Systematic Botany (2009), 34(1): pp. 207–219

Nitrate Reductase Phylogeny of Potato (*Solanum* sect. *Petota*) Genomes with Emphasis on the Origins of the Polyploid Species

Flor Rodríguez and David M. Spooner[1]

USDA, Agricultural Research Service, Vegetable Crops Research Unit, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison Wisconsin 53706-1590 U.S.A.
[1]Author for correspondence (david.spooner@ars.usda.gov)

Communicating Editor: Mark P. Simmons

Abstract—*Solanum* section *Petota* is taxonomically difficult, partly because of interspecific hybridization at both the diploid and polyploid levels. There is much disagreement regarding species boundaries and affiliation of species to series. Elucidating the phylogenetic relationships within the polyploids is crucial for an effective taxonomic treatment of the section and for the utilization of wild potato germplasm in breeding programs. We here infer relationships among the potato diploids and polyploids using nitrate reductase (*NIA*) sequence data in comparison to prior plastid phylogenies and: 1) examine genome types within section *Petota*, 2) show species in the polyploid series *Conicibaccata*, *Longipedicellata*, and in the Iopetalum group to be derived from allopolyploidization, 3) support an earlier hypothesis by confirming *S. verrucosum* as the maternal genome donor for the polyploid species *S. demissum* as well as species in the Iopetalum Group, 4) demonstrate that *S. verrucosum* is the closest relative to the maternal genome donor for species in ser. *Longipedicellata*, 5) support the close relationship between *S. acaule* and diploid species from series *Megistacroloba* and *Tuberosa*, and 6) show the North and Central American B genome species to be well distinguished from the A genome species of South America.

Keywords—allopolyploid, autopolyploid.

Polyploidy is an important mode of speciation. It can result from either duplication of a single genome (autopolyploidy) or from the combination of two or more differentiated genomes (allopolyploidy) (Masterson 1994; Ramsey and Schemske 1998; Soltis and Soltis 2000; Soltis et al. 2003). Almost all angiosperm species are believed to have polyploids in their ancestry (Masterson 1994; Soltis and Soltis 2000; Wendel 2000), and a better understanding of the processes of polyploidization and rapid diversification of the descendants of a single polyploidization event are of evolutionary interest (Wendel 2000; Soltis et al. 2003). Molecular markers have revealed that polyploid formation is a complex process, often characterized by extensive reticulation and lineage sorting that may not be fully captured by standard concepts of auto- vs. allopolyploidy (Wolfe 2001; Soltis et al. 2003).

Genome composition and repatterning after polyploidization is an important area of polyploid research. Kihara (1944) developed the conventional genome analysis methodology to determine the genome donors of allopolyploid species. This methodology involves the resynthesis, via artificial sexual hybridization, of a polyploid species with its presumed progenitor species followed by analysis of the chromosome pairing behavior in the hybrid. However, chromosomal pairing in hybrids can sometimes be dramatically influenced by only a few genes (Sears 1976). In addition, it is difficult to distinguish inter and intragenomic chromosomal pairing using conventional cytogenetic methods. Therefore, erroneous conclusions can be drawn because of the different interpretations of chromosome pairing. Structural chromosome rearrangements within members of sect. *Petota* have not yet been reported with convincing evidence, although they have been reported in close outgroups *Solanum* sects. *Etuberosum* and *Lycopersicon* (Tanksley et al. 1992; Perez et al. 1999) (see Table 1 for authors of taxonomic names).

An important evolutionary mechanism in sect. *Petota* is "endosperm balance number" (EBN), a strong crossing barrier that results in death of the endosperm in EBN-incompatible crosses. *Solanum* species have been assigned EBN numbers based on their ability to hybridize with each other (Johnston et al. 1980; Ortiz and Ehlenfeldt 1992; Hanneman 1994). Barring other crossing barriers, successful hybridization is expected when male and female gametes have matching EBN values, regardless of ploidy. The diploid chromosome number ($2x$) in sect. *Petota* is 24, so chromosome numbers vary from 24–72; in addition, there are rare occurrences of triploid and pentaploid populations. Ploidy (EBN) combinations in potato include $2x$(1EBN), $2x$(2EBN), $4x$(2EBN), $4x$(4EBN), and $6x$(4EBN). The consistency of EBN in predicting crossing success was evidenced by the fact that only one EBN value could be assigned to any species and that autotetraploidy or haploidy doubled or halved, respectively, the EBN class with which a given species could successfully hybridize (Johnston and Hanneman 1980, 1982). As a result, $2n$ gametes in the tuber-bearing potatoes provide opportunities for gene flow between species with different ploidy levels and/or different EBN levels (Den Nijs and Peloquin 1977).

The genome concept for potato species has been developed based on crossability success in interspecific combinations, hybrid sterility, hybrid viability, pollen fertility and in the degree of chromosomal homology (Marks 1955, 1965; Hawkes 1958, 1990; Irikura 1976; Ramanna and Hermsen 1981; López and Hawkes 1991; Matsubayashi 1991). Chromosome pairing relationships in interspecific hybrids and in polyploid species have been interpreted by genome formulas, although authors gave them different symbols. According to this hypothesis, five genomes (A, B, C, D, and P) are recognized in the tuber-bearing species of section *Petota*. Genome symbol E was given to species of section *Etuberosum* based on the specificity of meiotic behavior and sterility of their diploid hybrids with the A-genome tuber-bearing potato species (Ramanna and Hermsen 1979, 1981). Symbol L was proposed for tomato (section *Lycopersicon*) on the bases of preferential chromosome pairing and clear-cut parental genome discrimination by using genomic in situ hybridization (GISH) and amphidiploids of the LLEE type between tomato and *S. etuberosum* (Gavrilenko et al. 2001).

Hawkes (1990) divided the tuber-bearing potatoes in 19 series distributed among two superseries based on corolla morphology (superseries *Stellata* with stellate corollas and superseries *Rotata* with rotate corollas). He distinguished

TABLE 1. Species in *Solanum* section *Petota* and outgroup section *Etuberosum* analyzed in this study. AR=advanced rotate, PA=primitive rotate, AS=advanced stellate, PA=primitive stellate, nd=not determined.

Section, series (ser. abbreviation)	Species	Superseries	PI	Country	Ploidy	EBN	Genomes (Matsubayashi 1991)	Genomes (Hawkes 1990)	Plastid clade	NCBI sequence database accession numbers
Section *Petota* Dumort.										
Acaulia Juz. (ACA)	*Solanum acaule* Bitter	AR	310923	Bolivia	4*x*	2	AAA^aA^a	$A_2A_2A_3A_3$	4	FJ155386–87
Bulbocastana Rydb. Hawkes (BUL)	*S. bulbocastanum* Dunal	PS	347757	Mexico	2*x*	1	A^bA^b		2	FJ155402
	S. clarum Correll	PS	275202	Guatemala	2*x*	nd	A^bA^b		1	FJ155406
Conicibaccata Bitter (CON)	*S. agrimonifolium* Rydb.	AR	243349	Guatemala	4*x*	2	$A^{c1}A^{c1}C^aC^a$		4	FJ155388–91
	S. chomatophilum Bitter	AR	310991	Peru	2*x*	2	$A^{c2}A^{c2}$		3	FJ155405
	S. colombianum Bitter	AR	473462	Venezuela	4*x*	2			4	FJ155407–409
	S. longiconicum Bitter	AR	186568	Costa Rica	4*x*	nd	$A^{c1}A^{c1}C^1C^1$		4	FJ155432–33
	S. violaceimarmoratum Bitter	PR	473396	Bolivia	2*x*	2			unknown	FJ155474
Cuneoalata Hawkes (CUN)	*S. infundibuliforme* Phil.	PR	472857	Argentina	2*x*	2	AA		4	FJ155426–27
Demissa Buk. (DEM)	*S. iopetalum* (Bitter) J. G. Hawkes	AR	275183	Mexico	6*x*	4			unknown	FJ155397–98
	S. demissum Lindl.	AR	558482	Mexico	6*x*	4	$AADDD^dD^d$	A^1A^4[B,C,D]	4	FJ155410–15
	S. schenckii Bitter	AR	275261	Mexico	6*x*	nd			4	FJ155455–60
Lignicaulia Hawkes (LIG)	*S. lignicaule* Vargas	PS	473351	Peru	2*x*	1			4	FJ155431
Longipedicellata Buk. (LON)	*S. hjertingii* Hawkes	AR	186559	Mexico	4*x*	2			4	FJ155423–24
	S. stoloniferum Schltdl. A	AR	497994	Mexico	4*x*	2	AABB			FJ155418–21
	S. stoloniferum B	AR	558453	Mexico	4*x*	2	AABB		4	FJ155463–66
Megistacroloba Cárdenas and Hawkes (MEG)	*S. boliviense* Dunal	PR	597736	Bolivia	2*x*	2	AA		4	FJ155395
	S. megistacrolobum Bitter	PR	545979	Bolivia	2*x*	2	AA		4	FJ155434
	S. raphanifolium Cárdenas and Hawkes	PR	265862	Peru	2*x*	2	AA		4	FJ155453–54
Morelliformia Hawkes (MOR)	*S. morelliforme* Bitter and G. Muench	PS	275218	Mexico	2*x*	nd	A^mA^m		1	FJ155436
Pinnatisecta (Rydb.) Hawkes (PIN)	*S. stenophyllidium* Bitter	PS	255527	Mexico	2*x*	1			1	FJ155396
	S. cardiophyllum Lindl.	PS	595465	Mexico	2*x*	1	$A^{pi}A^{pi}$		2	FJ155403
	S. ehrenbergii (Bitter) Rydb.	PS	184762	Mexico	2*x*	1			1	FJ155416–7
	S. jamesii Torr.	PS	458424	US	2*x*	1	$A^{pi}A^{pi}$		1	FJ155428
	S. pinnatisectum Dunal	PS	347766	Mexico	2*x*	1	$A^{pi}A^{pi}$		1	FJ155451
	S. trifidum Correll	PS	255536	Mexico	2*x*	1			1	FJ155468
Piurana Hawkes (PIU)	*S. albornozii* Correll	AR	498206	Ecuador	2*x*	nd			3	FJ155392
	S. paucijugum Bitter	AR	561643	Ecuador	4*x*	nd			3	FJ155447–50
	S. tuquerrense Hawkes	AR	567849	Ecuador	4*x*	2	A^PA^PPP		3	FJ155469–72
Polyadenia Correll (POL)	*S. polyadenium* Greenm.	PS	161728	Mexico	2*x*	nd	$A^{po}A^{po}$		1	FJ155452
	S. *lesteri* Hawkes and Hjert.	PS	442694	Mexico	2*x*	nd			1	FJ155430
Tuberosa (Rydb.) Hawkes (TUB)	*S. andreanum* Baker	AR	320345	Colombia	2*x*	nd		A_1A_1	3	FJ155393
	S. berthaultii Hawkes	PR	265857	Bolivia	2*x*	2		A_1A_1	4	FJ155394

