*

The GBSSI sequences presented very low variation in the autogamous species, compared with the considerable diversity found in all allogamous species. These results are in concordance with isozyme data (Rick and Fobes, 1975a, b; Rick and Tanksley, 1981; Rick, 1983, 1986; Bretó, Asins, and Carbonell, 1993) and nRFLP data that showed self-incompatible species to harbor, on average, ten times more genetic variation within accessions than the self-compatible ones (Miller and Tanksley, 1990). In fact, more variation was found within a single accession of the self-incompatible species *S. peruvianum* than among all accessions tested of any one of the self-compatible species (Miller and Tanksley, 1990). The estimated amount of DNA polymorphism reflects breeding systems, consequently less within-species variation was found in the selfing species than in the outcrossing ones (Wolfgang and Langley, 1998). Similarly, low levels of DNA polymorphism were found in molecular linkage maps based on a cross between two closely related autogamous species, *S. lycopersicum* and *S. pimpinellifolium,* compared with the higher level of polymorphisms found in maps based on intraspecific crosses between *S. lycopersicum* and other allogamous species (Chen and Foolad, 1998). Results from different molecular markers demonstrated the difficulty of finding polymorphisms among self-compatible species. More variable markers, such as microsatellites used to identify tomato cultivars (Broun and Tanksley, 1996; Smulders et al., 1997), may be better suited to compare relationships among autogamous tomato species.

The low level of genetic variation found within self-compatible species may be explained by the role of autogamy that drives the loss of variation and fixation of alleles (Rick, 1984). The low variation among the self-compatible species showed by cpDNA (Palmer and Zamir, 1982), allozyme (Bretó, Asins, and Carbonell, 1993), nRFLP (Miller and Tanksley, 1990), molecular linkage maps (Chen and Foolad, 1998), and GBSSI sequences may reflect recent lineage divergence (Moniz de Sá and Drouin, 1996).

The maternal phylogenetic results from cpDNA restriction sites of Palmer and Zamir (1982) showed that the red- to orange- to yellow-fruited self-compatible autogamous species (*S. cheesmaniae, S. lycopersicum,* and *S. pimpinellifolium*) formed a clearly derived monophyletic group. The nRFLP showed the same results (Miller and Tanksley, 1990). The same group has been classified as series *Lycopersicon* in a morphological treatment by Child (1990; Fig. 2). The GBSSI data also were concordant in supporting the derived nature of the red-fruited autogamous species, and the monophyly of this group was confirmed (but with a low bootstrap value of 24%), with the exception of inclusion the single accession of *S. peruvianum* LA2163 from Yamalúc (Cajamarca Province, Peru). This northern accession came from the same area and is closely related genetically to the only self-compatible population of this species, *S. peruvianum* LA2157 from Tunel Chotano (Cajamarca Province, Peru). These northern Peruvian populations have been considered ancestral stocks (Rick, 1986), and *S. peruvianum* LA2163 may represent a progenitor population of this derived clade.

Critical results revealed by the GBSSI data are the para-

phyly of *S. peruvianum* and the close relationship of the northern populations with the self-compatible taxa, also supported by nRLFP (Miller and Tanksley, 1990). The morphological and crossability differences between the northern and southern populations of *S. peruvianum* have long been recognized. The northern populations were separated as *Lycopersicon* [*Solanum*] *peruvianum* var. *humifusum* C. H. Müll. based on plants with decumbent stems, very short trichomes, and leaves with few leaflets (Müller, 1940). Rick (1963, 1986) demonstrated that these northern races were partially crossable among themselves but with reduced crossability to the southern populations that grow south of 8° S latitude (Rick, 1986). In addition, Miller and Tanksley (1990) showed that these two groups could be distinguished by nuclear restriction fragment length polymorphism data, although only one population of the northern group was analyzed, *S. peruvianum* var. *humifusum* LA2150 (not included in our GBSSI study). Within the northern *S. peruvianum* clade, only one central Peruvian population, *S. peruvianum* LA1609, was also found. Interestingly, this accession is geographically and morphologically related to other two coastal races of *S. peruvianum,* LA107 and LA1373 (not included in our GBSSI study), that have been used as crossing bridge populations between the northern and the southern populations of *S. peruvianum* (Rick, 1963, 1986). The inclusion of *S. peruvianum* LA1609 within the northern population clade suggests a possibility of naturally occurring gene exchange between northern and southern groups.