Section, series (ser. abbreviation)	Species	Superseries	PI	Country	Ploidy	EBN	Genomes (Matsubayashi 1991)	Genomes (Hawkes 1990)	Plastid clade	NCBI sequence database accession numbers
	S. brevicaule Bitter	PR	498115	Bolivia	2*x*	2		A_1A_1	4	FJ155399–400
	S. bukasovii Juz.	PR	266385	Peru	2*x*	2	AA	A_1A_1	4	FJ155401
	S. gandarillasii Cárdenas	PR	265866	Bolivia	2*x*	2		A_1A_1	4	FJ155422
	S. immite Dunal	PR	365330	Peru	2*x*	nd		A_1A_1	3	FJ155425
	S. leptophyes Bitter	PR	458378	Peru	2*x*	2	AA	A_1A_1	4	FJ155429
	S. microdontum Bitter	PR	500036	Argentina	2*x*	2	AA	A_1A_1	4	FJ155435
	S. oplocense Hawkes A	PR	435079	Argentina	6*x*	4			4	FJ155437–42
	S. oplocense B	PR	442693	Bolivia	4*x*	4			4	FJ155443–46
	S. sparsipilum (Bitter) Juz. and Buk.	PR	310957	Peru	2*x*	2	AA	A_1A_1	1	FJ155461–62
	S. verrucosum Schltdl.	AR	161173	Mexico	2*x*	2	AA	A_1A_1	4	FJ155473
Yungasensa Correll (YUN)	*S. chacoense* Bitter	AS	275138	Argentina	2*x*	2	AA		4	FJ155404
	S. tarijense Hawkes	AS	442689	Argentina	2*x*	2			4	FJ155467
Section *Etuberosum* (Buk and Kameraz) A. Child (ETU)	*S. etuberosum* Lindl.	–	498311	Chile	2*x*	1	E^eE^e			FJ155385
	S. palustre Peopp.	–	558233	Chile	2*x*	1	E^bE^b			FJ155384
Section *Lycopersicon* (Miller) Wettstein	*S. lycopersicum* L.		GenBank Accession X14060		2*x*					

"primitive" and "advanced" forms of each superseries and recognized four groups based on morphological characters: primitive *Stellata*, advanced *Stellata*, primitive *Rotata*, and advanced *Rotata*. He hypothesized the evolution of the advanced *Rotata* morphology from primitive *Stellata*-like ancestors. Superseries *Stellata* consists of nine diploid series with putatively ancestral stellate corollas while superseries *Rotata* consists of ten series of varying ploidies, all with pentagonal to rotate corollas. He postulated that the ancestral wild potato species were 2*x*(1EBN), possessed B genomes, produced white stellate corollas, and originated in North or Central America in the late Cretaceous to Eocene eras. Subsequent dispersal of one or more of these species to South America took place before the sinking of the Central American land bridge during the mid-Eocene to Pliocene eras, gradually leading to the evolution of species with A genomes, rotate corollas, and species with 2EBN. He also postulated that when the bridge was restored in Pliocene times, a remigration of one or more of these diverged A genome species back into North and Central America allowed the hybridization and allopolyploidization with the native Mexican or Central American B genome taxa. This produced the tetraploid members of series *Longipedicellata* (AB). A second, comparatively recent migration of a second species (*S. verrucosum*, A genome) from South America formed the *Demissa* hexaploids (A_1A_4[B,C,D]); Hawkes 1990) by crossing with series *Longipedicellata* tetraploids and possibly other series. In North and Central America are another group of allopolyploids, members of series *Conicibaccata* (AC), but only series *Longipedicellata* was designated AB, and the source of the C and D genome donors of series *Conicibaccata* and *Demissa* is unknown.

Even though Hawkes did not specifically identify the B genome in any diploid species, he suggested that it be placed in the North-Central American diploid species (with the exclusion of *S. verrucosum*) when he hypothesized that the ancestral wild potato species were 2*x*(1EBN), produced white stellate corollas and possessed B genomes. Contrasting Hawkes (1990), Matsubayashi (1991) recognized only one genome group in all diploid tuber-bearing species (only genome A) after he observed a regular meiosis in the five diploid hybrid clones between *S. bulbocastanum* and *S. verrucosum*. This suggested that *S. bulbocastanum* possesses a genome similar to the *S. verrucosum* genome which was earlier designated as A. Then, he concluded that the rest of the Mexican diploids species in series *Morelliformia*, *Pinnatisecta*, and *Polyadenia* possess the same genome B because they are genomically closely related to *S. bulbocastanum*.

The presence of B, C, D, or P genomes in polyploid species was inferred by observing univalents in triploid artificial hybrids resulting from crosses of tetraploid species and diploid species of various putative genomes (Hawkes 1990; Matsubayashi 1991). For instance, the triploid hybrids between *S. stoloniferum* (4*x*, AABB) and *S. chacoense* (2*x*, AA) showed a strong tendency toward the formation of 12 bivalents and 12 univalents; the bivalents are attributed to homology between one of the two *stoloniferum* genomes and the *chacoense* genome (known as A), and the univalents arise from another one of the *stoloniferum* genomes that was named B (Matsubayashi 1955). This hypothesis was also supported by the meiotic behavior of the diploid plant of *S. stoloniferum* and in the tetraploid hybrids combining the two genomes of *S. stoloniferum* and the two identical genomes (AA) from an induced tetraploid *S. chacoense* (Matsubayashi 1991).

Segmental allopolyploidy has been proposed for *S. acaule* (AAA^aA^a) by comparing the frequency of multivalent formation at metaphase I in the species and its haploids

(Matsubayashi 1991). Gavrilenko (2007) also proposed segmental allopolyploidy for the tetraploid species *S. tuquerrense*, although Matsubayashi (1991) considered it as a strict allotetraploid (A^pA^pPP). However, the observation of a high frequency of trivalents at metaphase I in a triploid hybrid (AA^pP) of *S. tuquerrense* and *S. verrucosum* (AA) (Marks 1965) indicates partial homology of the A^p and P genomes.

Spooner et al. (1993) used plastid DNA restriction site data and morphological data to reinvestigate the relationships of potatoes, tomatoes, further outgroups in *Solanum*, and other genera of the Solanaceae. Their results confirmed placement of all of Hawkes's (1990) tuber-bearing species into sect. *Petota*, but excluded the non-tuber-bearing species, in contrast to the treatment of Hawkes (1990). Subsequent plastid DNA phylogenies documented section *Petota* (tuber-bearing) to be divided into four clades that often showed little relationship to Hawkes's taxonomic series (Spooner and Sytsma 1992; Rodríguez and Spooner 1997; Spooner and Castillo 1997). These clades contain: 1) North and Central American diploid species, with the exception of *S. bulbocastanum*, *S. cardiophyllum*, and *S. verrucosum*; 2) *S. bulbocastanum* and *S. cardiophyllum*; 3) all examined members of series *Piurana* and some South American species belonging to other series (*S. andreanum, S. chancayense*, *S. immite*, and *S. mochiquense* of ser. *Tuberosa*; *S. chomatophilum*, and *S. tundalomense* of ser. *Conicibaccata*; *S. huancabambense* of ser. *Yugansensia*; *S. sogarandinum* of ser. *Megistacroloba*); 4) all remaining South American species, *S. verrucosum* from Mexico, and North and Central American polyploid species. The interpretation of maternal gene trees (derived from plastid DNA data) needs to be corroborated with nuclear markers to infer progenitors of the polyploid species (Wendel and Doyle 1998).

Recent morphological and molecular studies of the polyploids showed phylogenetic results that bear upon taxonomic terminology used here. The tetraploid *S. demissum* was shown to be related to the South American tetraploid species *S. acaule* and *S. albicans*, not to other members of series *Demissa* (Spooner et al. 1995; Kardolus et al. 1998; Kardolus 1999; Nakagawa and Hosaka 2002). Spooner et al. (2004) used these results to place *S. acaule*, *S. albicans*, and *S. demissum* into an informal Acaulia group, and the other members of series *Demissa* (*S. hougasii*, *S. iopetalum*, and *S. schenckii*) in an informal Iopetalum Group. We use the terms Acaulia and Iopetalum Groups in the text, but show Hawkes's (1990) series classifications in Table 1.

Until recently, plastid DNA and nuclear ribosomal DNA have provided the major datasets for phylogenetic inference in flowering plants because of the ease of obtaining data due to high copy number. Recently, single- to low-copy nuclear DNA markers have been developed as powerful new tools for phylogenetic analyses (Wolfe et al. 1987; Doyle and Doyle 1999; Cronn et al. 2002; Sang 2002; Small et al. 2004). An ideal low-copy nuclear marker should maximize phylogenetic information but minimize homoplasy at the taxonomic level studied. Introns tend to provide better resolution for relationships among recently diverged species or within species (Tank and Sang 2001) because they evolve faster than exon regions and usually accommodate more insertion and deletion mutations (Sang 2002).

Low-copy markers have successfully revealed the origins and evolutionary history of polyploids, clarified their genomic nature, indicated their parental lineages, and identified hybridization events involved in their formation (Soltis and Soltis 1993; Wendel 2000; Soltis et al. 2003). Reconstruction of the evolutionary origins of allopolyploids using nuclear markers is relatively straightforward because homoeologous loci that are contributed by the diploid parents can be cloned and sequenced from an allopolyploid species. Analyses of these sequences, together with the gene sequences of the putative diploid parents, convert the reconstruction of reticulate evolution into the reconstruction of the diverged histories of parental lineages (Small et al. 1998; Cronn et al. 1999; Sang and Zhang 1999).