The GBSSI gene sequence results also showed the close relationship of two self-compatible species, *S. chmielewskii* and *S. neorickii,* with the northern *S. peruvianum* group. Similarly, the nRFLP supported a northern population, *S. peruvianum* var. *humifusum* LA2150, as the most closely related to *S. chmielewskii* and *S. neorickii*. In the maternal phylogenetic results from cpDNA restriction sites, *S. chmielewskii* also appeared closely related to a northern population, *S. peruvianum* LA1032 (not included in our GBSSI study), and Palmer and Zamir (1982) proposed the inclusion of *S. chmielewskii* as subspecies of *S. peruvianum*. In the phylogenetic GBSSI gene sequence results, the allogamous self-incompatible *S. chilense* was included within the basal polytomy with the southern races of *S. peruvianum*. A similar relationship was found with cpDNA restriction sites, and Palmer and Zamir (1982) also proposed the inclusion of *S. chilense* as subspecies of *S. peruvianum.*

Mitochondrial DNA (mtDNA) restriction fragment length studies were used to address relationships of nine *Lycopersicon* species and two closely related *Solanum* species (McClean and Hanson, 1986). The divergence estimates were based on the ratio between the number of shared fragments and all fragments evaluated. The mtDNA divergence is higher than that in cpDNA, indicating that the DNA of the two organelles is evolving at different rates (McClean and Hanson, 1986). Tomato mitochondrial genomes exhibit a small amount of divergence, and the variability was more limited within the self-compatible *S. lycopersicum* (0.04%) than between this species and the self-incompatible *S. sitiens* (2.7%). The inferred phylogeny from the GBSSI data confirms the position of this latter taxa as an outgroup species. Nevertheless, cpDNA, nRFLPs, and GBSSI results were not in agreement with the species relationships showed by mitochondrial results. These differences can be explained by the possibility of mitochondrial genome rearrangements that could result in estimates of divergences greater than the true values (McClean and Hanson,

1986). The incorrect assessment of homology of the mtDNA bands is another possible explanation that may account for the species relationship incongruence among data sets.

In agreement with crossing data, GBSSI gene sequence results distinguished the northern and southern *S. peruvianum* groups and also the close relationship of this species with *S. chilense* (Rick, 1963, 1979, 1986). Nevertheless, the molecular results of this study do not support the placement of *S. pennellii* and *S. habrochaites* within the Esculentum group (Rick, 1979), and *S. chmielewskii* and *S. neorickii* were more closely related to the northern *S. peruvianum* group than to the red- to orange- to yellow-fruited self-compatible species. The GBSSI results support a close relationship among green-fruited species, in agreement with morphological treatments of wild tomatoes (Müller, 1940; Luckwill, 1943).

Multiple GBSSI clones obtained from a single individual had identical sequences for *S. chilense* LA1930 and *S. cheesmaniae* LA1450 (Table 2). Nevertheless, multiple clones obtained from a single PCR performed on five different individuals, each one representing five *S. peruvianum* accessions, differed by nucleotide substitution and deletion/insertion events (Table 2). In four cases (*S. peruvianum* LA110, LA111, LA1336, and LA2152) more than two GBSSI sequence variants were identified. Theoretically, in diploid wild tomato species, a single nuclear gene as GBSSI has up to two different alleles in the heterozygote. True allelic variation at a single locus could explain the presence of two distinct sequences of GBSSI. However, the presence of more than two variants could be caused by artifactual errors that may occur as a consequence of nucleotide misincorporation during the PCR amplification of the GBSSI gene (Shimada and Tada, 1991). Another possibility to explain the observed intraindividual variation is the presence of more than one GBSSI locus (Salehuzzaman, Jacobsen, and Visser, 1993; Evans et al., 2000). In spite of the variation found among clones, all the variants used in with the cladistic analysis fell within the same clade.

The GBSSI DNA data, like other molecular markers used in tomatoes to date, were not able to completely differentiate very closely related, and most probably recent divergent species, especially the self-compatible ones. The GBSSI provide useful information about intraspecific variation in *S. peruvianum* and about character evolution regarding breeding systems and fruit color. A future paper will use these same accessions to investigate the morphological support for species boundaries and relationships.

LITERATURE CITED

BAMBERG, J. B., M. W. MARTIN, J. J. SCHARTNER, AND D. M. SPOONER. 1996. Inventory of tuber-bearing *Solanum* species. Catalog of potato germplasm. Potato Introduction Station, National Research Support Program-6 (NRSP-6), Sturgeon Bay, Wisconsin, USA.

BAUM, D. A., K. J. SYTSMA, AND P. C. HOCH. 1994. A phylogenetic analysis of *Epilobium* (Onagraceae) based on nuclear ribosomal DNA sequences. *Systematic Botany* 19: 363–388.

BOHS, L., AND R. G. OLMSTEAD. 1997. Phylogenetic relationships in *Solanum* (Solanaceae) base on *ndhF* sequences. *Systematic Botany* 22: 5–17.

BOHS, L., AND R. G. OLMSTEAD. 1999. *Solanum* phylogeny inferred from chloroplast DNA sequence data. *In* M. Nee, D. E. Symon, R. N. Lester, and J. P. Jessop [eds.], Solanaceae IV, advances in biology and utilization, 97–110. Royal Botanic Gardens, Kew, UK.

BÖRNER, C. 1912. *Solanum* L. und *Solanopsis* gen. nov. *Abhandlungen herausgegeben vom Naturwissenschaftlichen Verein zu Bremen* 21: 282.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.

BRETÓ, M. P., M. J. ASINS, AND E. A. CARBONELL. 1993. Genetic variability

in *Lycopersicon* species and their genetic relationships. *Theoretical and Applied Genetics* 86: 113–120.

BROUN, P., AND S. D. TANKSLEY. 1996. Characterization and mapping of simple repeat sequences in the tomato genome. *Molecular and General Genetics* 250: 39–49.

CHEN, F. Q., AND M. R. FOOLAD. 1998. A molecular linkage map of tomato based on a cross between *Lycopersicon esculentum* and *L. pimpinellifolium* and its comparison with other molecular maps of tomato. *Genome* 42: 94–103.

CHILD, A. 1990. A synopsis of *Solanum* subgenus *Potatoe* (G. Don) D'Arcy [*Tuberarium* (Dun.) Bitter (s.l.)]. *Feddes Repertorium* 101: 209–235.

CLARK, J. R., M. ROBERTSON, AND C. C. AINSWORTH. 1991. Nucleotide sequence of a wheat (*Triticum aestivum* L.) cDNA clone encoding the *waxy* protein. *Plant Molecular Biology* 16: 1099–1101.

D'ARCY, W. G. 1987. The circumscription of *Lycopersicon. Solanaceae Newsletter* 2: 60–61.

DONOGHUE, M. J., R. G. OLMSTEAD, J. F. SMITH, AND J. D. PALMER. 1992. Phylogenetic relationships of Dipsacales based on *rbcL* sequences. *Annals of the Missouri Botanical Garden* 79: 333–345.

DOYLE, J. J., V. KANAZIN, AND R. C. SHOEMAKER. 1996. Phylogenetic utility of histone H3 intron sequences in the perennial relatives of soybean (*Glycine*; Leguminosae) *Molecular Phylogenetics and Evolution* 6: 438–447.

DRY, I., A. SMITH, A. EDWARDS, M. BHATTACHARYYA, P. DUNN, AND C. MARTIN. 1992. Characterization of cDNAs encoding two isoforms of granule-bound starch synthase which show differential expression in developing storage organs of pea and potato. *Plant Journal* 2: 193–202.

EVANS, R. C., L. A. ALICE, C. S. CAMPBELL, E. A. KELLOGG, AND T. A. DICKINSON. 2000. The granule-bound starch synthase (GBSSI) gene in the Rosaceae: multiple loci and phylogenetic utility. *Molecular Phylogenetics and Evolution* 17: 388–400.

FARRIS, J. S. 1989. The retention index and homoplasy excess. *Systematic Zoology* 38: 406–407.

FARRIS, J. S., V. A. ALBERT, M. KÄLLERSJÖ, D. LIPSCOMB, AND A. G. KLUGE. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12: 99–124.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

FOSBERG, F. R. 1987. New nomenclatural combination for Galapagos plant species. *Phytologia* 62: 181–183.