Nitrate reductase (*NIA*) is a potential single-copy nuclear marker useful for phylogenetic purposes. It catalyzes NAD(P)H reduction of nitrate to nitrite, which is the first step of nitrate assimilation. The gene expressing this protein has been isolated from fungi, algae, and higher plants. *NIA* has three intron regions and has been shown to be single-copy in some plants, where it is useful for low-level phylogenetic analyses. *NIA* has been used in phylogenetic studies for *Betula* (Li et al. 2007), *Boerhavia* (Douglas 2007), *Memecylon* (Stone et al. 2006), and *Scaevola* (Howarth and Baum 2002). In this study, we used nitrate reductase sequence data from an intron of *NIA* and its flanking regions to investigate the phylogeny of *Solanum* sect. *Petota* (A, B, C, D genomes), using outgroups from sect. *Etuberosum* (E genome) and sect. *Lycopersicon* (L genome). We re-examine cladistic relationships that were based on plastid data and investigate the species contribution and/or possible genome differentiation in polyploid species in light of previous hypotheses based on morphology and cytogenetic data.

Materials and Methods

***Plant Materials and DNA Isolation*—**Forty-five ingroup accessions, representing 43 ingroup taxa from 14 series of Hawkes (1990) belonging to all four plastid DNA clades of sect. *Petota* were examined. These species were chosen to cover most of the series, EBN, and ploidy levels (Hawkes 1990), plastid DNA clades (Castillo and Spooner 1997; Spooner and Castillo 1997), and genomic groups (Matsubayashi 1991). Three outgroup species, *S. etuberosum* and *S. palustre* that belong to sect. *Etuberosum* (Table 1) and a published sequence from *S. lycopersicum* (GenBank Accession X14060) (sect. *Lycopersicon*) were also included. Herbarium vouchers are deposited at PTIS. Total DNA from each species was isolated using a standard CTAB protocol (International Potato Center 1999). DNA quality and quantity were estimated by comparison with CsCl-purified λ DNA digested with *Pst*I on ethidium bromide-stained agarose gels.

***Amplification, Cloning and Sequencing*—**A GenBank nitrate reductase tomato sequence (Accession X14060) was used to design primers flanking intron 3, for a total sequence of 1246 bp (436 bp of intron 3, 132 bp of exon 3 and 678 bp of exon 4). PCR amplifications were performed using AccuPower™ PCR PreMix (Bionexus Inc, Oakland, California) in 20 µL reactions consisting of 0.6 µM final concentration of each primer (NIA-3F: 5′ AAG TAC TGG TGT TGG TGC TTT TGG TC 3′, and NIA-R: 5′ TAT GTG TTT GCC AAC AGG TA 3′) and 50 ng of template genomic DNA. Amplifications were carried out in an MJ Research DNA Engine Dyad® Peltier Thermal Cycler (Watertown, Massachusetts) using an initial denaturation at 95°C for 3 min, followed by 30 cycles of 95°C for 45 s, 62°C for 2 min, 72°C for 2 min and with a final elongation at 72°C for 20 min to maximize A-tailing and increase efficiency of cloning into T-tailed vectors. The reactions were run on a 1.5% agarose gel with 1× TBE buffer for 5 hr. Bands were cut out and cleaned using Wizard®SV Gel Clean-Up System (Promega, Madison, Wisconsin) and resuspended in 50 µL of nuclease-free water.

Cleaned PCR products were cloned into Promega's pGEM-T Easy vector. Ligation, transformation, and plating were carried out following the manufacturer's instructions except that ligation and transformation reaction volumes were halved. Five positive colonies were sampled and cultured for diploid species, 10 for tetraploid and 20 for hexaploid species. Plasmid DNA was isolated using Wizard® Plus SV Minipreps DNA Purification Systems (Promega) and then was digested with *Eco*RI and checked on 1% agarose gels for the presence of inserts and to estimate concentrations.

Plasmids with the correctly sized insert were sequenced with 5 pmol of SP6 and T7 vector-specific primers in a 10 µL reaction using the ABI Big Dye dideoxynucelotide termination kit (Applied Biosystem, Foster City, California). Amplifications were carried out in an MJ Research DNA Engine Dyad® Peltier Thermal Cycler using an initial denaturation at 95°C for 3 min, followed by 30 cycles of 96°C for 25 s, 50°C for 20 s, 60°C for 5 min and with a final elongation at 72°C for 7 min. Excess dye terminators were removed using CleanSeq magnetic bead sequencing reaction clean up kit from Agencourt Biosciences (Beverly, Massachusetts). Sequences were resolved on an ABI 3730xl capillary-based automated DNA sequencer (Applied Biosystems) with 50 cm POP-7 polymer capillaries at the Biotechnology Center of the University of Wisconsin-Madison.

***Sequence Editing and Alignment*—**Sequences were edited with Staden package version-2003.0-beta (Staden 1996) and aligned using ClustalX version 1.81 (Thompson et al. 1997) at default parameters, except for the "percentage of delay divergence sequences" which was set to 15% after tests of various parameters. Further manual alignments were done in MacClade 4.06 PPC (Maddison and Maddison 2001) minimizing the number of gaps and preferring transitions instead of transversions. Indels were scored by the simple gap scoring method (Simmons and Ochoterena 2000) using Gap Recorder web interface (http://maen.huh.harvard.edu:8080/services/gap_recoder).

In order to identify pseudogenes a tomato nitrate reductase sequence (accession X14060) was added to the data set and exon sequences were translated in MacClade. DNA sequences were deposited in GenBank (Table 1), and the aligned dataset with gap codes is deposited in TreeBASE (study number S2214).

***Phylogenetic Analyses*—**Analyses were first done with the diploids and a second analysis was performed with all diploids and polyploids to infer cladistic contributors to the polyploids. Phylogenetic analyses based on maximum parsimony (MP) were performed using PAUP* 4.0b10 (Swofford 2002). The most parsimonious trees were found by heuristic searches under Fitch criteria and with equal weight for all characters. A heuristic search was performed using TBR branch swapping on one million random taxon addition sequences. A rooted strict-consensus tree was obtained using *S. etuberosum* and *S. palustre* as outgroups based on Spooner et al. (1993) and making those a monophyletic sister group to the ingroup. Bootstrap support (Felsenstein 1985) was estimated with 1,000 bootstrap replicates performing a TBR branch swapping on 100,000 random taxon addition sequences per replicate.

Maximum likelihood (ML; Felsenstein 1973) analysis was used to assess 16 models of sequence evolution representing the 16 pairwise combinations between four models of nucleotide substitution and four models of among-site rate variation. The four nucleotide substitution models tested were: the Jukes-Cantor model (Jukes and Cantor 1969), the Kimura two-parameter model (Kimura 1980), the Hasegawa-Kishino-Yano model (Hasegawa et al. 1985), and the general time reversible model (Tavaré 1986). The four models of among-site rate variation tested were (1) no variation among sites; (2) some proportion of sites, estimated via ML, being allowed to be invariable (I; Hasegawa et al. 1985); (3) rate variation being allowed to follow a discrete approximation to a gamma distribution (Γ; Yang 1994) with a shape parameter estimated using ML; and (4) some sites being allowed to be invariable, with the possibility of rate variation at the remaining sites following a gamma distribution (I + Γ; Gu et al. 1995; Waddell and Penny 1996). All parameters were estimated on one of the most parsimonious trees found before. The choice of trees used in parameter estimations should not be critical as long as they represent reasonable representations of the data and parameters do not fluctuate greatly from one tree to the next (Swofford et al. 1996). Likelihood scores between nested models of sequence evolution were compared using a likelihood ratio test (Felsenstein 1988) to determine the model that best fits the data.

A maximum likelihood heuristic search was conducted using a genetic algorithm as implemented in GARLI 0.951 (Zwickl 2006). The run was repeated three times from random starting trees using the default settings for the components of the genetic algorithm but the default termination conditions were changed allowing the software to terminate the run after 300,000 generations if no better scoring topology was encountered. A GTR + Γ + I model was used. To evaluate the stability of clades on the optimal tree, a bootstrap analysis was performed with 100 bootstrap replicates performing the same search, under the best model of evolution with parameters fixed.

Bayesian Markov chain Monte Carlo (MCMC) phylogenetic analysis (Yang and Rannala 1997) also was performed using MrBayes version 3.1 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003), as an independent measurement of phylogenetic relationship and clade composition. The GTR + Γ + I model for DNA data was chosen according to the criteria described above and the restriction site (binary) model for gaps. Model parameters were estimated separately for DNA and gaps; for gaps the restriction site (binary) model was used "lset coding=variable" which accounts for the ascertainment bias produced by characters that are constant (either state 0 or 1) in all taxa and are not observed. Tree searching using MrBayes was performed by running 6 linked chains, initiated from random trees with a sequential heat of 0.05 (determined empirically) for five million generations with trees sampled every 100 generations. At the end of the run, convergence was evaluated by visual inspection of a graph of likelihood as a function of generation. A conservative burn-in period was determined, and only postburn-in trees were saved. This analysis was conducted three times, and the majority-rule consensus trees from each run were compared to evaluate mixing. The three sets of postburn-in trees were then combined to form a majority rule consensus tree, and this pool was taken as the best representation of the posterior distribution of tree topology and model parameters (Huelsenbeck and Ronquist 2001). The proportion of searches in which any given node is found during the post burn-in portion of the chain constitutes the Bayesian posterior probability for that node.

***Hypothesis Testing*—**We tested 1) if phylogenetic trees reflect EBN, ploidy, genome designations or geographical patterns, 2) the monophyly of Hawkes's taxonomic series, 3) the monophyly of two divergent alleles in diploid species, 4) the monophyly of all North and Central American diploid species except *S. verrucosum*, 5) the phylogenetic placement of the North and Central American diploids exclusive of *S. verrucosum* as sister to the other tuber-bearing potatoes, 6) the monophyly of all South American diploid species, 7) the placement of *S. verrucosum* with all South American diploid species, 8) the allopolyploid origin of series *Longipedicellata*, *Conicibaccata,* and the Iopetalum group, 9) the autopolyploid origin of polyploids in series *Tuberosa*, *Piurana,* and Acaulia Group, and 10) the single origin of allopolyploids.

All hypotheses were examined under MP and Bayesian approaches. The Templeton test (Templeton 1983) was performed under MP. In a Bayesian framework, the trees found after stationarity were filtered to find those consistent with the constraint tree and the number of these trees was recorded. This number was divided by the total number of poststationarity trees to calculate the posterior probability of the hypothesis represented by the constraint tree. This value was used to calculate the *p* value for each hypothesis (C. Ané, pers. comm.).