HEINE, H. 1976. Flora de la Nouvelle Caledonie, vol. 7. Musée Nacional D'Histoire Naturelle, Paris, France.

HOLLE, M., C. M. RICK, AND D. G. HUNT. 1978. Catalog of collections of green-fruited *Lycopersicon* species and *Solanum pennellii* found in watersheds of Peru. *Tomato Genetics Cooperative Report* 28: 49–78.

HOLLE, M., C. M. RICK, AND D. G. HUNT. 1979. Catalog of collections of green-fruited *Lycopersicon* species and *Solanum pennellii* found in watersheds of Peru. *Tomato Genetics Cooperative Report* 29: 63–91.

KLÖSGEN, R. B., A. GIERL, Z. S. SCHWARZ-SOMMER, AND H. SAEDLER. 1986. Molecular analysis of the *waxy* locus of *Zea mays. Molecular and General Genetics* 203: 237–244.

KLUGE, A. G., AND J. S. FARRIS. 1969. Qualitative phyletics and the evolution of anurans. *Systematic Zoology* 18: 1–32.

KNAPP, S., AND D. M. SPOONER. 1999. A new name for a common Ecuatorian and Peruvian wild tomato species. *Novon* 9: 375–376.

LESTER, R. N. 1991. Evolutionary relationships of tomato, potato and pepino, and wild species of *Lycopersicon* and *Solanum. In* J. G. Hawkes, R. N. Lester, M. Nee, and N. Estrada [eds.], Solanaceae III: taxonomy, chemistry, evolution, 283–301. Royal Botanic Gardens, Kew, UK.

LINNAEUS, C. 1753. *Species Plantarum,* 1st ed. Holmiae, Stockholm, Sweden.

LUCKWILL, L. C. 1943. The genus *Lycopersicon*: an historical, biological, and taxonomic survey of the wild and cultivated tomatoes. *Aberdeen University Studies* 120: 1–44.

MACBRIDE, J. F. 1962. Flora of Peru: Solanaceae. *Field Museum of Natural History, Botanical Series* 13: 3–267.

MASON-GAMER, R. J., AND E. A. KELLOGG. 1996. Potential utility of the nuclear gene *waxy* for plant phylogenetic analysis. *American Journal of Botany* (supplement)83: 178.

MASON-GAMER, R. J., C. F. WEIL, AND E. A. KELLOGG. 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. *Molecular Biology and Evolution* 15: 1658–1673.

MCCLEAN, P. E., AND M. R. HANSON. 1986. Mitochondrial DNA sequence

divergence among *Lycopersicon* and related *Solanum* species. *Genetics* 112: 649–667.

MÉRIDA, A., J. M. RODRÍGUEZ-GALÁN, C. VINCENT, AND J. M. ROMERO. 1999. Expression of the granule-bound starch synthase I (*Waxy*) gene from snapdragon is developmentally and circadian clock regulated. *Plant Physiology* 120: 401–409.

MILLER, J. C., AND S. D. TANKSLEY. 1990. RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon. Theoretical and Applied Genetics* 80: 437–448.

MILLER, P. 1754. The gardeners dictionary, abridged 4th ed. London, UK.

MILLER, R. E., M. D. RAUSHER, AND P. S. MANOS. 1999. Phylogenetic systematics of *Ipomoea* (Convolvulaceae) based on ITS and *Waxy. Systematic Botany* 24: 209–227.

MONIZ DE SÁ, M., AND G. DROUIN. 1996. Phylogeny and substitution rates of angiosperm actin genes. *Molecular Biology and Evolution* 13: 1198–1212.

MÜLLER, C. H. 1940. A revision of the genus *Lycopersicon. United States Department of Agriculture, Miscellaneous Publication* 382: 1–28.

OLMSTEAD, R. G., AND J. D. PALMER. 1997. Implications for phylogeny, classification, and biogeography of *Solanum* from cpDNA restriction site variation. *Systematic Botany* 22: 19–29.

OLMSTEAD, R. G., J. A. SWEERE, R. E. SPANGLER, L. BOHS, AND J. D. PALMER. 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. *In* M. Nee, D. E. Symon, R. N. Lester, and J. P. Jessop [eds.], Solanaceae IV, advances in biology and utilization, 111–137. Royal Botanic Gardens, Kew, UK.