Results

***Sequence Alignment and Variability*—**Complete *NIA* sequences (no missing data) of the third intron and flanking exons in the diploid data set varied from 1,168 bp in *S. clarum* to 1,327 bp in the two outgroups; regarding only the ingroup taxa, *S. microdontum* was the longest with 1,301 bp. Of the 34 diploid species analyzed, 29 had a single sequence type, and five had two sequence types. The two sequences of *S. ehrenbergii* and *S. raphanifolium* resolved with their respective species. In the case of *S. infundibuliforme*, one allele resolved with *S. berthaultii* and *S. tarijense* and the other with *S. brevicaule*, all in subclade 4a (Fig. 1). On the other hand, one allele of *S. brevicaule* and *S. sparsipilum* were placed in the subclade 4a and the other in the subclade 4b.

The average GC content was 39%, varying from 44–30% in exon and intron regions, respectively. No stop codons were found when exons were translated. The aligned DNA data matrix for diploids was 1,442 characters long. The percentage of gaps varied from 8% in the two outgroups (*S. etuberosum* and *S. palustre*) to 19% in *S. clarum*; their scores added an additional 78 characters, which makes a total of 1,520 characters. From those, 355 were variable and 151 parsimony informative, 116 of 151 were nucleotide substitutions and 35 were gap characters (Table 2).

In the data set with all diploids and polyploids, the length of the sequences ranged from 1,159 bp in one allele from *S. stoloniferum* to 1,327 bp in outgroups; within polyploids the longest sequence was found in one allele of *S. agrimonifolium* and *S. colombianum* with 1,298 bp. All 12 polyploid taxa analyzed had more than one allele. As in the diploids, polyploids

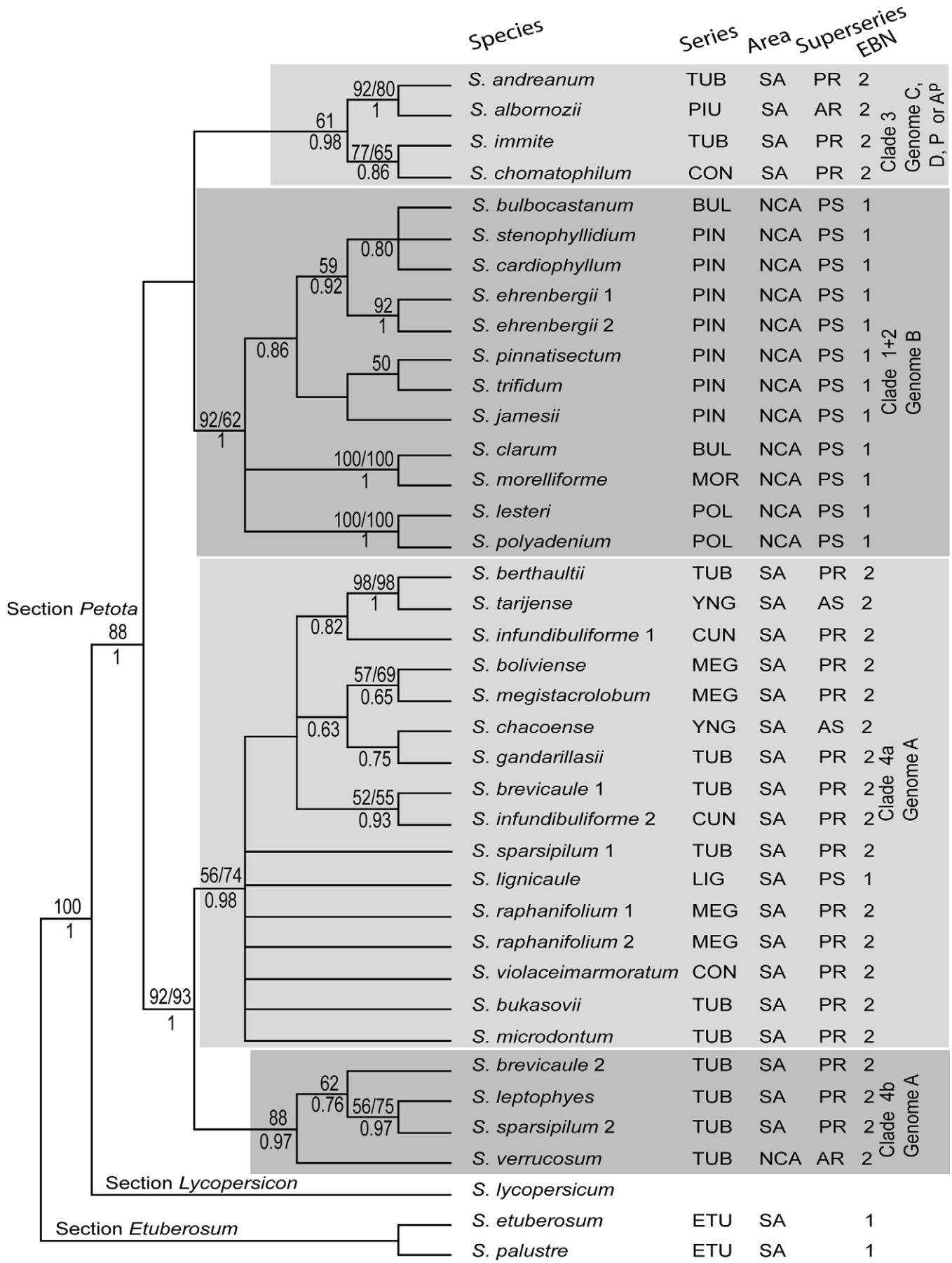

FIG. 1. Strict consensus of 104 most parsimonious trees for diploid species (L = 5112, CI = 0.764, RI = 0.779; DNA and indel characters included). *Solanum etuberosum* and *S. palustre* were used as outgroups. Bootstrap values higher than 50% are indicated above branches, the first value refers to Maximum Parsimony and the second to Maximum Likelihood analyses; below branches are the posterior probability values. Numbers after the species name indicate allelic variants. Clade designations follow those of the plastid phylogenies (Spooner and Sytsma 1992; Rodríguez and Spooner 1997; Spooner and Castillo 1997) differing here only in uniting clades 1 and 2 and separating clade 4 into 4a and 4b. Series, superseries, and EBN (Endosperm Balance Number) codes or designations follow Table 1; area codes are SA = South America and NCA = North and Central America.

showed the same GC content, no stop codons were found, but the aligned DNA data matrix contained 1,445 characters. Indels were frequent in the intron region, and ranged from 8% (in outgroups) to 20% (in the shortest allele of *S. stoloniferum*) of the total aligned DNA data matrix. Within the polyploids, the lowest gap percentage (10%) was found in the shortest alleles of *S. agrimonifolium* and *S. colombianum*. Gap scores added 87 additional characters, thus a total sum of 1,532 characters, of which 427 sites were variable and 216 were parsimony informative (174 from DNA sequences and 42 from gaps; Table 2).

***Model Selection*—**Likelihood models that account for rate variation, either by allowing for some proportion of invariant sites (I) or by allowing for gamma-distributed rate variation among sites (Γ), resulted in the greatest increase in ML scores in both data sets. Gamma-distributed rate variation was the single parameter that provided the most improvement. However, there was noticeable heterogeneity among substitution types when six types were considered, as in the case of the GTR models. The model that incorporated the most parameters, GTR + I + Γ, fit both data sets best. The log ML score for diploids was -4,621.523 and for the complete data set -5,905.884.

***Phylogenetic Results, Diploid Species*—**We first constructed a phylogeny of the diploid species to reveal relationships of the putative genomic donors for comparison to the results from the polyploid species. Maximum parsimony analysis produced 104 equally most-parsimonious trees with a length of 514 steps, a consistency index (CI; Kluge and Farris 1969) of 0.764 and retention index (RI; Farris 1989) of 0.779. A strict consensus tree with bootstrap values is shown in Fig. 1). Maximum likelihood analysis using the parameters estimated previously yielded a tree with a log-likelihood of -4,619.667 (Supplementary Fig. 1). The three independent Bayesian MCMC runs (each with two internal runs) mixed well as indicated by the convergence diagnostic (the standard deviation of split frequencies) which reached a value less than 0.005 (Tree shown in Supplementary Fig. 2).

To aid comparison with earlier phylogenetic results we refer to the clades as in the four-clade plastid phylogeny of Spooner and Castillo (1997). The MP (Fig. 1), ML, and Bayesian analyses identified three main clades (Supplementary Figs. 1, 2); the Bayesian 50% majority rule consensus tree placed them as a polytomy, while MP placed clade 4 sister to both clade 3 and 1 + 2 (named relative to the plastid phylogenies that separates these species into two clades; Spooner and Sytsma 1992; Rodríguez and Spooner 1997; Spooner and Castillo 1997), and ML placed clade 1 + 2 as sister to the rest, although in both cases with less than 50% support. Clades in ML were less supported than in MP and Bayesian analysis. This could be due to the fact that gaps were coded as binary characters and it allowed their use only in MP and Bayesian analyses, which added 35 parsimony informative characters.

Within clade 1 + 2, which is formed by North and Central American diploid species exclusive of *S. verrucosum*, MP and Bayesian topologies were nearly identical; the only difference was the placement of *S. trifidum* sister to *S. pinnatisectum* in the MP tree, and sister to *S. bulbocastanum*, *S. cardiophyllum*, *S. ehrenbergii,* and *S. stenophyllidium* in Bayesian. The ML differed from the other two in that it did not place the two alleles of *S. ehrenbergii* together, instead allele 1 was close to *S. stenophyllidium,* but only with < 50% support. The ML also placed *S. trifidum* as in the Bayesian analysis, but with only 51% bootstrap support. The four species in clade 3 (Piurana clade) resolved identically in all tree topologies. In clade 4, Bayesian and MP analyses identified two subclades (4a and 4b). Clade 4b had an identical topology in both analyses, but clade 4a showed much more resolution in Bayesian analysis. The ML produced a very well defined clade 4a and species are placed as in MP and Bayesian phylogenies, but species that belong to clade 4b were in a grade.

Given that there are not any ML programs in which mixed models are implemented that could allow the use of gap data in the analysis, MP and Bayesian analyses showed more resolution, and we base further discussion only on results with gap information included. The MP and Bayesian analyses strongly supported the monophyly of the North and Central American diploid species exclusive of *S. verrucosum* (clade 1 + 2, BS = 92, PP = 1.00). Trees found by forcing *S. bulbocastanum* and *S. cardiophyllum* to be separate from all North and Central diploids, and to group with South American diploid species (suggested by plastid phylogenetic study; Spooner and Castillo 1997), were significantly longer (29 steps) than the most parsimonious tree ($p < 0.0001$, defined as the probability of obtaining a more extreme t value with the two-tailed test under the null hypothesis of no difference between the two trees). The Bayesian tree was similar, in that none of the trees in the posterior distribution were consistent with the constraint. Therefore, the *NIA* sequence data rejected the possibility suggested by the plastid phylogenetic study.

Clade 1 + 2—The North and Central American diploid clade (1 + 2) has two well defined subclades with strong support in both MP and Bayesian phylogenies. One clade resolved *S. clarum* and *S. morelliforme*; they occur in similar habitats, with *S. morelliforme* an epiphyte, and *S. clarum* growing under the shade of trees in moss in an epiphytic like condition. Both species are difficult to cross to the other North and Central American diploids (Graham and Dionne 1961; Marks 1968). Another clade resolved *S. lesteri* and *S. polyadenium*; both species were also united by plastid-DNA restriction site data (Spooner and Sytsma 1992), morphological data (Lara-Cabrera and Spooner 2005) and AFLP data (Lara-Cabrera and Spooner 2004). They have densely pubescent leaves with the same type of glandular trichomes, share a characteristic mousy odor, and were placed by Hawkes (1990) in series *Polyadenia*. *Solanum jamesii* and *S. pinnatisectum* form a well supported clade in the Bayesian MCMC tree. These are the only species in sect. *Petota* that have pinnatifid pseudostipules (Lara-Cabrera and Spooner 2005) and were united by AFLP data (Lara-Cabrera and Spooner 2004).

Series *Pinnatisecta* is supported as monophyletic but only with the inclusion of *S. bulbocastanum* (series *Bulbocastana*), but with under 50% bootstrap support. However, trees found by forcing series *Pinnatisecta* to be monophyletic (excluding *S. bulbocastanum*) were only three steps longer than the MP trees, and the Templeton test failed to reject the monophyly of series *Pinnatisecta* ($p = 0.18$–0.53). Bayesian analyses found that 0.0003% of the trees in the posterior distribution were consistent with the topological predictions of this hypothesis (monophyly of series *Pinnatisecta*). Thus, the monophyly of series *Pinnatisecta* could not be rejected.

Forcing the two species of series *Bulbocastana* (*S. bulbocastanum* and *S. clarum*) to be monophyletic resulted in trees were 12 steps longer and statistically less likely (Templeton test, $p = 0.0013$–0.0339) than the MP tree without this constraint. Bayesian analysis also rejected the monophyly of series

Table 2. Summary of sequence variation among diploids and polyploids in the complete data set.

Data set	Region	Total characters	Variable characters	Number and percentage of parsimony informative characters
Diploids	exon	816	95	35 (4%)
	intron	626	182	81 (13%)
	DNA	1442	277	116 (8%)
	gaps	78	78	35 (45%)
	DNA & gaps	1520	355	151 (10%)
Diploids and polyploids	exon	816	129	64 (8%)
	intron	629	211	110 (18%)
	DNA	1445	340	174 (12%)
	gaps	87	87	42 (48%)
	DNA & gaps	1532	427	216 (14%)

Bulbocastana, because none of the trees in the posterior distribution were consistent with the constraint.

Given that the only difference between MP and Bayesian topologies in clade 1 + 2 was the placement of *S. trifidum*, the two topologies were tested. The Templeton test failed to reject the Bayesian topology (p = 0.74–0.83, Supplementary Fig. 2) and Bayesian analysis could not reject the MP topology either (PP = 0.26, (Fig. 1). Therefore, both scenarios are plausible, because the contradiction was not significant.

Clade 3—Clade 3 contained species from three of Hawkes's (1990) series: *S. albornozii* (*Piurana*), *S. andreanum* (*Tuberosa*), *S. immite* (*Tuberosa*) and *S. chomatophilum* (*Conicibaccata*). *Solanum andreanum*, from Colombia and Ecuador, is the most northern representative of clade 3. Hawkes (1990) placed *S. andreanum* in series *Tuberosa*, but its association with other members of series *Piurana* is strongly supported by plastid DNA (Spooner and Castillo 1997) and the present results. *Solanum chomatophilum* and *S. immite* are also placed inside of this clade, which is, again, supported by the plastid DNA data. To test if the placement of those species, belonging to other series, with the Piurana clade was due to stochastic noise, we forced *S. andreanum* and *S. immite* to group with the other ten species in series *Tuberosa* in clade 4. The Templeton test and Bayesian analysis rejected this possibility ($p < 0.0001$, PP = 0). The same result was found when *S. chomatophilum* was forced to resolve with *S. violaceimarmoratum*, another species in series *Conicibaccata* (p = 0.002–0.008, PP = 0). Furthermore, the posterior probability of clade 3 in the Bayesian phylogeny (Supplementary Fig. 2) was 0.96 and only 0.04% of the trees sampled were not in agreement in clade 3 to be monophyletic. Hence, there was strong evidence that those four species form a monophyletic group.

Clade 4—*Solanum verrucosum*, the only North and Central American diploid species with an A genome, was resolved with the A-genome South American diploid species in clade 4. It is the only North and Central American species that has a clear crossability relationship to the South American species with which it shares the same EBN level. Trees found by forcing *S. verrucosum* to fall in the North and Central American diploid clade (clade 1 + 2) were significantly longer (12 steps) than the most parsimonious tree. This constraint provided a significantly less likely explanation of the data than the best tree in MP (Templeton test, p = 0.0005–0.0282), with similar rejection of the hypothesis by Bayesian analysis (PP = 0). Furthermore, in the Bayesian phylogeny (Supplementary Fig. 2) clade 4 (with *S. verrucosum*) had a posterior probability of 1. Consequently, the placement of *S. verrucosum* elsewhere on the tree has a PP = 0, and *NIA* data strongly supported *S. verrucosum* as a member of clade 4.

There were four species in clade 4 with two alleles: *S. brevicaule*, *S. infundibuliforme*, *S. sparsipilum*, and *S. raphanifolium*. Alleles from *S. brevicaule* and *S. sparsipilum* were split between subclades 4a and 4b. When either *S. brevicaule* or *S. sparsipilum* alleles were forced to form a clade the resulting trees were significantly longer than unconstrained trees (12 and 13 steps longer respectively; p = 0.003–0.235 and p = 0.002–0.163, respectively) and the posterior probability of monophyly of these groups was zero in both cases. It suggested a possible hybrid origin of these taxa, which needs exploration with additional molecular markers.

The other two taxa that had two divergent alleles, *S. infundibuliforme* and *S. raphanifolium*, have alleles placed in the same subclade (4a), but each allele resolved with different species. Alleles from the same species were forced to group together and the shortest tree compatible with these constraints were only one step longer than the unconstrained trees, which is considered insignificant in the Templeton test (p = 0.71–0.78 and 0.706–0.808 respectively). Therefore, stochastic noise alone can explain this result and the possibility that those two alleles resolve together could not be rejected. However, Bayesian analysis showed that there is no evidence that those two alleles should group together. Additional molecular markers are needed to confirm these findings.

In clade 4a there are three series, *Megistacroloba*, *Tuberosa*, and *Yungasensa*, with more than one species per series. To test if the classification made by Hawkes (1990) is supported by *NIA* sequence data, species belonging to these series were forced to be monophyletic; Templeton ($p < 0.001$ in all three cases) and Bayesian analyses (PP = 0 in all three cases) strongly rejected this possibility.

To understand the evolutionary history of section *Petota*, factors known to affect hybridization, such as EBN, ploidy, geographical barriers and genome type were analyzed in the optimal phylogenetic trees in MP (Templeton test) and in Bayesian analysis. Within diploids, phylogenetic trees did not reflect EBN designations perfectly, but still showed a good fit (Fig. 1). It suggested that only one reversal from EBN 2–1 (*S. lignicaule*) is required to explain the distribution of EBN on the resulting trees. Additionally, species did not group based on Hawkes's four superseries (primitive *Stellata*, advanced *Stellata*, primitive *Rotata*, and advanced *Rotata*) based on flower shape. The Templeton test ($p < 0.016$) and Bayesian analysis (PP = 0) rejected all four of these groups. Two of these species, *S. berthaultii* and *S. tarijense*, were placed in putatively widely divergent superseries *Rotata* and *Stellata*, respectively, by Hawkes (1990). *NIA* results place them on an exclusive lineage, supporting a decision by Spooner et al. (2007), based on morphological and AFLP data, to synonymize both under *S. berthaultii*.

Phylogenetic Results, Polyploid Species—Of 1,532 characters, 427 were variable and 216 were parsimony informative. Parsimony analysis generated 9,482 most parsimonious trees (length = 712, CI = 0.680, RI = 0.846; see tree in Supplementary Fig. 3). Maximum likelihood analysis using the parameters estimated previously yielded a single tree (-ln likelihood = 5899.7467; Supplementary Fig. 4). Tree topologies resulting from three independent Bayesian runs were identical with similar clade posterior probabilities. The majority rule consensus tree of the pool of the three sets of post burn-in trees produced 42 clades receiving posterior probabilities of ≥ 0.95 (Fig. 2).