PALMER, J. D., AND D. ZAMIR. 1982. Chloroplast DNA evolution and phylogenetic relationships in *Lycopersicon. Proceedings of the National Academy of Sciences, USA* 79: 5006–5010.

PERALTA, I. E. 2000. Phylogeny of wild tomatoes (*Solanum* L. section *Lycopersicum* [Mill.] Wettst. subsection *Lycopersicum*) based on morphology and *waxy* gene sequences. Ph.D. dissertation, Plant Breeding and Plant Genetics Program, University of Wisconsin, Madison, Wisconsin, USA.

PERALTA, I. E., AND D. M. SPOONER. 2000. Classification of wild tomatoes: a review. *Kurtziana* 28: 45–54.

RICK, C. M. 1960. Hybridization between *Lycopersicon esculentum* and *Solanum pennellii*: phylogenetic and cytogenetic significance. *Proceedings of the National Academy of Sciences, USA* 46: 78–82.

RICK, C. M. 1963. Barriers to interbreeding in *Lycopersicon peruvianum. Evolution* 17: 216–232.

RICK, C. M. 1971. *Lycopersicon. In* I. L. Wiggins and D. M. Porter [eds.], Flora of the Galapagos Islands, 468–471. Stanford University Press, Palo Alto, California, USA.

RICK, C. M. 1973. Potential genetic resources in tomato species: clues from observations in native habitats. *In* A. M. Srb [ed.], Genes, enzymes, and populations, 255–269. Plenum, New York, New York, USA.

RICK, C. M. 1979. Biosystematic studies in *Lycopersicon* and closely related species of *Solanum. In* J. G. Hawkes, R. N. Lester, and A. D. Skelding [eds.], The biology and taxonomy of Solanaceae. Linnean Society Symposium Series 7, 667–677. Academic Press, New York, New York, USA.

RICK, C. M. 1982. Genetic relationships between self-incompatibility and floral traits in the tomato species. *Biologisches Zentralblatt* 101: 185–198.

RICK, C. M. 1983. Tomato (*Lycopersicon*). *In* S. D. Tanksley and T. J. Orton [eds.], Isozymes in plant genetics and breeding part B, 147–165. Elsevier Science, Amsterdam, The Netherlands.

RICK, C. M. 1984. Evolution of mating systems: evidence from allozyme variation. *In* V. L. Chopra, B. C. Joshi, R. P. Sharma, and H. C. Bansal [eds.], Genetics: new frontiers. Proceedings of the XV International Congress of Genetics, vol. 4, Applied genetics, 215–221. Bowker, Epping, Essex, UK.

RICK, C. M. 1986. Reproductive isolation in the *Lycopersicon peruvianum* complex. *In* W. G. D'Arcy [ed.], Solanaceae, biology and systematics, 477–495. Columbia University Press, New York, New York, USA.

RICK, C. M. 1991. Tomato resources of South America reveal many genetic treasures. *Diversity* 7: 54–56.

RICK, C. M., AND J. F. FOBES. 1975a. Allozymes of Galapagos tomatoes: polymorphism, geographic distribution and affinities. *Evolution* 29: 443–457.

RICK, C. M., AND J. F. FOBES. 1975b. Allozyme variation in the cultivated tomato and closely related species. *Bulletin of the Torrey Botanical Club* 102: 376–384.

RICK, C. M., J. F. FOBES, AND S. D. TANKSLEY. 1979. Evolution of mating systems in *Lycopersicon hirsutum* as deduced from genetic variation in electrophoretic and morphological characters. *Plant Systematics and Evolution* 132: 279–298.

RICK, C. M., AND M. HOLLE. 1990. Andean *Lycopersicon esculentum* var. *cerasiforme*: genetic variation and its evolutionary significance. *Economic Botany* 44: 69–78.

RICK, C. M., E. KESICKI, J. F. FOBES, AND M. HOLLE. 1976. Genetic and biosystematic studies on two new sibling species of *Lycopersicon* from Interandean Peru. *Theoretical and Applied Genetics* 47: 55–68.

RICK, C. M., H. LATERROT, AND J. PHILOUZE. 1990. A revised key for the *Lycopersicon* species. *Tomato Genetics Cooperative Report* 40: 31.

RICK, C. M., AND S. D. TANKSLEY. 1981. Genetic variation in *Solanum pennelli*: comparison with two other sympatric tomato species. *Plant Systematics and Evolution* 139: 11–45.