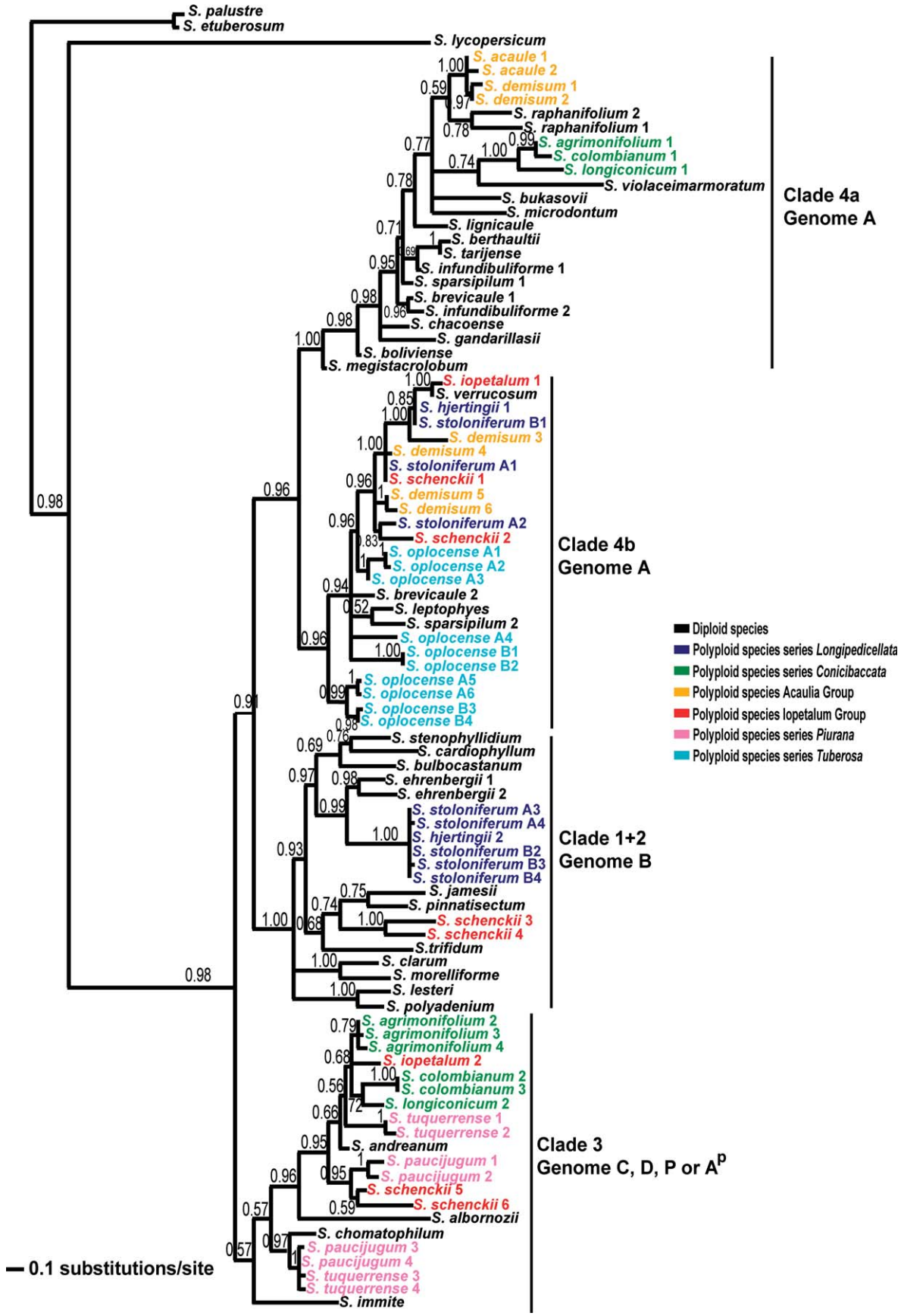

FIG. 2. Bayesian phylogram of all species, diploid and polyploid (DNA and indel characters included). Branch lengths are drawn in proportion to the estimated number of substitutions per site and represent an average of the branch length of all trees sampled in the Markov chain that have that branch. Posterior probability values are indicated for nodes with a probability value greater than 0.50. Numbers after the species name indicate allele variants. Letters after the species name indicate another accession for the same species.

The MP (Supplementary Fig. 3), ML (Supplementary Fig. 4), and Bayesian (Fig. 2) analyses found nearly identical species level relationships, but as with diploids, MP and Bayesian analyses showed more resolution. Clades 1 + 2 and 4 are very well differentiated in the three analyses. Clade 3 showed low posterior probability in Bayesian analysis, and bootstrap support less than 50% in MP and ML.

All of the polyploids showed consistent patterns of cladistic relationship depending on their taxonomic grouping as was also shown with *GBSSI* (*waxy*) sequence but with less resolution (Spooner et al. 2008) than the present *NIA* data. In all tetraploid species of series *Longipedicellata* two divergent alleles were identified, with one type placed in clade 1 + 2 and the other in clade 4b. Their alleles in clade 1 + 2 formed a strongly supported monophyletic clade, but their alleles in clade 4b did not (Fig. 2). *NIA* rejected the monophyly of alleles of series *Longipedicellata* in clade 4b (Templeton $p = 0.0047–0.0331$; PP = 0).

Tetraploid species of series *Conicibaccata* had alleles partitioning into clades 3 and 4a, but only the alleles in clade 4a were monophyletic with strong support. Trees forcing all their alleles in clade 3 forming a monophyletic clade were only one step longer than the unconstrained tree, and the Templeton test failed to reject their monophyly ($p = 0.763–0.827$). Furthermore, the constraint occurred in 25% of trees in the Bayesian posterior distribution. Thus, the monophyly of the alleles in clade 3 of series *Conicibaccata* could not be rejected.

Polyploid species of the Iopetalum group (6*x*) had clones in clade 1 + 2, 3 and 4b (*S. schenckii*) or in 3 and 4b (*S. iopetalum*). The two alleles of *S. schenckii* in clade 1 + 2 formed an exclusive lineage with strong support. But alleles of both species found in clades 3 and 4b failed to resolve. Trees forcing them to group were 26 and 15 steps longer, respectively, than the MP tree without the constraint. The Templeton test rejected the monophyly of all alleles of the Iopetalum group in clades 3 and 4b ($p = 0.0001–0.0004$ and $p = 0.0055–0.0333$, respectively). Furthermore, none of the trees in the posterior distribution were consistent with any of these two constraints. Hence, *NIA* data did not support the monophyly of their alleles in clades 3 and 4b.

Species of the Acaulia Group (*S. acaule* 4*x* and *S. demissum* 6*x*), had all clones exclusively in clade 4, the former in subclade 4a and the latter in 4a and 4b. Their alleles in clade 4a formed an exclusive lineage with strong support. But, the four alleles of *S. demissum* in clade 4b were not monophyletic, confirmed by the Templeton test ($p = 0.014–0.048$) and Bayesian analysis (PP = 0).

Discussion

As pointed out in the introduction, molecular phylogenetic analyses of sect. *Petota* often show little relationship to Hawkes's taxonomic series (Spooner and Sytsma 1992; Rodríguez and Spooner 1997; Spooner and Castillo 1997). *NIA* data only support the monophyly of series *Pinnatisecta* and *Polyadenia*. The monophyly of diploid species in series *Bulbocastana*, *Conicibaccata*, *Megistacroloba, Tuberosa*, and *Yungasensa* were rejected in MP and Bayesian analyses. On the other hand, the data show a clear relationship between *S. albornozii* (series *Piurana*), *S. chomatophilum* (series *Conicibaccata*), and *S. andreanum* and *S. immite* (series *Tuberosa*), which all grouped together in clade 3 (the Piurana clade). The close relationship of these species is also supported by most species having coriaceous and glossy leaves (Correll 1962; Hawkes 1990; Ochoa 1999) and moniliform tubers (Spooner and Salas 2006; Ames et al. 2008). Clearly, the taxonomy of section *Petota* is in need of revision.

In the diploid (Fig. 1) and polyploid (Fig. 2) trees, four (A, B, E, L) of the six genomes present in the three sections (*Etuberosum*, *Lycopersicum*, *Petota*) under study formed distinct and divergent clades. The two exceptions were the C and D genomes, which were mixed in clade 3 (Fig. 2). Genome E, represented by species of section *Etuberosum*, was sister to tomatoes and potatoes followed by genome L of *S. lycopersicum,* which was sister to potatoes. In section *Petota,* the two genomes A and B are well differentiated, with B in clade 1 + 2 and A in clade 4. These results indicate that *NIA* sequences of genomes derived from the diploid ancestors with genomes A and B have remained clearly differentiated in the polyploids. *NIA* sequences in polyploid species with C and D genomes (in clade 3) are well differentiated from A and B genomes in clades 1 + 2 and 4, but are not well differentiated from each other in clade 3. Thus, all allopolyploid species have alleles in different clades according to their genome composition with one type in common, genome A.

Species from series *Longipedicellata* were postulated to have AB genomes (Table 1), a result supported here as their alleles resolve simultaneously in both clade 1 + 2 (genome B) and 4b (genome A). Matsubayashi (1955, 1991) indicated that series *Longipedicellata* possessed the A genome but the exact species contributor was not identified. Plastid DNA sequence data (Spooner and Sytsma 1992; Castillo and Spooner 1997; Spooner and Castillo 1997) suggested that their maternal donor is in clade 4. Based on both results, we conclude that the maternal donor had genome A and the paternal donor had genome B. A possible paternal species genome donor is the Mexican diploid *S. ehrenbergii,* given that it is their closest relative in the phylogenetic tree. Mexican diploid *S. verrucosum* seems to be the maternal species genome donor because all alleles from this series resolve with it and it occurs in the geographic area where this hybridization may have occurred. Further molecular and cytogenetic analyses are needed to confirm these results.

The tetraploid species from series *Conicibaccata* (putative AC genomes; Hawkes 1990, Matsubayashi 1991), have alleles in clades 3 and 4a. Their alleles in clade 4a resolve with the only diploid species from this series included in this study, *S. violaceimarmoratum*, genome A. The other alleles are in clade 3 which contains the diploids S. *albornozii*, *S. andreanum*, *S. chomatophilum,* and *S. immite* and the tetraploids *S. paucijugum* and *S. tuquerrense*. Matsubayashi (1991) designated *S. tuquerrense* as genome A^pP. The relationship among these four genomes (A, A^p, C, and P) is not clear. The genome A^p or P could be C or a slight variant of it. But, the data clearly show that these tetraploid species of series *Conicibaccata* contain two distinct and divergent alleles of *NIA*. Moreover, *NIA* data showed that for the tetraploid species in series *Conicibaccata* the maternal genome donor is genome A (clade 4a) and the paternal donor is in clade 3 with genome A^p, P, or C. Further analyses using more species and nuclear orthologous genes are needed to settle genomic donors more precisely.

All species of the Iopetalum group are hexaploids. They were postulated to have $ABB^{(1–4)}$ genomes by Marks (1955), A_1A_4(B,C,D) by Hawkes (1990), AB^sB^d by Irikura (1976), A_1A_4B by Ramanna and Hermsen (1979) or $ADD^{a,b,d,g}$ by

Matsubayashi (1991) (Table 1). *NIA* data support *S. schenckii* and *S. iopetalum* to have genome A (clade 4b), genome B (clade 1 + 2) and A^P, C, D, or P (clade 3). Their maternal genome donor is supported as *S. verrucosum* (genome A). The identity of the paternal donor remains unknown, but likely comes from B genome or the unidentified A^P, C, D, or P genomes.