ROHDE, W., D. BECKER, AND F. SALAMINI. 1988. Structural analysis of the *waxy* locus from *Hordeum vulgare*. *Nucleic Acids Research* 16: 7185–7186.

RYCHLIK, W. 1998. Oligo® primer analysis software, ver. 6 for Macintosh. Molecular Biology Insights, Cascade, Colorado, USA.

SALEHUZZAMAN, S. N. I. M., E. JACOBSEN, AND R. G. F. VISSER. 1993. Isolation and characterization of a cDNA encoding granule-bound starch synthase in cassava (*Manihot esculenta* Crantz) and its antisense expression in potato. *Plant Molecular Biology* 23: 947–962.

SEITHE, A. 1962. Die Haararten der Gatung *Solanum* L. und ihre taxonomische Verwertung. *Botanische Jahrbücher für Systematik, Pflanzengeshichte und Pflanzengeographie* 81: 261–336.

SHIMADA, H., AND Y. TADA. 1991. Rapid isolation of a rice *waxy* sequence: a simple PCR method for the analysis of recombinant plasmids from intact *Escherichia coli* cells. *Gene* 98: 243–248.

SHURE, M., S. WESSLER, AND N. FEDOROFF. 1983. Molecular identification and isolation of the *Waxy* locus in maize. *Cell* 35: 225–233.

SMULDERS, M. J. M., G. BREDEMEIJER, W. RUS-KORTEKAAS, P. ARENS, AND B. VOSMAN. 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicun esculentum* cultivars

and accessions of other *Lycopersicum* species. *Theoretical and Applied Genetics* 97: 264–272.

SOLTIS, D. E., AND P. S. SOLTIS. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. *In* D. E. Soltis, P. S. Soltis, and J. J. Doyle [eds.], Molecular systematics of plants, II: DNA sequencing, 1–42. Kluwer Academic, Dordrecht, The Netherlands.

SPOONER, D. M., G. J. ANDERSON, AND R. K. JANSEN. 1993. Chloroplast DNA evidence for the interrelationships of tomatoes, potatoes, and pepinos (Solanaceae). *American Journal of Botany* 80: 676–688.

SWOFFORD, D. L. 1998. Paup*: phylogenetic analysis using parsimony, version 4. 0d64. Sinauer, Sunderland, Massachusetts, USA.

SYTSMA, K. J. 1990. DNA and morphology: inference of plant phylogeny. *Trends in Ecology and Evolution* 5: 104–110.

TANKSLEY, S. D., AND F. LOAIZA-FIGUEROA. 1985. Gametophytic self-incompatibility is controlled by a single major locus on chromosome 1 in *Lycopersicon peruvianum*. *Proceedings of the National Academy of Sciences, USA* 82: 5093–5096.

TAYLOR, I. B. 1986. Biosystematics of the tomato. *In* J. G. Atherton, and J. Rudich [eds.], The tomato crop: a scientific basis for improvement, 1–34. Chapman and Hall, London, UK.

VAN DER LEIJ, F. R., R. G. F. VISSER, A. S. PONSTEIN, E. JACOBSEN, AND W. J. FEENSTRA. 1991. Sequence of the structural gene for granule-bound starch synthase of potato (*Solanum tuberosum* L.) and evidence for a single point deletion in the *amf* allele. *Molecular Gene Genetics* 228: 240–248.

WANG, S. J., K. W. YEH, AND C. Y. TSAI. 1999. Molecular characterization and expression of starch granule-bound starch synthase in the sink and source tissues of sweet potato. *Physiologia Plantarum* 106: 253–261.

WANG, Z., Z. WU, Y. XING, F. ZHENG, X. GUO, W. ZANG, AND M. HONG. 1990. Nucleotide sequence of rice *waxy* gene. *Nucleic Acids Research* 18: 5898.

WHITSON, M. K., AND P. MANOS. 1999. Phylogenetic utility of sequence data from two regions of nuclear DNA in clarifying sectional relationships within *Physalis*. Abstracts, XVI International Botanical Congress, Saint Louis, Missouri, USA.

WOLFGANG, S., AND C. H. LANGLEY. 1998. DNA polymorphism in *Lycopersicon* and crossing-over per physical length. *Genetics* 150: 1585–1593.