Regarding the Acaulia group, all alleles of *S. acaule* resolve with two of the six alleles of *S. demissum* in clade 4a (genome A). The other four alleles of *S. demissum* are in clade 4b (also genome A) and resolve with *S. verrucosum* and alleles from series *Longipedicellata*. Results confirm the close relationship between *S. acaule* and *S. demissum* shown by Spooner et al. (1995), Kardolus et al. (1998), Kardolus (1999), and Nakagawa and Hosaka (2002). This suggests that *S. demissum* possibly is composed of two types of genome A, which are only slightly different.

Hawkes (1990) stated that *S. acaule* is allotetraploid and that one of its parents could come from series *Megistacroloba* and another from series *Tuberosa*. Matsubayashi (1991) gave *S. acaule* a genome formula of AA^a suggesting that it is a segmental allotetraploid with close homology to diploid species from series *Commersoniana* and *Tuberosa* but less homology with the genome of *S. bulbocastanum*, and had at least one genome in common with series *Longipedicellata* species (here we recognize as A). *NIA* data support the close relationship between *S. acaule* and diploid species from series *Megistacroloba* and *Tuberosa* as postulated by Hawkes (1990).

Polyploid species from series *Tuberosa* (*S. oplocense*), series *Piurana* (*S. tuquerrense*), and *S. paucijugum* (placed in series *Conicibaccata* by Hawkes [1990] but supported as part of the Piurana clade by Castillo and Spooner [1997]) have alleles in only one clade. *Solanum oplocense* alleles are in clade 4b, thus showing only one genome composition (genome A), and *S. paucijugum* and *S. tuquerrense* are in clade 3.

In summary, *NIA* sequence data provide strong evidence that polyploid species in series *Conicibaccata*, series *Longipedicellata*, and in the Iopetalum group are derived from allopolyploidization through hybridization between different ancestral species (different genomes). These data also show evidence that *S. verrucosum* is the maternal genome donor for polyploid species in the Iopetalum Group, confirming Hawkes' hypotheses. Furthermore, it provides evidence that *S. verrucosum* is the maternal genome donor for species in series *Longipedicellata*. On the other hand, it also suggests that polyploids species from series *Tuberosa* and *Piurana* that were included in this study have allopolyploid origins, but have the same genome A coming from different parents.

This analysis also gave strong support for the hypothesis of a North and Central American B-genome origin of the tuber-bearing members of *Solanum* section *Petota*, with eventual hybridization leading to allopolyploidy and giving rise to the series *Longipedicellata,* the Iopetalum Group and *S. demissum*. This hypothesis is supported by the monophyletic nature of B genome species (North and Central American diploid species) and the grouping of some alleles of series *Longipedicellata,* the Iopetalum Group and *Demissa* with *S. verrucosum*, which is the only Mexican diploid species with genome A.

This work answered many evolutionary questions about species in section *Petota*, and revealed relationships among A, B, C, D, E, and L genomes, but it also raised new questions. Future studies could use four new approaches to address the limitations of the current study. First, more genes could be used to broaden the genome coverage, since one single-copy gene covers only the evolution of a portion of one chromosome. Second, the use of more diploid species could investigate other genome contributors to allopolyploid species, especially in addressing the question of the origin of genomes C and D, which are still uncertain and cannot be inferred with the present data. Third, more accessions of each allopolyploid species could address the question of its single or multiple origins. Fourth, GISH techniques could be used to complement this study, labeling with different fluorescence the genomic DNA of the species that we identified as possible progenitors of the polyploids and hybridizing them with the genomic DNA of their polyploids, as has been initiated by Pendinen et al. (2008). The same strategy can be used to infer phylogenetic relationships among cultivated potatoes which, when added to the information generated here, will contribute to the design of crossing strategies capable of incorporating wild species into the cultivated potatoes.

Acknowledgments. We thank Ph.D. committee members Cecile Ané, Dennis Halterman, Shelley Jansky, and Kenneth Sytsma; and also David Baum for advice and review, two anonymous reviewers and Associate Editor Mark Simmons for invaluable comments and criticism, and the staff of the US Potato Genebank for germplasm. This paper represents partial fulfillment of a Ph.D. Degree for FR in Plant Breeding and Plant Genetics at the University of Wisconsin-Madison; this research was supported by the USDA and by NSF DEB 0316614 to David Spooner.

Literature Cited

Ames, M., A. Salas, and D.M. Spooner. 2008. A morphometric study of species boundaries of the wild potato *Solanum* series *Piurana* (Solanaceae) and putatively related species from seven other series in *Solanum* sect. *Petota. Systematic Botany* 33: 566–578.

Castillo, R. and D.M. Spooner. 1997. Phylogenetic relationships of wild potatoes, *Solanum* series *Conicibaccata* (sect. *Petota). Systematic Botany* 22: 45–83.

Correll, D.S. 1962. *The potato and its wild relatives.* Contributions from Texas Research Foundation: Botanical Studies vol. 4. Renner, Texas: Texas Research Foundation.

Cronn, R.C., R.L. Small, and J.F. Wendel. 1999. Duplicated genes evolve independently after polyploid formation in cotton. *Proceedings of the National Academy of Sciences USA* 96: 14406–14411.

Cronn, R.C., R.L. Small, T. Haselkorn, and J.F. Wendel. 2002. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany* 89: 707–725.

Douglas, N. 2007. A nuclear phylogeny of *Boerhavia* (Nyctaginaceae) suggests that expansion of North American deserts has facilitated multiple recent range expansions from the Gulf of California region. [Abstract] Chicago: Botany and Plant Biology 2007.

Doyle, J.J. and J.L. Doyle. 1999. Nuclear protein-coding genes in phylogeny reconstruction and homology assessment: some examples from Leguminosae. Pp. 229–254 in *Molecular systematics and plant evolution*, ed. R.J. Gornall. London: Taylor and Francis.

Den Nijs, T.P.M. and S.J. Peloquin. 1977. 2*n* gametes in potato species and their function in sexual polyploidization. *Euphytica* 26: 585–600.

Farris, J.S. 1989. The retention index and the rescaled consistency index. *Cladistics* 5: 417–419.

Felsenstein, J. 1973. Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on discrete characters. *Systematic Zoology* 22: 240–249.

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

Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics* 22: 521–565.

Gavrilenko, T. 2007. Potato Cytogenetics. Pp. 203–216 in *Potato Biology and Biotechnology: Advances and Perspectives*, ed. D. Vreugdenhil. Amsterdam: Elsevier B. V.

Gavrilenko, T., R. Thieme, and V.-M. Rokka. 2001. Cytogenetic analysis of *Lycopersicon esculentum* (+) *Solanum etuberosum* somatic hybrids and

their androgenetic regenerants. *Theoretical and Applied Genetics* 103: 231–239.

Graham, K.M. and L.A. Dionne. 1961. Crossability relationships of certain diploid Mexican *Solanum* species. *Canadian Journal of Genetics and Cytology* 3: 121–127.

Gu, X., X.-Y. Fu, and W.-H. Li. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Molecular Biology and Evolution* 12: 546–557.

Hanneman, R.E. Jr. 1994. Assignment of Endosperm Balance Numbers to the tuber-bearing Solanums and their close non-tuber-bearing relatives. *Euphytica* 74: 19–25.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating the human-ape split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22: 160–174.

Hawkes, J.G. 1958. Taxonomy, cytology and crossability. Pp. 1–43 in *Kartoffel*, ed. 2, vol. 3, eds. H. Kappert, and W. Rudorf. Berlin: Paul Parey.

Hawkes, J.G. 1990. *The potato. Evolution, biodiversity and genetic resources*. Washington: Belhaven Press.

Howarth, D.G. and D.A. Baum. 2002. Phylogenetic utility of a nuclear intron from nitrate reductase for the study of closely related plant species. *Molecular Phylogenetics and Evolution* 23: 525–528.

Huelsenbeck, J.P. and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics (Oxford, England)* 17: 754–755.

International Potato Center (CIP). 1999. *Molecular Biology Laboratory Protocols: Plant Genotyping*, eds. M. Ghislain, D. Zhang, and M.R. Herrera. Crop Improvement and Genetic Resources Department Training Manual. 2nd Edition. Lima: CIP.

Irikura, Y. 1976. Cytogenetic studies on the haploid plants of tuber-bearing *Solanum* species: II. Cytogenetic investigations on haploid plants and interspecific hybrids by utilizing haploidy. (In Japanese, with English summary.). Research Bulletin, Hokkaido. *National Agricultural Research Station* 115: 1–80.

Johnston, S.A. and R.E. Hanneman Jr. 1980. Support of the Endosperm Balance Number hypothesis using some tuber-bearing *Solanum* species. *American Journal of Potato Research* 57: 7–14.

Johnston, S.A. and R.E. Hanneman Jr. 1982. Manipulations of the Endosperm Balance Number hypothesis utilizing some tuber-bearing *Solanum* species. *Science* 217: 446–448.

Johnston, S.A., T.P.M. den Nijs, S. J. Peloquin, and R. E. Hanneman Jr. 1980. The significance and genic balance to endosperm development in interspecific crosses. *Theoretical and Applied Genetics* 57: 5–9.

Jukes, T.H. and C.R. Cantor. 1969. Evolution of protein molecules. Pp. 21–132 in *Mammalian protein metabolism*, ed. N. H. Munro. New York: Academic Press.

Kardolus, J.P. 1999. Morphological variation within series *Acaulia* Juz. (*Solanum* sect. *Petota*). Pp. 257–274. in *Solanaceae IV: Advances in biology and utilization*, eds. M. Nee, D.E. Symon, R.N. Lester, and J.P. Jessop. Kew: Royal Botanic Gardens.

Kardolus, J.P., H.J. van Eck, and R.G. van den Berg. 1998. The potential of AFLPs in biosystematics: a first application in *Solanum* taxonomy (Solanaceae). *Plant Systematics and Evolution* 210: 87–103.

Kihara, H. 1944. Discovery of the DD-analyzer, one of the ancestors of *Triticum vulgare* (Japanese). *Agriculture and Horticulture (Tokyo)* 19: 13–14.

Kimura, M. 1980. A simple model for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16: 111–120.

Kluge, A.G. and J.S. Farris. 1969. Quantitative phyletics and the evolution of the anurans. *Systematic Zoology* 18: 1–32.

Lara-Cabrera, S.I. and D.M. Spooner. 2004. Taxonomy of Mexican diploid wild potato (*Solanum* sect. *Petota) species:* AFLP data. *Plant Systematics and Evolution* 248: 129–142.

Lara-Cabrera, S.I. and D.M. Spooner. 2005. Taxonomy of Mexican diploid wild potatoes: (*Solanum* sect. *Petota*) morphological and microsatellite data. Pp. 199–205 in *A Festschrift for William G. D'Arcy, the legacy of a taxonomist*, eds. R.C. Keating, V.C. Hollowell, and T.B. Croat. Monographs in Systematic Botany from the Missouri Botanical Garden vol. 104. St. Louis: Missouri Botanical Garden.

Li, J., S. Shoup, and Z. Chen. 2007. Phylogenetic relationships of diploid species of *Betula* (Betulaceae) inferred from DNA sequences of nuclear nitrate reductase. *Systematic Botany* 32: 357–365.

López, L.E. and J.G. Hawkes. 1991. Cytology and genome constitution of the tuber-bearing *Solanum* species in series *Conicibaccata*. Pp.327–346 in *Solanaceae III, taxonomy, chemistry and evolution*, eds. J.G. Hawkes, R.N. Lester, M. Nee, and N. Estrada-R. Kew: Royal Botanic Gardens.

Maddison, D.R. and W.P. Maddison. 2001. MacClade 4.03: Analysis of phylogeny and character evolution, v. 4.08. Sunderland: Sinauer Associates.

Marks, G.E. 1955. Cytogenetic studies in tuberous *Solanum* species: I. Genomic differentiation in the Group Demissa. *Journal of Genetics* 53: 262–269.

Marks, G.E., 1965. Cytogenetic studies in tuberous *Solanum* species: III. Species relationships in some South and Central American species. *The New Phytologist* 64: 293–306.

Marks, G.E. 1968. Structural hybridity in a tuberous *Solanum* hybrid. *Canadian Journal of Genetics and Cytology* 10: 18–23.

Masterson, J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* 264: 421–424.

Matsubayashi, M. 1955. Studies on the species differentiation in the section *Tuberarium* of *Solanum*. III. Behavior of meiotic chromosomes in F_1 hybrid between *S. longipedicellatum* and *S. schenckii*, in relation to its parent species. *Science Reports of the Hyogo University of Agriculture* 2: 25–31.

Matsubayashi, M. 1991. Phylogenetic relationships in the potato and its related species. Pp. 93–118 in *Chromosome engineering in plants: genetics, breeding, evolution, Part B*, eds. T. Tsuchiya, and P.K. Gupta. Amsterdam: Elsevier Science Publishers BV.

Nakagawa, K. and K. Hosaka. 2002. Species relationships between a wild tetraploid potato species, *Solanum acaule* Bitter, and its related species as revealed by RFLPs of chloroplast and nuclear DNA. *American Potato Journal Research* 79: 85–98.

Ochoa, C.M. 1999. *Las papas de sudamerica: Peru (Parte I)*. Lima: International Potato Center.

Ortiz, R. and M.K. Ehlenfeldt. 1992. The importance of Endosperm Balance Number in potato breeding and the evolution of tuber-bearing *Solanum* species. *Euphytica* 60: 105–113.

Pendinen, G., T. Gavrilenko, J. Jiang, and D.M. Spooner. 2008. Allopolyploid speciation of the tetraploid Mexican potato species *S. stoloniferum* and *S. hjertingii* revealed by genomic in situ hybridization. *Genome* 51: 714–72.

Perez, F., A. Menendez, P. Dehal, and C.F. Quiros. 1999. Genomic structural differentiation in *Solanum*: comparative mapping of the A- and E-genomes. *Theoretical and Applied Genetics* 98: 1183–1193.

Ramanna, M.S. and J.G.T. Hermsen. 1979. Genome relationships in tuber-bearing Solanums. Pp. 647–654 in *The biology and taxonomy of the Solanaceae*, eds. G. Hawkes, R.N. Lester, and A.D. Skelding. Linnean Society of London Ser. 7. London: Academic Press.

Ramanna, M.S. and J.G.T. Hermsen. 1981. Structural hybridity in the series *Etuberosa* of the genus *Solanum* and its bearing on crossability. *Euphytica* 30: 15–31.

Ramsey, J. and D.W. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 29: 467–501.

Rodríguez, A. and D.M. Spooner. 1997. Chloroplast DNA analysis of *Solanum bulbocastanum* and *S. cardiophyllum*, and evidence for the distinctiveness of *S. cardiophyllum* subsp. *eherenbergii* (sect. *Petota*). *Systematic Botany* 22: 31–43.

Ronquist, F. and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics (Oxford, England)* 19: 1572–1574.

Sang, T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology* 37: 121–147.

Sang, T. and D. Zhang. 1999. Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five *Paeonia* species based on *Adh* gene phylogenies. *Systematic Botany* 24: 148–163.

Sears, E.R. 1976. Genetic control of chromosome pairing in wheat. *Annual Review of Genetics* 10: 31–51.

Simmons, M.P. and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49: 369–381.

Small, R.L., R.C. Cronn, and J.F. Wendel. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17: 145–170.

Small, R.L., J.A. Ryburn, R.C. Cronn, T. Seelanan, and J.F. Wendel. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. *American Journal of Botany* 85: 1301–1315.

Soltis, D.E. and P.S. Soltis. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* 12: 243–273.

Soltis, D.E., P.S. Soltis, and J.A. Tate. 2003. Advances in the study of polyploidy since plant speciation. *The New Phytologist* 161: 173–191.

Soltis, P.S. and D.E. Soltis. 2000. The role of genetics and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences USA* 97: 7051–7057.

Spooner, D.M. and R. Castillo. 1997. Reexamination of series relationships of South Ameri can wild potatoes (Solanaceae: *Solanum* sect. *Petota):* evidence from choloroplast DNA restriction site variation. *American Journal of Botany* 84: 671–685.

Spooner, D.M. and A. Salas. 2006. Structure, biosystematics, and genetic resources. Pp. 1–39 in *Handbook of potato production, improvement, and post-harvest management*, eds. J. Gopal and S.M. Paul Khurana. Binghampton, New York: Haworth's Press, Inc.

Spooner, D.M. and K.J. Sytsma. 1992. Reexamination of series relationships of Mexican and Central American wild potatoes (*Solanum* sect. *Petota):* Evidence from chloroplast DNA restriction site variation. *Systematic Botany* 17: 432–448.

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.

Spooner, D.M., D. Fajardo, and G.J. Bryan. 2007. Species limits of *Solanum berthaultii* Hawkes and *S. tarijense* Hawkes and the implications for species boundaries in *Solanum* sect. *Petota. Taxon* 56: 987–999.

Spooner, D.M., F. Rodríguez, Z. Polgár, H.E. Ballard Jr., and S.H. Jansky. 2008. Genomic origins of potato polyploids: GBSSI gene sequencing data. *The Plant Genome, a supplement to Crop Science* 48(S1): S27–S36.

Spooner, D.M., R.G. van den Berg, and J.B. Bamberg. 1995. Examination of species boundaries of *Solanum* series *Demissa* and potentially related species in series *Acaulia* and series *Tuberosa* (sect. *Petota). Systematic Botany* 20: 295–314.

Spooner, D.M., R.G. van den Berg, A. Rodríguez, J. Bamberg, R.J. Hijmans, and S.I. Lara-Cabrera. 2004. Wild potatoes (*Solanum* section *Petota*) of North and Central America. *Systematic Botany Monographs* 68: 1–209.

Staden, R. 1996. The Staden sequence analysis package. *Molecular Biotechnology* 5: 233–241.

Stone, D., G. Walters, and J.-P. Ghogue. 2006. *Memecylon batekeanum*, a new species from Southeastern Gabon, and a note on the circumscription of *Memecylon* sect. *Mouririoidea* (Melastomataceae). *Novon* 16: 276–280.

Swofford, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4.0b3a PPC. Sunderland: Sinauer Associates.

Swofford, D.L., G.J. Olsen, P.J. Waddell, and D.M. Hillis. 1996. Phylogenetic inference. Pp. 407–514 in *Molecular Systematics*, eds. D.M. Hillis, C. Moritz, and B. Mable. Sunderland: Sinauer Associates.

Tank, D.C. and T. Sang. 2001. Phylogenetic utility of the glycerol-3-phosphate acyltransferase gene: evolution and implications in *Paeonia* (Paeoniaceae). *Molecular Phylogenetics and Evolution* 19: 421–429.

Tanksley, S.D., M.W. Ganal, J.P. Prince, M.C. de Vicente, M.W. Bonierbale, P. Broun, M. Fulton, J.J. Giovannnoni, S. Grandillo, G.B. Martin, R. Messegeur, J.C. Miller, L. Miller, A.H. Paterson, O. Pineda, M. S. Roder, R.A. Wing, W. Wu, and N.D. Young. 1992. High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132: 1141–1160.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* 17: 57–86.

Templeton, A.R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* 37: 221–244.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 24: 4876–4882.

Waddell, P.J. and D. Penny. 1996. Evolutionary rate of apes and humans from DNA sequences. Pp. 53–73 in *Handbook of symbolic evolution*, eds. A.J. Locke, and C.R. Peters. Oxford: Clarendon Press.

Wendel, J.F. 2000. Genome evolution in polyploids. *Plant Molecular Biology* 42: 225–249.

Wendel, J.F. and J.J. Doyle. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. Pp. 265–296 in *Molecular systematics of plants II: DNA sequencing*, eds. D.E. Soltis, P.S. Solis, and J.J. Doyle. Boston: Kluwer Academic Publishers.

Wolfe, K.H. 2001. Yesterday's polyploids and the mystery of diploidization. *Nature Reviews. Genetics* 2: 333–341.

Wolfe, K.H., W.-H. Li, and P.M. Sharp. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences USA* 84: 9054–9058.

Yang, Z. 1994. Estimating the pattern of nucleotide substitution. *Journal of Molecular Evolution* 39: 105–111.

Yang, Z. and B. Rannala. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo method. *Molecular Biology and Evolution* 14: 717–724.

Zwickl, D.J. 2006. *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion.* Ph.D. thesis. Austin: The University of Texas.